



Separating response-execution bias from decision bias: Arguments for an additional parameter in Ratcliff's diffusion model

Andreas Voss^{1*}, Jochen Voss² and Karl Christoph Klauer¹

¹Albert-Ludwigs-Universität Freiburg, Germany

²University of Leeds, UK

Diffusion model data analysis permits the disentangling of different processes underlying the effects of experimental manipulations. Estimates can be provided for the speed of information accumulation, for the amount of information used to draw conclusions, and for a decision bias. One parameter describes the duration of non-decisional processes including the duration of motor-response execution. In the default diffusion model, it is implicitly assumed that both responses are executed with the same speed. In some applications of the diffusion model, this assumption will be violated. This will lead to biased parameter estimates. Consequently, we suggest accounting explicitly for differences in the speed of response execution for both responses. Results from a simulation study illustrate that parameter estimates from the default model are biased if the speed of response execution differs between responses. A second simulation study shows that large trial numbers ($N > 1,000$) are needed to detect whether differences in response-execution times are based on different execution times.

1. Introduction

In this paper, we consider the diffusion model for fast binary decisions (Ratcliff, 1978). The model maps cognitive processes on different model parameters. Most important are the speed of information intake (*drift rate*) and the amount of information that is required to come to a decision (*threshold*). The model also captures biases in decisions. Such decision biases are mapped on the *starting-point* of the diffusion process. Biases in starting-point are often referred to as *response biases* (drawing on the analogous concept from signal detection theory; Macmillan & Creelman, 1991), reflecting the preference for one response over another. However, we consider the label '*response bias*' to be misleading, because the starting-point influences the decision process rather

*Correspondence should be addressed to Dr Andreas Voss, Social Psychology and Methodology, Albert-Ludwigs-Universität Freiburg, 79085 Freiburg, Germany (e-mail: voss@psychologie.uni-freiburg.de).

than response execution. Consequently, we refer to biases based on the starting-point as *decision biases*.

In Ratcliff's model, the total response time (RT) is composed of the decision time as described by the diffusion process, and a non-decision component that comprises processes of encoding and response execution. In previous applications of the diffusion model, it has always been – implicitly – assumed that non-decision components are equal for both alternative responses. However, this assumption may be violated in many psychological tasks. Consider, for example, research on response priming (e.g., Elsner & Hommel, 2001) or response repetition effects (e.g., Campbell & Proctor, 1993): in these domains, the main hypotheses posit that a primed or repeated response is *executed* faster compared to a non-primed or non-repeated response. In this paper, we will refer to such differences in the *speed of response execution* between alternative responses as a *response-execution bias* or *response tendency*. This bias is theoretically independent of the decision bias discussed above, although both may often be elicited by the same tasks and may therefore occur simultaneously.

As we will show, parameters estimated with the default diffusion model (i.e., without modelling response-execution biases) might be notably distorted if a response-execution bias is inherent in the data. To permit explicit mapping of bias, we propose to extend the diffusion model by including an additional parameter.

2. The diffusion model and predicted RT distributions

2.1. The rationale of the diffusion model

In this section, we will provide a short introduction to Ratcliff's diffusion model (e.g., Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004; Voss, Rothermund, & Voss, 2004). Reviews of recent developments in this area are provided by Ratcliff and McKoon (2008) and Wagenmakers (2009).

The diffusion model is used to describe fast binary decisions. According to the model, information is accumulated continuously until evidence is sufficiently clear. The information accumulation is characterized by a constant drift and random noise. Two thresholds (for the two alternative decisions) form a corridor in which the process runs until one of the thresholds is hit. Then the response is initiated.

The diffusion model in its simplest form (see Figure 1) is characterized by the drift rate, threshold separation, starting-point and RT constant.¹ We will first introduce this simple model before discussing the role of additional inter-trial variability parameters.

The average rate of information uptake is called the drift rate (v). Large drift rates result in fast decisions and few errors. The distance between thresholds (a) reflects the amount of information that is necessary to come to a conclusion. A large span between thresholds indicates a conservative decision style, that is, the decision is based on a large amount of information. In this case, the decision making is slow and few errors occur. A liberal decision style is represented by a small distance between thresholds.

In some cases, different amounts of information may be used for the alternative decision outcomes. This is reflected by the starting-point (z) of the diffusion process.

¹ Additionally, the diffusion process is characterized by the amount of noise. This is reflected by the diffusion constant, which operates as a scaling parameter. The diffusion constant is not estimated but can be freely chosen.

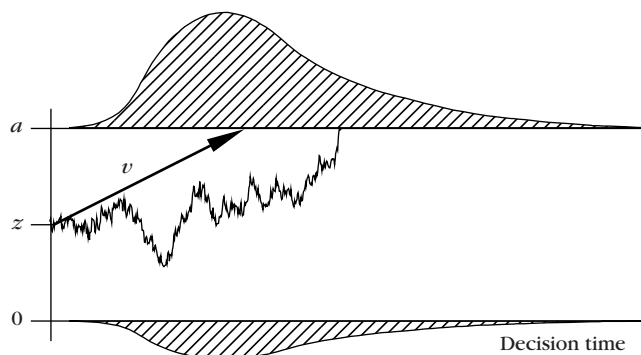


Figure 1. Ratcliff's diffusion model.

If the process starts above $a/2$, comparatively less information is sufficient to come to the decision outcome associated with the upper threshold (decision A), while more information supporting B is needed before this decision is initiated. If the starting-point is closer to the upper threshold, the mean RTs of the response connected to the upper threshold will be shorter than the mean RTs of the other response. Simultaneously, the portion of responses corresponding to the upper threshold will increase with increasing starting-point. In the simple diffusion model, any difference in mean RTs between lower and upper distributions is mapped on the z parameter.

The last parameter of the simple diffusion model is the RT constant (t_0). This parameter comprises all extra decision processes, that is, pre-decision components (e.g., stimulus encoding) and post-decision processes (e.g., response execution). RTs are modelled as the sum of decision times and RT constant. Therefore, t_0 contributes equally to correct and error RTs.

Ratcliff's complete diffusion model allows for inter-trial variability of drift, starting-point, and RT constant. These parameters influence the shape of the predicted RT distributions, and can account for small differences in the mean RTs of correct responses and errors (e.g., Ratcliff & Rouder, 1998): inter-trial variability of the starting-point can explain fast error responses, while inter-trial variability of the drift causes slow errors.

2.2. Applications of the diffusion model

The diffusion model has been applied successfully to a variety of tasks (see Spaniol, Madden, & Voss, 2006, for a review). One important field of application is the analysis of ageing effects on cognitive processes (e.g., Ratcliff, Thapar, & McKoon, 2006, 2007). In this domain, research focuses on the question whether slowing of RTs is based on slower information processing (drift), on slower response execution (RT constant), or on a more conservative decision style (threshold). An important finding is that age-related slowing in cognitive tasks is in large part due to conservative threshold settings and slower motor responses, while cognitive processing is often quite robust.

Other applications aim at improving the understanding of specific cognitive tasks. For example, Klauer, Voss, Schmitz, and Teige-Mocigemba (2007) analysed data from the implicit association test (Greenwald, McGhee, & Schwartz, 1998) with a diffusion model analysis: results of these analyses helped to clarify why different test blocks have different difficulties. In the same study, the inter-individual differences in parameter values were used successfully to predict explicit attitudes. This suggests that the

diffusion model may be used as a diagnostic tool (Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007).

There is a growing number of further applications of the diffusion model (see Wagenmakers, 2009). The model is likely to become even more popular, due to recent publications of different software tools for diffusion model analyses (*fast-dm*: Voss & Voss, 2007; *dmat*: Vandekerckhove & Tuerlinckx, 2007, 2008; *EZ-diffusion*: Grasman, Wagenmakers, & van der Maas, 2009; Wagenmakers, van der Maas, Dolan, & Grasman, 2008; Wagenmakers, van der Maas, & Grasman, 2007).

3. Mapping response tendencies: A second RT constant

The diffusion model as described above might be used to detect decision biases. Consider, for example, results by Voss, Rothermund, and Brandtstädter (2008): they showed within a motivated-perception design that participants were more liberal in the decision that an ambiguous stimulus is positive (i.e., a small amount of information is enough) and were more conservative in classifying a stimulus as negative (i.e., a high amount of information is used). In the diffusion model, this effect is reflected by a starting-point which is close to the ‘positive’ threshold.

The observed starting-point asymmetry in the data of Voss *et al.* (2008) is ultimately a consequence of shorter RTs for positive decisions (i.e., when a stimulus was classified as positive). In the diffusion model, differences in mean RTs between two alternative responses always map on the starting-point parameter (z), indicating different response criteria (Ratcliff, 1985).² However, different response criteria are not the only theoretically plausible source of differences in mean RTs between two alternative responses. Imagine, for example, a task in which stimuli have to be classified by pressing a key on the right- or left-hand side of a keyboard. Independent of the classification rules, a right-handed person may be faster with the ‘right’ response compared with the ‘left’ response. If this were mapped on the starting-point parameter, one would conclude that a more liberal criterion is associated with the corresponding decision. From a theoretical point of view, however, an effect of handedness is based on speed of response execution.

There are different psychological paradigms in which response facilitation or response inhibition are discussed as explanations of the observed effects (cf. affective priming: Fazio, Sanbonmatsu, Powell, & Kardes, 1986; flanker task: Eriksen, Coles, Morris, & O’Hara, 1985).³ In its default form, the diffusion model is unable to map response tendencies and it is therefore not possible to test contrasting approaches against each other (e.g., enhanced perception vs. speeded response execution). Notice that both approaches predict differences in latencies between responses. Therefore, we suggest applying different RT constants for the two responses: in the extended model, the RT constant $t_{0,\text{upper}}$ is added to the decision time if the process terminates at the upper threshold, and $t_{0,\text{lower}}$ is added for all responses at the lower threshold. We will recode these parameters as $t_0 = 0.5(t_{0,\text{upper}} + t_{0,\text{lower}})$ and $d = t_{0,\text{lower}} - t_{0,\text{upper}}$.

² If data are analysed in terms of correct versus erroneous responses (rather than response ‘A’ vs. response ‘B’) differences in mean RTs can be captured to a certain degree by the inter-trial variability of starting-point or drift as well (Ratcliff & Rouder, 1998). However, inter-trial variability cannot explain that, for example, response ‘A’ is always faster than response ‘B’, if this pattern holds for stimuli of type A (response ‘A’ is correct) and of type B (response ‘B’ is correct).

³ However, the response-execution bias can only explain effects in RT data and not in accuracy data. Since there are accuracy effects in affective priming, response-execution biases cannot explain the total effect.

Hence, t_0 reflects the mean RT constant, and d reflects the response-execution bias. In the extended model, differences in mean RTs for the two responses can be mapped by asymmetries either in criteria (*decision bias*) or in response-execution times (*response-execution bias*).

4. Simulation study I: Recovering parameter values

The aim of this study is to test whether parameter values can be recovered reliably with the extended model. For this purpose, data sets were simulated from four parameter sets characterized by (a) no bias in starting-point or RT constants, (b) a bias in starting-point but not in RT constants, (c) a bias in RT constants but not in starting-point, and (d) both types of bias. From these simulated data sets, parameters were re-estimated using models with different restrictions.

4.1. Simulating data sets

Table 1 shows the parameter values used for the simulation of data sets. Parameter set 1 represents a symmetric model without response-execution bias or starting-point bias. In parameter sets 2 and 4, the starting-point z lies above $a/2$, indicating a decision bias in favour of response A (upper threshold). In parameter sets 3 and 4, a response-execution bias is simulated: response A is executed 50 ms faster compared with response B.⁴

From each of the four parameter sets of Table 1, 100 data sets with 500 trials each (half of the trials using v_1 and half using v_2) were simulated using an adapted version of the *construct-samples* tool of *fast-dm release 29* (Voss & Voss, 2007; precision = 5.0).

4.2. Re-estimating parameter values

From each of the 400 simulated data sets, parameters were re-estimated using an adaptation of *fast-dm-29* (precision = 4.0).⁵ For each data set parameters for four models with different restrictions were fitted: in the 'fixed' model, neither decision bias nor response-execution bias were allowed ($z = a/2$ and $d = 0$).⁶ For the second model, only a response-execution bias was prevented ($d = 0$). The third model did not allow for a decision bias ($z = a/2$). Finally, the 'complete' model estimated both z and d . In summary, each of the 4 (parameter sets) \times 100 data sets was analysed with 4

⁴ Note that both biases operate in the same direction: in both cases, response B is delayed. Therefore, we address the case in which mean RTs differ between responses. Theoretically, it is possible that a response-execution bias and an opposing decision bias cancel out each other. However, we assume that this theoretical possibility will rarely occur in psychological studies.

⁵ Inclusion of a second RT constant in *fast-dm* is a relatively small change and will be included in one of the next releases of the program. A problem arises if the Kolmogorov-Smirnov (KS) statistic (i.e., the maximum vertical distance between predicted and observed cumulative distribution function) is used as an optimization criterion: in the extended model, changing a single RT constant influences only one of the two predicted RT distributions (upper or lower). Thus, when the point of maximum distance and the changed RT constant correspond to opposite sides, the KS statistic will be unaffected by changes of one of the RT constants within a certain range and minimizing the vertical distance no longer results in unique parameter values. To solve this problem, the sum of the maximum vertical distances of the upper and the lower distribution was used as minimization criterion.

⁶ Starting-point (z) was fixed to $0.5a$ rather than to an absolute value to force a symmetric model. The figures report absolute values of z rather than z/a and, consequently, there is still variation of starting values across sets of recovered parameters from the 'fixed models' (see Figures 2 and 3), because of the variance in the estimated threshold separation parameters (a).

Table 1. Parameter values used for the simulation of data sets (Study 1)

Parameter	Parameter set			
	1	2	3	4
Threshold separation	$a = 1.00$	$a = 1.00$	$a = 1.00$	$a = 1.00$
Starting-point	$z = 0.50$	$z = 0.65$	$z = 0.50$	$z = 0.65$
Drift rate 1	$v_1 = 1.00$	$v_1 = 1.00$	$v_1 = 1.00$	$v_1 = 1.00$
Drift rate 2	$v_2 = -1.00$	$v_2 = -1.00$	$v_2 = -1.00$	$v_2 = -1.00$
RT constant	$t_0 = 0.400$	$t_0 = 0.400$	$t_0 = 0.400$	$t_0 = 0.400$
RT bias	$d = 0.000$	$d = 0.000$	$d = 0.050$	$d = 0.050$

Note. Moderate variability was assumed ($s_z = 0.2$; $s_v = 0.5$; $s_{t_0} = 0.1$). A diffusion constant of $s = 1$ was adopted. If other diffusion constants are of interest parameters a , z , and v can be transformed by multiplication with the new value of s .

(parameter constraints) different estimation procedures. The different restrictions were used to test the effect of ‘wrong’ restrictions (e.g., fixing $d = 0$, when $d = 0.05$ in the original data set). Additionally, it was thus possible to test whether the ‘true’ source of RT differences between both responses could be identified.

Means and standard deviations for the parameter estimates are presented in the Figures 2 and 3. Results are shown for parameter sets without biases (Figure 2a), with a decision bias on the starting-point (Figure 2b), with a response-execution bias (Figure 3a), and with both types of bias (Figure 3b). The estimates for the variability parameters are presented in the Appendix.

4.3. Parameter set 1: No bias

Figure 2a shows the mean estimates for data from parameter set 1. All parameters were estimated accurately by all four models. This was expected, because no restrictions conflicted with the original parameter values.

4.4. Parameter set 2: Decision bias

In parameter set 2, the starting-point was biased towards the upper threshold. Therefore, restrictions in models 1 and 3 ($z = a/2$) do not allow accurate recovery of the original values. As evident in Figure 2b, an erroneous restriction of the starting-point leads to a biased estimation of drift rates, that is, the low starting-point value results in an inflation of the estimates for the drift. If z was restricted, but the response-execution bias parameter d was estimated (model 3), an erroneous response-execution bias of 50 ms was estimated. Note that this value matches the value of $d = 0.05$ that was chosen for the model including a response-execution bias. This agreement suggests that both biases have comparable effects on the RT distributions. In the complete model (model 4), the mean estimates for the starting-point and response-execution bias were close to the correct values.

4.5. Parameter set 3: Response-execution bias

Parameter set 3 represents the critical case of data sets with inherent response tendencies ($d = 0.05$). Figure 3a depicts the results of the parameter estimation procedure. In the standard analysis ($d = 0$), the starting-point is notably overestimated,

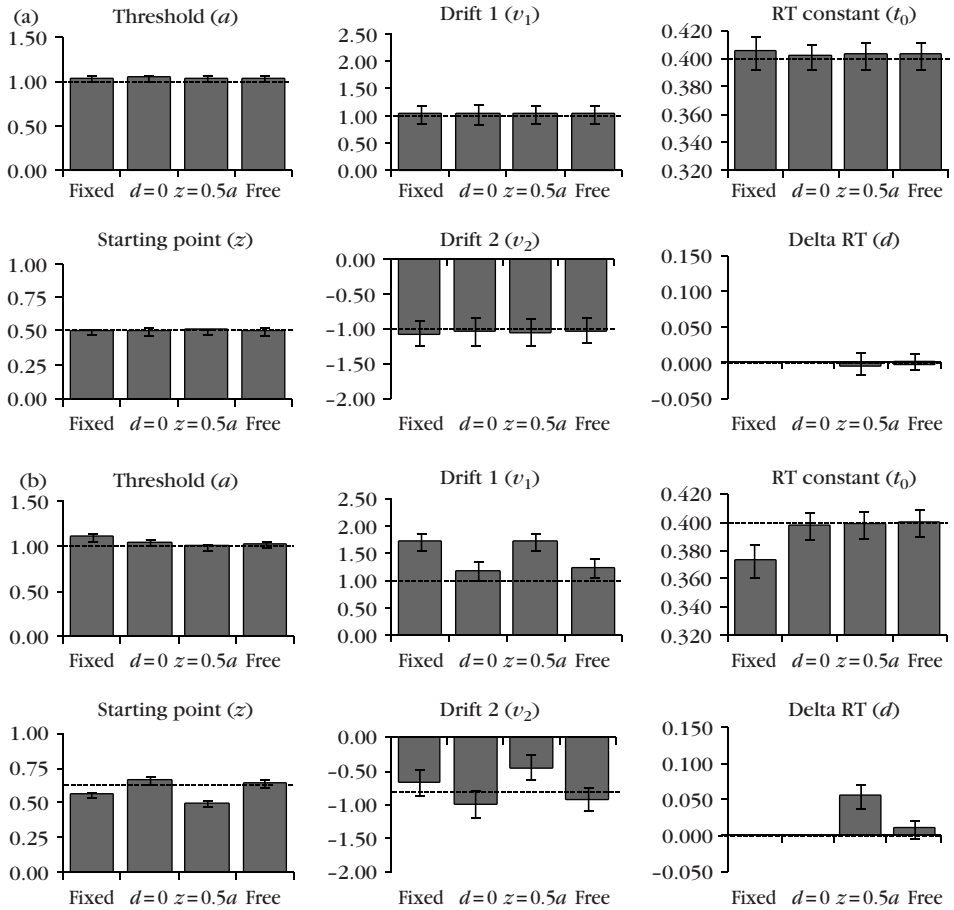


Figure 2. Re-estimated parameter values from Study 1. (a) Data from parameter set 1 (neither decision bias nor response-execution bias). (b) Data from parameter set 2 (decision bias only). Original (correct) values are indicated by the slashed lines. Error bars indicate ± 1 SD.

while the estimates for the drift rate fall below the true values. The models accounting for d yield much more precise results. However, if z is not fixed to the correct value, then the bias is dispersed to a certain degree between z and d , that is, results show a bias in z (although there was none in the data), while the estimate for the response-execution bias falls short of the true value of 50 ms.

4.6. Parameter set 4: Decision bias and response-execution bias

Figure 3b depicts the results for the data sets with decision bias and response-execution bias. The restricted models show the compensatory effects of both bias parameters: in the model without response-execution bias ($d = 0$), the decision bias is overestimated ($z > 0.65$); in the model without decision bias ($z = a/2$), the response-execution bias is overestimated ($d > 0.05$). Estimates are most accurate in the complete model accounting for both types of biases. However, in the complete model again the estimates for the response-execution bias are too low.

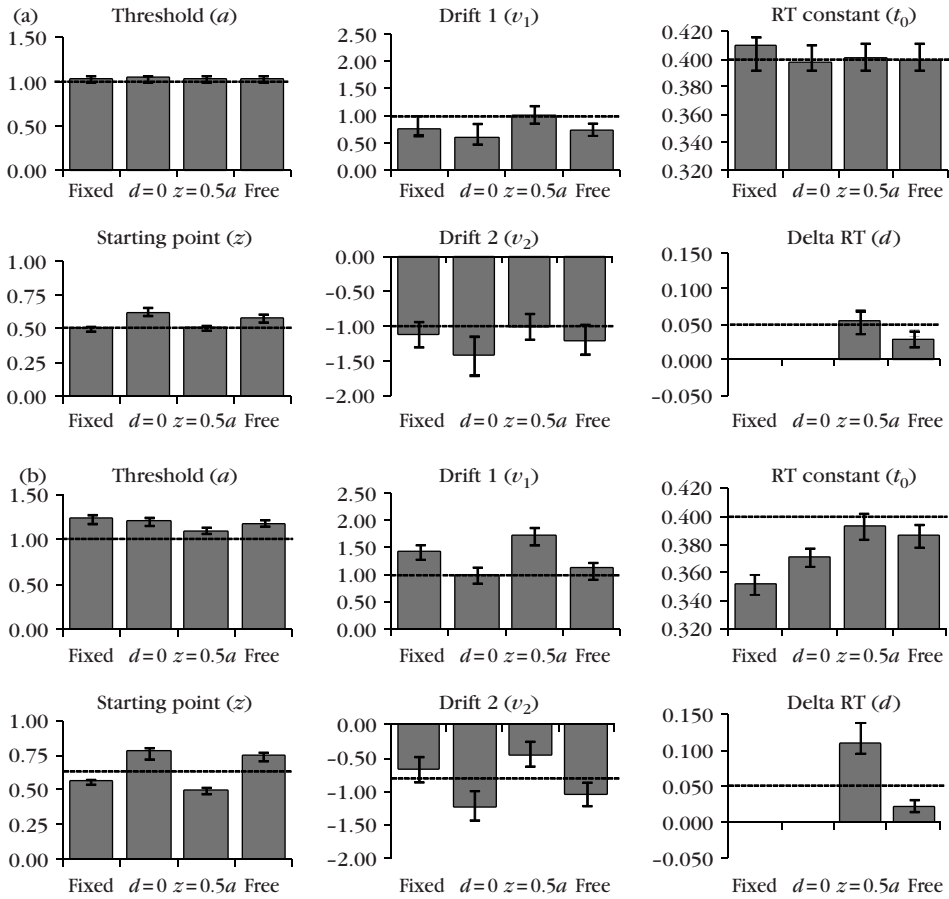


Figure 3. Re-estimated parameter values from Study 1. (a) Data from parameter set 3 (response-execution bias only). (b) Data from parameter set 4 (decision bias and response-execution bias). Original (correct) values are indicated by the slashed lines. Error bars indicate ± 1 SD.

4.7. Model fit

The optimization criterion of *fast-dm*, the Kolmogorov-Smirnov (KS) statistic, is reported here as an indicator of model fit (Voss et al., 2004). Separate KS values were computed for the two trial types (v_1 vs. v_2). Table 2 shows the mean KS values across both trial types. Low values indicate a good model fit. The model without biases is taken as a baseline condition here. Accordingly, a KS value below .030 is regarded as a good model fit. As evident in Table 2, the fit index exceeds this value in each case in which a model restriction violated original values from parameter sets.

4.8. Discussion

Study 1 was conducted to demonstrate that incorrect restrictions in the diffusion model result in notably biased parameter estimates. Specifically, this was illustrated for the case of response-execution biases, that is, for a situation in which one

Table 2. Model fit indices of Study 1

Restrictions	Parameter set							
	1		2		3		4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Fixed	0.029	0.005	0.056 ^a	0.008	0.046 ^a	0.007	0.078 ^a	0.007
$d = 0$	0.028	0.004	0.028	0.004	0.034 ^a	0.005	0.032 ^a	0.005
$z = 0.5$	0.026	0.004	0.036 ^a	0.006	0.027	0.004	0.037 ^a	0.007
Complete	0.027	0.003	0.027	0.003	0.029	0.004	0.030	0.005

Note. The table shows the KS statistic, which is the maximum vertical distance between predicted and empirical RT distributions. Fixed: $d = 0$ and $z = a/2$; complete: no restrictions of d or z .

^a Restrictions conflicted with the original values.

response is executed faster than the alternative response. Default diffusion model analyses so far did not account for differences in speed of response execution.

A response-execution bias that is not accounted for by the model is mapped on a (congruent) bias on starting-point and – simultaneously – on an (incongruent) bias on the drift (see Ratcliff & McKoon, 2008, for a discussion of the criterion of drift rates). For example, when the response linked to the upper threshold is executed faster, and this is not mapped in the model, the starting-point will be overestimated and the drift will be underestimated.

The results also revealed difficulties with estimating d and z at the same time, at least in data sets with 500 trials as used in the present studies. If nonetheless both kinds of bias are to be investigated within one model, we recommend the use of samples with substantially larger trial numbers. If it is of interest, however, which of the two biases accounts better for the data, a comparison of models incorporating either response-execution bias or starting-point bias might be successful: results from Study 1 suggest that – on average – models accounting for the ‘correct’ type of bias have a better fit. We investigate this issue more closely in Study 2.

5. Simulation study 2: Identifying response-execution biases in individual data sets

Study 1 showed that it can be challenging to get reliable estimates for z and d simultaneously. Including both parameters makes the model so flexible that the solution of a parameter search is not robust. Therefore, we recommend using models that allow for either starting-point asymmetries or response-execution biases, but not for both kinds of bias. Study 2 systematically examines how many trials are needed to reliably decide which of these two models better accounts for an individual data set. For this purpose, 1,000 data sets were simulated from parameter sets with starting-point asymmetries (but no response-execution bias) and from parameter sets with response-execution bias (but no bias on starting-point) for four different sample sizes (500, 1000, 2000, and 5000 trials per data set). Subsequently, fit indices were compared between these two models. Model fit was assessed using the KS statistic.

5.1. Simulating data sets and re-estimating parameter values

In a first step, 1,000 parameter sets were generated by randomly drawing values for a , v_1 , v_2 , t_0 , s_z , s_v and s_{t0} from uniform distributions (see Table 3). From each of these parameter sets two models were derived: for the *starting-point bias model*, z was calculated as $0.65a$, and d was set to 0; for the *response-execution model*, d was set to 0.05 s, and z was set to $0.5a$. For all parameter sets and both models, data sets with 500, 1000, 2000, and 5000 trials were simulated using a modified version of the *construct-samples* (precision = 4) tool of *fast-dm*. In each case, two stimulus types were simulated; that is, v_1 and v_2 , respectively, were used for half of the trials of each data set.

Table 3. Ranges for parameter values for the simulated data sets from Study 2

Parameter	Minimum	Maximum
a	0.80	1.30
v_1	1.00	4.00
v_2	-4.00	-1.00
t_0	0.20	0.40
s_z	0.00	$0.25a$
s_v	0.00	1.00
s_{t0}	0.00	0.10

Note. The decision bias parameter was set to $z = 0.5a$ or $0.65a$, respectively. The response-execution bias parameter was set to $d = 0.05$ or 0.00 , respectively. A diffusion constant of $s = 1$ was adopted. If other diffusion constants are of interest parameters a , z , v , s_z and s_v can be transformed by multiplication with the new value of s .

From each of the resulting 8,000 data sets two parameter sets were re-estimated using an adapted version of *fast-dm* (precision = 3), whereby either z was fixed to $0.5a$ (*response-execution model*) or d was fixed to 0 (*starting-point bias model*).

5.2. Results

Each data set was analysed with regard to whether the *response-execution model* or *starting-point bias model* fitted the data better in terms of the KS statistic. Figure 4 shows the frequency of false classifications as a function of the original model and sample size. As expected, the probability of false model selection declines with increasing number of trials. Results indicate that a response-execution bias of $d = 0.05$ is more often correctly classified compared with a starting-point bias of $z = 0.65a$. In the former case, approximately 900 trials are sufficient for 95% correct classifications, while in the latter case about 1,750 trials are needed to achieve the same accuracy.

5.3. Discussion

Study 2 demonstrates that it is possible to decide for an individual data set whether a bias in mean RTs is based on differences in response execution (response-execution bias) or on differences in information accumulation (starting-point bias) with reasonable accuracy. Because the diffusion model with its many parameters is very flexible in fitting RT distributions with various shapes (in particular, if the variability parameters are included), many trials are needed to allow for reliable separation of both models: results

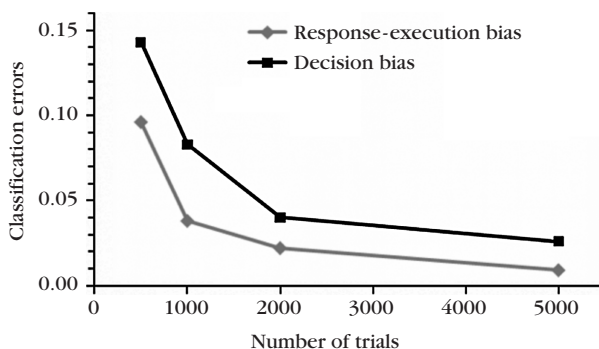


Figure 4. Percentage of data sets for which the estimation procedure with ‘wrong’ restriction (i.e., incompatible with the original parameter set) resulted in a better fit than the models with ‘correct’ restrictions (Study 2). The grey line represents data sets simulated from the response-execution bias model; the black line shows results from the decision-bias model.

of the present experiment indicate that at least 1,000 trials are needed, even in the case of large effects (i.e., about 50 ms difference between mean latencies for the two responses). To get an estimate for the number of trials required when effects are notably smaller, similar simulation studies can be repeated for other parameter sets.

6. Model mimicry: Locating the effect of wrong parameter specifications in the RT distributions

As shown above, it is difficult to separate the response-execution bias model from the starting-point bias model. That is, even if there are invalid model specifications (e.g., $d = 0$), the ‘wrong’ model can mimic the empirical RT distribution. In this section, we further analyse where exactly the predicted RT distribution differs from the empirical distributions in case of wrong specifications.

For this purpose, two ‘non-random’ data sets were simulated⁷ for parameter sets using the means of the parameter distributions of Study 2 ($a = 1.05$; $v_1 = 2.50$; $v_2 = -2.50$; $t_0 = 0.300$; $s_z = 0.13$; $s_v = 0.50$; $s_{t0} = 0.05$), one set with $d = 0.05$ and $z = 0.50a$ and one set with $d = 0$ and $z = 0.65a$ ($N = 10,000$ trials, half of these based on v_1 and half based on v_2). Parameters were re-estimated using ‘wrong’ restrictions, that is, fixing $d = 0$ for the data generated based on the response-execution bias model and fixing $z = 0.5a$ for the data generated on basis of the starting-point bias model.

Distributions of the data were compared with predicted distributions from the recovered parameters. For this purpose, so-called combined cumulative distribution functions (cCDF; for more detailed information, see Voss & Voss, 2008; Voss *et al.*, 2004) were used. The advantage of this kind of diagram is that two (predicted) distributions are mapped by one graph, and can thus be compared easily with the corresponding graphs given by the data. The cCDF graphs combine the RT distributions for the two responses by mapping the times from the lower distribution of the diffusion model on

⁷The data points were located on a grid of regularly spaced quantiles of the predicted distributions.

the negative half-axis (all RTs from the lower distribution get a negative sign). The lower distribution is mirror-imaged, that is, the tail of the distribution is depicted on the far left of the graph and the starting edge of the RT distribution appears close to the origin (Figure 5). The height of the plateau around zero indicates the total mass of the lower distribution.

Figure 5 shows the cDFs for predicted RT distributions of original parameters and of parameter estimates from the models with restriction that are incompatible with the fitted data. Despite the wrong restrictions, deviations are rather small: in particular, the leading edges of the RT distributions are recovered very well. The strongest deviations are evident at the tails of the RT distributions: for the response-execution bias model, the predicted cDF of stimulus 1 approaches 0 on the left-hand side and 1 on the right-hand side more slowly than the cDF of the original data. That is, the predicted number of slow responses exceeds the actual number. The opposite is true for stimulus 2: here, the predicted cDF approaches 0 and 1 faster than the original cDF, that is, the tails of the RT distributions are underestimated. The pattern of results is reversed for the starting-point bias model: here, the tails of distributions are underestimated for stimulus 1 and overestimated for stimulus 2.

In summary, these graphs suggest that the tails of RT distributions might be especially indicative of these kinds of biases: to make a decision, one has to consider separately stimuli corresponding to the faster response and stimuli corresponding to the slower response: if the percentage of slow responses is *overestimated* for stimuli requiring the faster response and the percentage of slow responses is *underestimated* for stimuli requiring the alternate response, this may be taken as an indicator of an unaccounted-for response-execution bias. The opposite pattern of results, that is, an

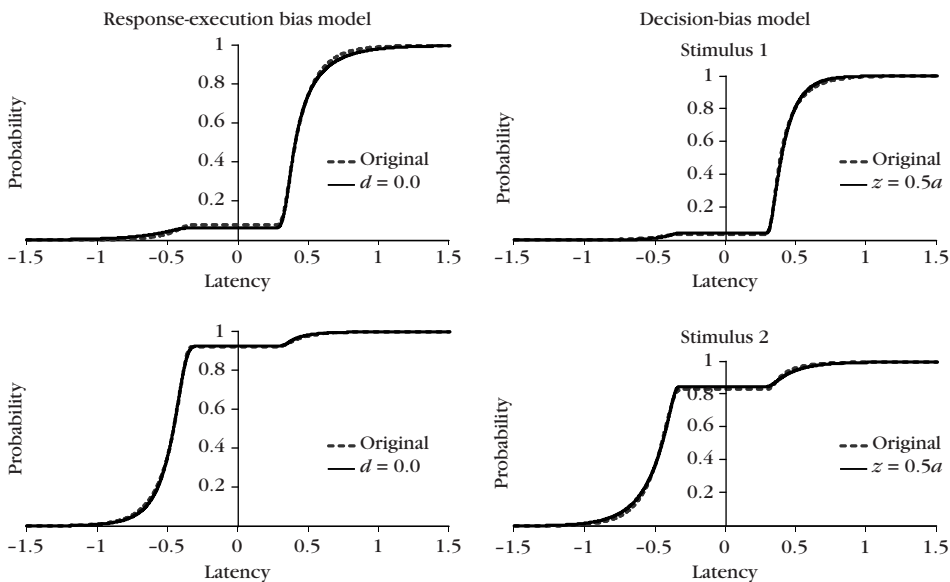


Figure 5. Model mimicry. The cDFs for the response-bias model (left) and for the decision-bias model (right). Original cDFs are demonstrated by the dotted lines. The solid black lines give the cDFs from parameters estimated with incorrect restrictions, that is, response-bias data is mimicked with a decision-bias model and vice versa.

underestimation of slow latencies for the stimuli with fast responses and vice versa, suggests that there might be a decision bias in the data.

However, we emphasize that the location of mismatch between data distributions and predicted distributions might strongly depend on the optimization criteria: the application of the KS statistic leads to a nearly exact recovery of the leading edges of the RT distributions of correct responses, which comes at the cost of a poor recovery of the last quantiles of the RT distributions, and – generally – of error distributions. This might be different if, for example, the chi-squared statistic were applied.

7. General discussion

In this paper, we introduced an extension to Ratcliff's (1978) diffusion model. In the default diffusion model, the assumption is implicitly made that there is no (systematic) difference between the speed of *execution* of both possible responses. The model's extension aims to allow mapping differences in speed of response execution for the two different responses. With estimation procedures using the default model (e.g., *fast-dm*: Voss & Voss, 2007, 2008; *dmat*: Vandekerckhove & Tuerlinckx, 2007, 2008), differences in response speed *between conditions* can be mapped if the t_0 parameter is allowed to vary between conditions. However, in this case, t_0 depends on *condition* and not on *response*, that is, within each condition again the same execution time is assumed for the two alternative responses 'A' and 'B'.

Study 1 investigates the effects of wrong model restrictions (especially the assumption that there is no response-execution bias) and checks whether it is possible to recover a response-execution bias accurately. Results indicate that outcomes of a diffusion model analysis are seriously affected if an existing response-execution bias is not incorporated in the model. Specifically, the estimate for the starting-point will be biased towards the threshold linked to the faster response, while the drift will be biased in the opposite direction. At the same time, our findings demonstrate that it might be challenging to estimate both d and z simultaneously. One strategy to deal with this problem is the comparison of model fit between models that account either for a starting-point bias or for a response-execution bias. Unfortunately, there is no exact way to test whether the inclusion of a new parameter improves the fit significantly. However, rules of thumb can be generated to decide which restrictions are sensible. If the fit between models differs only a little (e.g., by less than 1 standard deviation; see Table 2), the more parsimonious model should be chosen (if the number of parameters is equal, of course, the better fitting model is chosen). If the difference in fit is large (e.g., more than 1 standard deviation), the inclusion of the additional parameter seems to be justified. Of course, this heuristic is developed *post hoc* and needs to be tested with broader simulation studies, including other parameter values and trial numbers.

While the heuristic explicated above allows for comparison of model fit for a group of participants, it does not help to decide for an individual data set whether the decision-bias model or the response-execution model accounts for the data better. For this purpose in Study 2, we examined how many trials were needed to ensure that the true model resulted in a better model fit with a probability of $p = .95$. Results indicated that $N = 1,000$ to $2,000$ trials were needed to reliably detect the true model, even for large asymmetries as used in the present study. Consequently, for many psychological experiments, the collected amount of data will not suffice to detect a moderate response-execution bias. It might be helpful to compare fit

indices at a group level, as in Study 1. If data sets are even smaller ($N < 500$), the problem of discrimination between the alternate models obviously becomes even more severe.

7.1. Potential biases of parameter estimates from the default model

How can the biases of the parameter estimates from the default model be explained? To answer this question, one must examine the influence of the diffusion model parameter on the predicted RTs: first of all, it is important to note that the drift rate has parallel effects on predicted RTs for correct responses and errors: the larger the drift, the faster the mean RTs. As long as the starting-point is symmetrical ($z = 0.5a$) and all inter-trial variabilities are zero, the predicted RTs for errors and correct responses are exactly identical. There is, however, a strong influence of the starting-point on predicted RTs: if the starting-point moves upward, the predicted exit times for the upper threshold decrease, and times for the lower threshold increase. To put it the other way around, if there is an empirical difference in mean RTs, the default diffusion model will map this on the starting-point.

By moving the starting-point, differences in mean RTs of the upper and lower threshold can be accommodated. However, moving the starting-point upward increases the percentage of predicted responses for the upper threshold. If the asymmetric starting-point is an artefact, that is, if it is not based on real different threshold settings for the information accumulation, the number of predicted responses will be overestimated at the threshold towards which the starting-point has moved. In the multidimensional parameter search, this will be compensated for by an opposite effect on drift rate. In summary, if response 'B' is delayed, the starting-point will be biased towards 'A' and the drift will be biased towards 'B'.

7.2. Potential benefits of additional information from the extended model

In the above discussion, we concentrated on drawbacks for the validity of the parameter estimates (especially for z and v) if only one RT constant is adopted. However, this is not the only – and perhaps not the most important – reason to use the extended diffusion model. The most fundamental benefit of the extended model is that it allows for testing new empirical hypotheses. For any theory predicting response preparation, response facilitation, response inhibition, or response conflicts, the crucial tests are concerned with differences in the *speed of execution* of the alternative responses (but see footnote 3). For this purpose, different RT constants as implemented in the extended diffusion model are required.

Imagine, for example, the case of affective priming (Fazio *et al.*, 1986). There is a lively debate about whether this effect is equivalent to semantic priming (i.e., the prime pre-activates the target concept via spreading activation), or whether affective priming works similarly to response priming (i.e., the prime prepares a specific response). In the former case, affectively congruent priming should facilitate target perception (i.e., increase the drift), while in the latter case, a congruent prime should decrease the RT constant of the correct response (and increase the RT constant of the alternative response). With the extended model, it is possible to differentiate between these hypotheses empirically. Note, however, that response execution alone cannot explain the complete affective priming effect if effects are not limited to RTs but are also present in errors.

One caveat concerns the interpretation of the RT constant. The RT constants map all non-decision parts of information processing, that is, stimulus encoding as well as response execution. Therefore, it is theoretically possible that differences between the RT constants are based on differences in encoding times. However, in most experimental designs, stimuli do not differ in complexity, and no large differences in encoding times are to be expected.

We are optimistic that the extension of the diffusion model might prove useful for analysing data from many psychological paradigms. To make this kind of analysis accessible to a broader community of researchers, the extension will be implemented in one of the next releases of the *fast-dm* software (Voss & Voss, 2007; available from www.psychologie.uni-freiburg.de/Members/voss/fast-dm).

References

- Campbell, K. C., & Proctor, R. W. (1993). Repetition effects with categorizable stimulus and response sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1345–1362.
- Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 229–240.
- Eriksen, C. W., Coles, M. G., Morris, L. R., & O'Hara, W. P. (1985). An electromyographic examination of response competition. *Bulletin of the Psychonomic Society*, *23*(3), 165–168.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229–238.
- Grasman, R. P. P., Wagenmakers, E.-J., & van der Maas, H. L. J. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *Journal of Mathematical Psychology*, *53*, 55–68.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.
- Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology*, *93*(3), 353–368.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.
- Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review*, *92*(2), 212–225.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333–367.
- Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin and Review*, *13*(4), 626–635.
- Ratcliff, R., Thapar, A., & McKoon, G. (2007). Application of the diffusion model to two-choice tasks for adults 75–90 years old. *Psychology and Aging*, *22*, 56–66.
- Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*(3), 414–429.

- Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 101-117.
- Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin and Review*, *14*(6), 1011-1026.
- Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, *40*, 61-72.
- Voss, A., Rothermund, K., & Brandstädter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, *44*, 1048-1056.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory and Cognition*, *32*(7), 1206-1220.
- Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767-775.
- Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion-model parameters. *Journal of Mathematical Psychology*, *52*, 1-9.
- Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, *21*(5), 641-671.
- Wagenmakers, E.-J., van der Maas, H. L. J., Dolan, C. V., & Grasman, R. P. P. (2008). EZ does it! Extensions of the EZ-diffusion model. *Psychonomic Bulletin and Review*, *15*(6), 1229-1235.
- Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin and Review*, *14*(1), 3-22.

Received 13 August 2008; revised version received 17 September 2009

Appendix

Means and standard deviations of inter-trial variability of starting-point (s_z), drift rate (s_v), and RT constant (s_{t0}) as a function of parameter set and model restrictions (results from Study 1).

Restrictions	Parameter set							
	1		2		3		4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Fixed								
s_z	0.37	0.13	0.21	0.17	0.45	0.26	0.45	0.27
s_v	0.54	0.40	0.48	0.18	0.88	0.80	0.88	0.80
s_{t0}	0.07	0.06	0.11	0.06	0.02	0.03	0.02	0.03
$d = 0$								
s_z	0.35	0.09	0.30	0.14	0.20	0.14	0.20	0.15
s_v	0.54	0.33	0.26	0.11	0.97	0.38	0.97	0.38
s_{t0}	0.04	0.04	0.06	0.04	0.06	0.03	0.06	0.03
$z = 0.5a$								
s_z	0.36	0.10	0.24	0.15	0.21	0.21	0.21	0.21
s_v	0.53	0.39	0.25	0.14	0.71	0.60	0.71	0.60
s_{t0}	0.06	0.05	0.05	0.04	0.05	0.05	0.05	0.05
Complete								
s_z	0.36	0.10	0.32	0.13	0.24	0.13	0.24	0.13
s_v	0.50	0.31	0.28	0.09	0.90	0.38	0.90	0.38
s_{t0}	0.04	0.04	0.05	0.04	0.04	0.03	0.04	0.03

Note. Fixed: $d = 0$ and $z = a/2$; complete: no restrictions on d or z .