



When accuracy rates and mean response times lead to false conclusions: A simulation study based on the diffusion model

Veronika Lerche^a and Andreas Voss^a

^aPsychologisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany

Abstract ■ The diffusion model belongs to the class of sequential sampling models. Based on information about RT distributions from binary tasks, the model allows the separation of distinct processes underlying decision-making. With the present simulation study, we demonstrate that diffusion modeling has essential advantages over the mere analysis of behavioral variables such as accuracy rate or mean RT. Specifically, our results show that qualitatively different sets of diffusion model parameters can lead to the same pattern of these aggregated variables, which renders the interpretation of these measures ambiguous. Even apparently clear patterns of results can be based on completely different parameter sets, which reflect completely different cognitive processes. For example, if one group's performance (or performance in one experimental condition) is superior in both mean RT and accuracy rate (i.e., fast responses and few errors), this does not necessarily mean that in this group (or condition) information processing is faster (drift rate of the diffusion model). Accordingly, we conclude that the mere analysis of behavioral variables can result in false conclusions.

Keywords ■ diffusion model, validity, mathematical models, reaction time methods. **Tools** ■ fast-dm.

veronika.lerche@psychologie.uni-heidelberg.de

[10.20982/tqmp.16.2.p107](https://doi.org/10.20982/tqmp.16.2.p107)

Acting Editor ■
Cheng-Ta Yang
(Department of
Psychology, Na-
tional Cheng Kung
University)

Reviewers
■ Two anonymous re-
viewers.

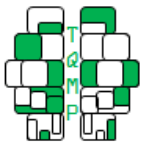
Introduction

Binary response time (RT) tasks are employed in many different fields of psychology, such as in cognitive, motivational, or clinical psychology. Most often, researchers examine the data of these tasks in terms of mean RTs (of correct responses) and/or accuracy rates. If participants in condition A are faster than participants in condition B, it is usually assumed that participants' performance is *better* in condition A than B, that is, that information processing is more efficient. Such differences in RTs, however, become difficult to interpret if participants in condition A are not only faster, but also make more errors than participants in condition B. In this situation, it is unclear whether information processing in condition A is more efficient or whether participants are merely less cautious. Such speed-accuracy trade-offs pose a severe problem to the interpretation of data from RT tasks. If interpretations are exclu-

sively based on mean RTs or accuracy rates, invalid conclusions might be drawn.

Sequential sampling models like the diffusion model (Ratcliff, 1978) go beyond the analysis of aggregated data (as mean RTs or accuracy rates); the diffusion model, for example, takes into account information from the entire RT distributions. Validation studies have shown that the parameters of the diffusion model capture specific psychological processes (Arnold, Bröder, & Bayen, 2015; Lerche & Voss, 2019; Voss, Rothermund, & Voss, 2004). For example, the diffusion model allows to compute a direct measure of speed of information accumulation with other processes such as speed-accuracy trade-offs or motoric response speed "partialled out". Thus, the parameters have specific meanings allowing for more valid interpretations of the effects of studies.

When introducing the diffusion model to other researchers, we frequently noticed that the advantage of dif-



fusion modeling over the analysis of aggregated behavioral variables is not intuitive. Often, researchers argue that from the pattern of behavioral variables alone, it would already be possible to deduce which processes differ between groups or conditions. Consequently, they do not see the need to bother themselves with the application of complex mathematical models that require more efforts than just computing mean RTs or accuracy rates. For example, one common assumption is that if participants in condition A are faster than and at least as accurate as participants in condition B, participants in condition A must have a higher processing speed (i.e., an increased drift rate of the diffusion model). Importantly, this conclusion is not always true. In the present paper, we show results from a simulation study in which we demonstrate that from a pattern of mean RTs and accuracy rates alone it is not possible to infer which cognitive processes vary between conditions. In fact, we illustrate that many different parameter constellations result in very similar patterns of mean RTs and accuracy rates.

In the following, we first give a basic introduction to the diffusion model. Then, we demonstrate how changes of values of single parameters affect predicted mean RT and accuracy rate. Finally, we present our simulation study in which we vary the three main diffusion model parameters (drift rate, threshold separation, and non-decision time) simultaneously and illustrate that substantially different parameter sets can lead to the same accuracy rates and mean RTs.

Introduction to Diffusion Modeling

The diffusion model (Ratcliff, 1978) assumes that in each trial of a binary task participants continuously collect information about two response options until the accumulated information suffices to make a decision (for review articles about the diffusion model, see e.g., Ratcliff, Smith, Brown, & McKoon, 2016; Voss, Nagler, & Lerche, 2013; Wagenmakers, 2009). The model has been applied successfully to many different experimental tasks such as lexical decision tasks (word vs. non-word; e.g., Gomez, Perea, & Ratcliff, 2013; Ratcliff, Gomez, & McKoon, 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008), perceptual tasks (e.g., Ratcliff, 2014; Ratcliff, Thapar, & McKoon, 2003; Voss, Rothermund, & Brandtstädter, 2008), or recognition memory tasks (e.g., Ratcliff & McKoon, 2015; Spaniol, Madden, & Voss, 2006; White, Ratcliff, Vasey, & McKoon, 2010). As recent research demonstrated, the model can be applied not only to tasks with short RTs of a maximum of 1.5 seconds—which was often recommended for diffusion modeling in the past (e.g., Ratcliff & Frank, 2012; Ratcliff, Thapar, Gomez, & McKoon, 2004)—but also to more complex tasks taking up to several seconds per trial (Lerche, Christmann,

& Voss, 2018; Lerche et al., *in press*; Lerche & Voss, 2019).

An illustration of a diffusion process is given in Figure 1. In the figure, thresholds are associated with correct (upper threshold) and erroneous responses (lower threshold). The diffusion process is characterized by a systematic drift with a certain direction (toward upper or lower threshold) and speed (i.e., steepness). Furthermore, random Gaussian noise (where the standard deviation is the so-termed diffusion coefficient) adds to the process, resulting in the erratic path. If a diffusion process reaches one of the two thresholds, the decision has been made and the corresponding response (e.g., a key press) is triggered. In the following, we will shortly introduce the model parameters.

First, there is the threshold separation (parameter a), that is, the distance between the two boundaries, which defines the amount of information needed to reach a decision. If individuals are instructed to respond cautiously (*accuracy instruction*), thresholds are further apart than when individuals are told to respond as quickly as possible (e.g., Ratcliff & Rouder, 1998; Voss et al., 2004). Second, there is drift rate (parameter ν) which informs about the direction (positive: upper threshold, negative: lower threshold) and speed of information accumulation (higher values indicating higher speed). It has been shown that easier tasks go along with higher drift rates (e.g., Arnold et al., 2015; Voss et al., 2004) and also that performance of more intelligent individuals is characterized by higher drift rates (e.g., Ratcliff, Thapar, & McKoon, 2011; Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007; Schubert, Hagemann, Voss, Schankin, & Bergmann, 2015).

Third, there is the starting point (z , or the relative starting point $z_r = z/a$) which defines the position from which the accumulation process starts. If—like in Figure 1—the starting point is situated in the center between thresholds, there is no prior bias for either of the two thresholds. Such prior biases can arise if one of the two response options has a higher expected value (e.g., Leite & Ratcliff, 2011; Voss et al., 2004). Then, the starting point will be located closer to the threshold associated with the respective response.

The diffusion model analysis takes into account that—next to the decision process—also non-decisional processes contribute to the reaction time. The duration of all non-decision processes is mapped by an additional parameter (t_0). These processes include the encoding of information which precedes the accumulation of information and the motoric response which succeeds the decision. Duration of non-decisional processes is, for example, affected by the response modality with more complex motoric responses going along with higher non-decision time (Gomez, Ratcliff, & Childers, 2015; Voss et al., 2004).

In addition to the four main diffusion model parameters discussed above (a , ν , z_r , and t_0), it is often as-



Figure 1 ■ Illustration of the main parameters of the diffusion model. The thresholds—here associated with correct and erroneous responses—are separated by the distance a . One exemplary diffusion process is illustrated that initiates at the starting point z which is here centered between the two thresholds. The process of information accumulation moves with drift rate ν until it reaches one of the two thresholds (here, the upper one). Parameter t_0 —for simplification depicted left to the decision process—includes the time needed for both pre- and post-decisional processes.

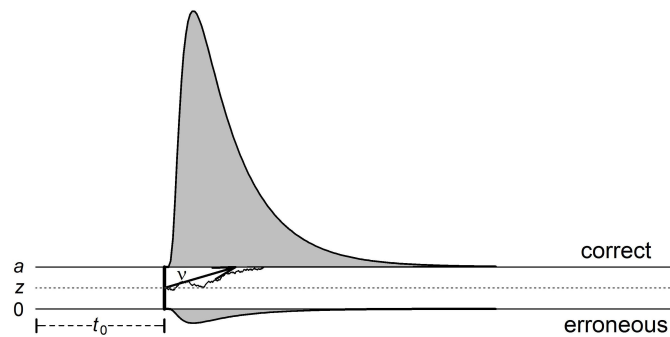


Table 1 ■ Influence of increases of diffusion model parameters a , ν , and t_0 on mean RT and accuracy rate

Increase of	Interpretation	Mean RT	Accuracy rate
a	distance between thresholds	↑	↑
ν	speed of information accumulation	↓	↑
t_0	non-decision time (e.g., encoding, motoric response)	↑	—

sumed that processes can vary from trial to trial (so-called intertrial variabilities). Specifically, the parameters drift (s_ν), starting point (s_{zr}) and non-decision time (s_{t0}) are assumed to vary between trials. Despite their theoretical importance (amongst others in explaining mean differences in RTs of correct and erroneous responses; Ratcliff & Rouder, 1998), the practical advantage of including the three intertrial variability parameters in the model is limited. More specifically, whereas the inclusion of s_{t0} often helps to capture influences of fast contaminants, s_{zr} and s_ν cannot be estimated reliably and fixing rather than estimating these parameters can improve the estimation of the main diffusion model parameters (Lerche & Voss, 2016; see also Boehm et al., 2018; van Ravenzwaaij, Donkin, & Vandekerckhove, 2017). In the following section, we describe the relationships between diffusion model parameters and the behavioral variables mean RT and accuracy rate.

Relationships between Diffusion Model Parameters and Behavioral Variables

Table 1 describes the effects of separately increasing diffusion model parameters (exemplarily for the main param-

eters a , ν and t_0) on the behavioral variables mean RT and accuracy rate. Increasing threshold separation (response caution) influences mean RT and accuracy rate in the same direction: Both variables increase. An increase of drift rate (processing speed), on the other hand, has opposing effects, with accuracy rate rising and mean RT decreasing. Whereas threshold separation and drift rate affect both behavioral variables, an increase in non-decision time solely affects mean RT.

Obviously, Table 1 is based on two simplifications. First, it illustrates only the *general direction of effects*, that is, whether there is a positive or negative effect of a parameter on the behavioral variables (i.e., mean RT and accuracy). For example, according to Table 1, drift rate affects both mean RT and accuracy and so does threshold separation. But what about effect sizes? Does, for example, drift rate affect mean RT and accuracy similarly or does it affect one of these variables more strongly? Ratcliff and McKoon (2008), who examined a large number of data sets ($N = 18$), found a consistent pattern: Drift rate was strongly related to accuracy rate, whereas threshold separation was related primarily to mean RT. For example,

**Table 2** ■ Effects of a hypothetical manipulation (condition B) on accuracies and mean RTs

Case	Effect on Accuracy	Effect on Mean RT	Accuracy	Mean RT (in ms)
1	no effect	no effect	.94	550
2	worse	worse	.90	600
3	better	worse	.98	600
4	no effect	worse	.94	600
5	worse	no effect	.90	550

in a study by Ratcliff, Thapar, and McKoon (2004), drift rate was highly correlated with accuracy rate (younger adults: $r = .74$, older adults: $r = .77$) whereas there was no such consistent pattern for mean RT (younger participants: correct RT: $r = -.014$, error RT: $r = .37$; older participants: correct RT: $r = -.40$, error RT: $r = -.18$). On the other hand, threshold separation was closely related to mean RT (younger participants: correct RT: $r = .81$, error RT: $r = .72$; older participants: correct RT: $r = .87$, error RT: $r = .87$), but it was not correlated with accuracy rate (younger participants: $r = .018$, older participants: $r = -.025$).

Second, in the table the selective influence of *single parameters* is shown, that is, the assumption is made that other parameters remain constant. However, in real data, typically more than one model parameter varies—at least to some extent—as a function of experimental condition or group. Thus, if for example participants are better in terms of accuracy and RT, this does not necessarily mean that this difference is based on drift rate. Also, a complex combination of all three parameters can result in such a pattern of behavioral variables. In our simulation study, we demonstrate that different parameter sets go along with the same pattern of mean RTs and accuracy rates.

Simulation Study

In our simulation study, we generated a large number of parameter sets and examined which of these sets result in the same patterns of mean RTs and accuracy rates. Thereby, we examine whether one and the same pattern of behavioral variables can originate from many—possibly qualitatively completely different—parameter sets.

Method

Our simulation study is based on a simple experimental design that is very common in psychological studies: We assume a design with one independent variable that has two levels (in the following called conditions A and B). For this design, we consider five different possible patterns of results (“cases”).

Specifically, condition A is the standard condition with an accuracy rate of .94 and a mean RT for correct responses (in the following, termed “mean RT”) of 550 ms. The five

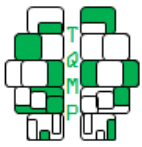
different patterns of the behavioral variables accuracy rate and mean RT in conditions A and B are displayed in Table 2. In case 1, we assumed identical values for accuracy rate and mean RT for both conditions A and B. If such data were collected in an empirical study, researchers who base their analyses on the behavioral variables would come to the conclusion that there are no differences in performance between the two conditions. In case 2, performance in condition B is worse than in condition A, both regarding mean RT (600 ms) and accuracy (.90). In case 3, performance in condition B is worse regarding mean RT (600 ms), but better regarding accuracy (.98). Finally, in cases 4 and 5, one behavioral variable is not affected, while the other variable indicates a decreased performance (case 4: mean RT = 600 ms, case 5: accuracy = .90).

To test which cognitive processes can explain the postulated behavioral effects, we generated a large number of parameter sets. More specifically, for each of the parameters a , ν , and t_0 we used 100 different values. Parameter a varied between 0.500 and 1.985 (step size: 0.015), parameter ν between 1.000 and 4.960 (step size: 0.040), and parameter t_0 between 0.100 and 0.595 (step size: 0.005). The minimum and maximum values for each parameter are based on typical values reported in the diffusion model literature. Thus, the parameter values of almost all diffusion model studies will be captured by these ranges. In total, we generated $100^3 = 1,000,000$ different parameter sets. Then, for each parameter set, we simulated 10,000 trials (assuming a deterministic process) using *construct-samples* which is part of the program *fast-dm* (Voss & Voss, 2007, 2008; Voss, Voss, & Lerche, 2015). Finally, for each simulated data set, we computed mean RT and accuracy rate, rounded to two decimal places for accuracy and to 10 ms for mean RT.

Results

For each of the five cases, we counted the number of parameter sets that resulted in the assumed values of the behavioral variables. Note that due to rounding of numbers, all simulations were counted as hits, when accuracy differed less than 0.005 from the postulated value and mean RT differed less than 5 ms from the postulated value.

Our analyses showed that, for example, a total of $N = 670$ different parameter sets lead to an accuracy rate of .94



and a mean RT of 550 ms (case 1). Thus, an impressive number of different parameter constellations leads to the same pattern of results. The numbers of different parameter sets for all five cases are given in Table 3 (second column) at the end of the article.

Next, we further examined the parameter sets that could explain the results. For this purpose, we first selected an example parameter set that was representative for the standard condition A ($a = 0.965$, $\nu = 2.840$, $t_0 = 0.400$). This example set is approximately in the center of the range of each parameter.¹ We assume that these values represent the true cognitive processes in condition A. This allows us to compare whether the parameters that possibly explain performance in condition B, are similar to parameters from condition A, or—if not—in which direction they changed.

A parameter can either be identical to the parameter of the standard condition ($=$), or smaller ($<$), or larger ($>$). As there are three parameters (a , ν , and t_0) and three different relationships ($=$, $<$, $>$), there is a total of $3^3 = 27$ possible patterns of relations. For each of these patterns, we counted the number of parameter sets. For example, the parameter set $a = 0.890$, $\nu = 3.040$ and $t_0 = 0.425$ for condition B would be classified as ($<$, $>$, $>$) because it is smaller in a , and larger in both ν and t_0 compared to the parameters of the standard condition A. If a parameter differed only slightly from the parameter of the standard condition, we defined this parameter as identical to the standard condition ($=$). Specifically, absolute deviations up to 0.074, 0.198, and 0.025 were regarded as irrelevant for parameters a , ν , and t_0 , respectively (i.e., as identical to the standard condition). These limits are 5% of the range of each parameter used for the construction of the data sets (e.g., for drift: $(4.960-1.000) \times .05 = 0.198$). We used these limits so that only empirically relevant deviations are considered.

The results of the analyses are presented in Table 3 at the end of the article. For each case, the relationship patterns are sorted by frequency (see N -rows). Only combinations are shown that were actually found to match the postulated relationship patterns (frequency of at least 1). For each pattern, we give one example parameter set. As can be seen in the table for case 1, the behavioral variables from the standard condition can also result from qualitatively completely different parameter sets, for example, with a being smaller and ν and t_0 being larger than in the standard condition ($N = 291$) or with a being larger and ν and t_0 being smaller than in the standard condition ($N = 282$). The plots in Figure 2 illustrate the three

most frequent relationship patterns for case 1. The upper plot is based on the values of the standard condition (i.e., $a = 0.965$, $\nu = 2.840$, $t_0 = 0.400$). The middle plot illustrates the ($>$, $<$, $<$) relationship pattern ($a = 1.670$, $\nu = 1.600$, $t_0 = 0.100$) and the lower plot the ($<$, $>$, $>$) pattern ($a = 0.890$, $\nu = 3.040$ and $t_0 = 0.425$). It can be seen that the RT distributions—despite having the same mean RT and accuracy rate—have different positions and forms.

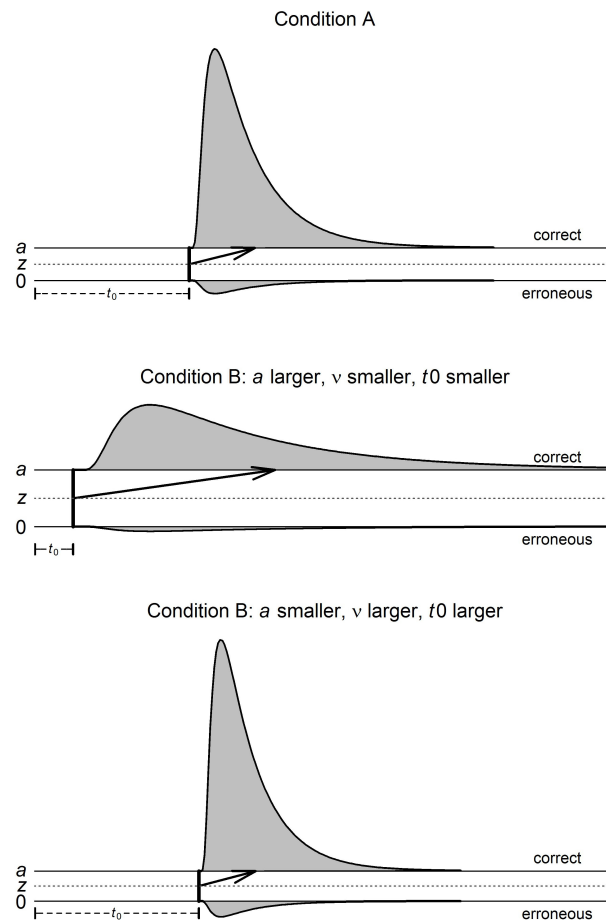
In case 2, condition B is characterized by a worse performance compared to condition A regarding both accuracy rate (lower) and mean RT (higher). At this point, researchers will typically infer that processing speed (i.e., drift rate) is lower in condition B than in condition A. However, this is not necessarily true. For example, a ($<$, $>$, $>$) relation pattern ($N = 144$) could also explain this pattern of behavioral variables. Thus, in condition B, participants might actually be faster in their speed of information accumulation than in condition A. The higher error rate in condition B is then a result of the reduced threshold separation. Participants are less cautious which leads to more errors and faster RTs. The higher non-decision time component counteracts these fast RTs so that—in sum—participants make more errors and are slower in condition B than in condition A.

Also in case 3, researchers might interpret the behavioral variables incorrectly. As condition B has a higher accuracy but a worse mean RT, they might conclude that participants in condition B are more cautious which is captured in the threshold separation parameter of the diffusion model. However, the same pattern of behavioral results could also be the consequence of a reduced threshold separation if this goes along with an increase in both drift rate and non-decision time ($N = 228$). The higher accuracy rate is then explained by the higher drift rate and the slower RTs are due to the higher non-decision time. Thus, as these examples show, even seemingly clear patterns of results might not be interpreted correctly. Again, the diffusion model analysis helps to disentangle the true processes involved here.

In additional analyses, we only considered large differences in parameter values between condition A and condition B. Specifically, we set the limits to 10% (in contrast to the 5% criterion applied in the previous analyses) of the range of each parameter used for the construction of the data sets (e.g., for drift: $(4.960-1.000) \times .10 = 0.396$). Even if only such large differences are considered, the numbers of simulations matching the most common patterns ($>$, $<$, $<$) and ($<$, $>$, $>$) are only slightly reduced. For example, in case 1, the number of parameter sets of the ($>$, $<$, $<$) pat-

¹Specifically, across the 670 parameter sets that result in an accuracy rate of .94 and a mean RT of 550 ms, we first computed the median of parameter a and selected all parameter sets with parameter a equal to the median ($N = 8$). Then, we followed the same approach consecutively for parameters ν ($N = 2$) and t_0 which resulted in one remaining parameter set.

Figure 2 ■ Three exemplary, completely different parameter sets resulting in the same pattern of mean RT (550 ms) and accuracy rate (.94).



tern drops from 282 for the 5% criterion to 242 for the 10% criterion (see Table 4 at the end of the article for details).

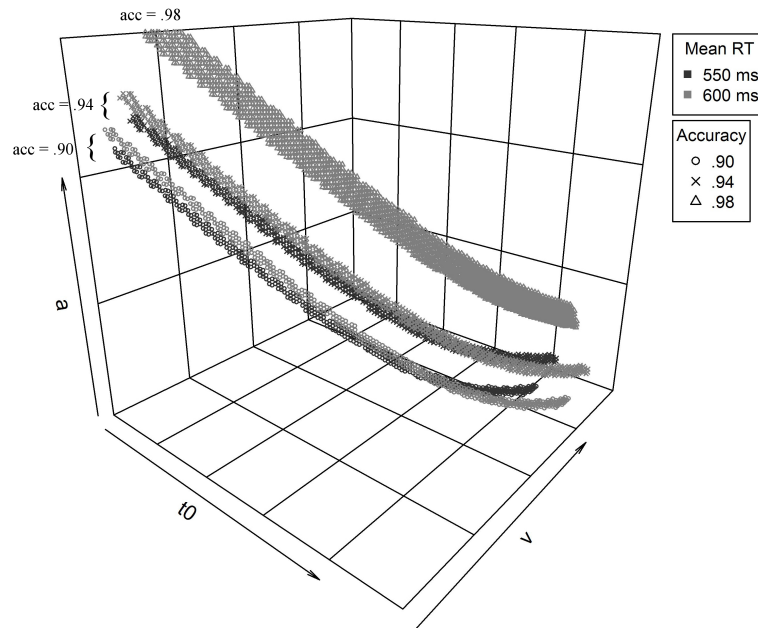
In Figure 3, all parameter sets are depicted that result in one of the five patterns of accuracy rate and mean RT. Each dot symbolizes one parameter set. The plot illustrates that very different parameter sets result in the same accuracy and mean RT. Furthermore, it can be seen that there is a clear relationship between the three parameters. A higher threshold separation goes along with lower non-decision time and drift rate. This is plausible as a higher threshold separation leads to fewer errors so that drift rate does not need to be high. A high threshold separation and small drift rate both lead to higher RTs so that non-decision time can be smaller. Importantly, these relationships do not mean that the different parameter sets cannot

be separated in the parameter estimation procedure. In our simulation study, we only considered accuracy rates and mean RTs and examined which parameter sets result in the same values of these behavioral variables. Our study demonstrates that based on these behavioral variables alone, it is not possible to disentangle different accounts. However, diffusion modeling is not just based on accuracy rates and mean RTs, but on entire distributions of response times. With this higher amount of information it becomes possible to estimate parameters reliably and thus distinguish between different possible patterns of model parameters (for exemplary simulation studies, see Lerche & Voss, 2018).

Figure 3 further illustrates that for higher accuracy rates, more parameter sets result in the same pattern of



Figure 3 ■ Scatter plot of all parameter sets that result in the same accuracy and mean RT.



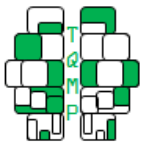
behavioral variables (for the exact values, see the second column in Table 3). This also suggests that for experiments with high accuracies more trials are needed to allow for reliable parameter estimation because the model must be able to discriminate between the numerous possible parameter sets (Lerche, Voss, & Nagler, 2017). Finally, the plot reveals that if non-decision time is larger and threshold separation lower, the faster mean RT (550 ms) goes along with higher drift rates than the slower mean RT (600), whereas for higher threshold separations and lower non-decision times, the faster mean RT goes along with smaller drift rates than the slower mean RT.

Discussion

The popularity of the diffusion model has greatly risen in the last two decades (Voss et al., 2013). Whereas in the past the model was mainly used by mathematical psychologists, now researchers of many different fields of psychology apply this model (which is most probably attributable to the availability of user-friendly software solutions, e.g., Vandekerckhove & Tuerlinckx, 2008; Voss & Voss, 2007; Wagenmakers, van der Maas, & Grasman, 2007; Wiecki, Sofer, & Frank, 2013). Consequently, it becomes more and more

important to provide tutorials that demonstrate why the model can be useful, how the model parameters are best estimated and in which cases the model cannot be applied or parameter estimates need to be interpreted with caution. Critically, up to date, the number of articles that provide guidelines for researchers interested in diffusion modeling is limited. There are a few articles that examine how many trials are required for diffusion modeling (Lerche & Voss, 2017; Lerche et al., 2017), which estimation method performs best (Lerche et al., 2017; Ratcliff & Childers, 2015; Ratcliff & Tuerlinckx, 2002; Wiecki et al., 2013) and which model parameters are difficult to estimate (Boehm et al., 2018; Lerche & Voss, 2016; van Ravenzwaaij et al., 2017). Furthermore, there is a newly developed tool that allows to illustrate the influence of changing model parameters on RT distributions (Alexandrowicz, 2018). With this article we want to contribute to this line of research.

In the present paper, we aim at clarifying a common misconception regarding the relationships between diffusion model parameters and behavioral variables. More specifically, we have repeatedly talked to researchers who assume that it is easy to deduce diffusion model parameters from patterns of behavioral variables. For exam-



ple, if in one condition participants are both faster and make fewer errors, the common assumption is that here speed of information accumulation (drift rate parameter of the diffusion model) must be increased. Similarly, if individuals are slower but at the same time more accurate, this might be attributed to an effect in threshold separation which defines how much information is accumulated before a decision is made. Often, these inferences will hold; however, this will not always be the case because—as we demonstrated in this article based on a simulation study—qualitatively different sets of model parameters can result in the same mean RTs and accuracy rates. For example, individuals might be faster and more accurate in their responses as a result of the combination of (1) being more cautious (higher threshold separation), (2) being faster in their motor response (reduced non-decision time), and (3) simultaneously being slower in information processing (lower drift rate).

Furthermore, if two conditions do not differ at all in terms of mean RT and accuracy rate, there can still be substantial differences in cognitive processes as measured by diffusion model parameters. The diffusion model considers information about entire RT distributions (e.g., including information of standard deviations and skew), thereby allowing to disentangle cognitive processes that cannot be separated based on analyses of mean RTs or accuracy rates alone.

Interestingly, the large validation project by Dutilh et al. (2018) showed that it might—at least in some cases—also be possible to draw accurate inferences from the inspection of summary statistics alone. In this project, the participating research teams were instructed to analyze data of several pseudo-experiments and find out in which of the four psychological processes (ease of information processing, caution, decision bias, non-decision time) two conditions differed. Importantly, the research teams were blind regarding the manipulations. Whereas most teams opted to apply either the diffusion model or the linear ballistic accumulator model (Brown & Heathcote, 2008), two teams (Evans & Brown = EB; van Maanen = MA) decided to make their inferences without applying any model, but by mere inspection of summary statistics. Strikingly, both teams performed well and one team (EB) even outperformed most other teams. Certainly, one needs to consider that both teams consisted of researchers who were very experienced in the application of sequential sampling models. If novices had used the same strategy, it is questionable whether their conclusions would be as accurate. Furthermore, it is important to mention that the EB team examined next to median RTs and accuracy rates also RT quantiles and thus used also information about the shape of distributions of response times. Finally, it should

be considered that in the project of Dutilh et al. (2018) experimental effects were rather large; if smaller effects are to be detected, the analysis of summary statistics alone will probably be less successful.

In many studies, like in the experiments by Dutilh et al. (2018), researchers want to find out in which of the diffusion model parameters two conditions differ. Therefore, they will allow all main diffusion model parameters to vary between the conditions. In other studies, the experimental design might make variations of diffusion model parameters impossible or unlikely. Imagine, for example, a task with two types of stimuli. Researchers want to find out whether the two types of stimuli are encoded differently (parameter: non-decision time) and/or differ in how well they are processed (parameter: drift rate). The stimuli are presented intermixed. According to the conception of the diffusion model, threshold separation is set prior to the presentation of the stimulus. It is, thus, not possible that threshold separation differs between the two types of stimuli because prior to the presentation of the stimulus, participants do not know yet which stimulus will be shown. Accordingly, in such an experiment, modelers might only allow drift rate and non-decision time to vary between the different types of stimuli, and estimate one value for threshold separation across conditions. Thereby, the model becomes more restricted and there are fewer combinations of parameters that result in the same pattern of mean RTs and accuracy rates. Thus, the number of different combinations of parameters that go along with the same mean RTs and accuracy rates depends on characteristics of the specific study.

Note that our simulation study was limited to a variation of the three main diffusion model parameters threshold separation (a), drift rate (ν) and non-decision time (t_0). Similar simulations as presented in this article could be done including also the other diffusion model parameters (starting point z , and inter-trial variabilities), examining more complex models (e.g., with a bias for one of the two response alternatives). In fact, with an increase in the number of considered model parameters the model becomes even more flexible to account for certain patterns of mean RTs and accuracy rates. The general message of such additional simulations would be the same, that is, that an interpretation of behavioral variables can lead to wrong conclusions because various different parameter sets might be responsible for a certain pattern of behavioral data.

As outlined above, our simulation study shows that very different patterns of parameters result in the same patterns of behavioral variables. The diffusion model is able to discriminate between these different parameter patterns because information about RT distributions



(rather than only summary statistics) are considered. Note that the diffusion model is also a falsifiable model (for a discussion of this argument, see e.g., Heathcote, Wagenmakers, & Brown, 2014; see Roberts & Pashler, 2000, for a more general discussion of model fit in theory testing).

Critically, there are high demands on the parameter estimation procedure which should result in unbiased estimates. Several empirical and simulation studies have compared different estimation procedures in terms of reliability of parameter estimation (e.g., Lerche & Voss, 2017; Lerche et al., 2017; Ratcliff & Childers, 2015; Ratcliff & Tuerlinckx, 2002; Wiecki et al., 2013). Among other things, the studies showed that different estimation procedures might be the optimal choice depending on features of the task, like the number of trials or possible contaminations of data. Whereas for large trial numbers there are small differences between different estimation procedures, for smaller trial numbers it is crucial to use a reliable procedure. For example, the optimization criterion maximum likelihood seems suitable also for small trial numbers (100 or even fewer trials), whereas the chi-square optimization criterion does not provide reliable estimates for such trial numbers (Lerche et al., 2017).

Importantly, even if a model fits the data of a study well, this does not necessarily mean that the estimated parameter values reflect the true cognitive processes. Therefore, it is important to conduct validation studies to examine the validity of the parameters of a model. For the diffusion model, there are a few studies testing the validity of the main diffusion model parameters. Among these studies are experimental validation studies (so-called selective influence studies, Arnold et al., 2015; Lerche & Voss, 2019; Voss et al., 2004) that selectively manipulated processes assumed to underlie diffusion model parameters and analyzed whether these manipulations affect the respective parameter in the predicted way (convergent validity) and take no influence on other parameters (discriminant validity). Furthermore, there are correlational studies showing, for example, relationships of drift rate with general intelligence (Ratcliff et al., 2011; Schmiedek et al., 2007; Schubert et al., 2015), and lately also with domain-specific intelligence (Lerche et al., *in press*). However, the rather small number of validation studies is in striking contrast to the high number of application studies that assume that the validity of the model parameters is given. In our eyes, future research should continue to thoroughly examine the validity of the diffusion model parameters.

Authors' note

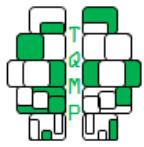
The research was supported by a grant from the German Research Foundation to Andreas Voss (Grant Nr. VO1288/2-2).

References

- Alexandrowicz, R. W. (2018). The diffusion model visualizer: An interactive tool to understand the diffusion model parameters. *Psychological Research*, 53, 1430–2772. doi:[10.1007/s00426-018-1112-6](https://doi.org/10.1007/s00426-018-1112-6)
- Arnold, N. R., Bröder, A., & Bayen, U. J. (2015). Empirical validation of the diffusion model for recognition memory and a comparison of parameter estimation methods. *Psychological Research*, 79(5), 882–898. doi:[10.1007/s00426-014-0608-y](https://doi.org/10.1007/s00426-014-0608-y)
- Boehm, U., Annis, J., Frank, M. J., Hawkins, G. E., Heathcote, A., Kellen, D., ... Wagenmakers, E.-J. (2018). Estimating across-trial variability parameters of the diffusion model: Expert advice and recommendations. *Journal of Mathematical Psychology*, 87, 46–75. doi:[10.1016/j.jmp.2018.09.004](https://doi.org/10.1016/j.jmp.2018.09.004)
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57(3), 153–178. doi:[10.1016/j.cogpsych.2007.12.002](https://doi.org/10.1016/j.cogpsych.2007.12.002)
- Dutilh, G., Annis, J., Brown, S. D., Cassey, P., Evans, N. J., Grasman, R. P. P., ... Donkin, C. (2018). The quality of response time data inference: A blinded, collaborative assessment of the validity of cognitive models. *Psychonomic Bulletin & Review*, 26, 1051–1069. doi:[10.3758/s13423-017-1417-2](https://doi.org/10.3758/s13423-017-1417-2)
- Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked versus unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance*, 39(6), 1731–1740. doi:[10.1037/a0032333](https://doi.org/10.1037/a0032333)
- Gomez, P., Ratcliff, R., & Childers, R. (2015). Pointing, looking at, and pressing keys: A diffusion model account of response modality. *Journal of Experimental Psychology: Human Perception and Performance*, 41(6), 1515–1523. doi:[10.1037/a0039653](https://doi.org/10.1037/a0039653)
- Heathcote, A., Wagenmakers, E. J., & Brown, S. D. (2014). The falsifiability of actual decision-making models. *Psychological Review*, 121(4), 676–678. doi:[10.1037/a0037771](https://doi.org/10.1037/a0037771)
- Leite, F. P., & Ratcliff, R. (2011). What cognitive processes drive response biases? A diffusion model analysis. *Judgment and Decision Making*, 6(7), 651–687.
- Lerche, V., Christmann, U., & Voss, A. (2018). Impact of context information on metaphor elaboration: A diffusion model study. *Experimental Psychology*, 65(6), 370–384. doi:[10.1027/1618-3169/a000422](https://doi.org/10.1027/1618-3169/a000422)
- Lerche, V., von Krause, M., Voss, A., Frischkorn, G. T., Schubert, A.-L., & Hagemann, D. (in press). Diffusion modeling and intelligence: Drift rates show both domain-



- general and domain-specific relations with intelligence. *Journal of Experimental Psychology: General*.
- Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in Psychology*, 7, 1324. doi:[10.3389/fpsyg.2016.01324](https://doi.org/10.3389/fpsyg.2016.01324)
- Lerche, V., & Voss, A. (2017). Retest reliability of the parameters of the Ratcliff diffusion model. *Psychological Research*, 81(3), 629–652. doi:[10.1007/s00426-016-0770-5](https://doi.org/10.1007/s00426-016-0770-5)
- Lerche, V., & Voss, A. (2018). Speed-accuracy manipulations and diffusion modeling: Lack of discriminant validity of the manipulation or of the parameter estimates? *Behavior Research Methods*, 50(6), 2568–2585. doi:[10.3758/s13428-018-1034-7](https://doi.org/10.3758/s13428-018-1034-7)
- Lerche, V., & Voss, A. (2019). Experimental validation of the diffusion model based on a slow response time paradigm. *Psychological Research*, 83(6), 1194–1209. doi:[10.1007/s00426-017-0945-8](https://doi.org/10.1007/s00426-017-0945-8)
- Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods*, 49(2), 513–537. doi:[10.3758/s13428-016-0740-2](https://doi.org/10.3758/s13428-016-0740-2)
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. doi:[10.1037/0033-295x.85.2.59](https://doi.org/10.1037/0033-295x.85.2.59)
- Ratcliff, R. (2014). Measuring psychometric functions with the diffusion model. *Journal of Experimental Psychology: Human Perception and Performance*, 40(2), 870–888. doi:[10.1037/a0034954](https://doi.org/10.1037/a0034954)
- Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision*, 2(4), 237–279. doi:[10.1037/dec0000030](https://doi.org/10.1037/dec0000030)
- Ratcliff, R., & Frank, M. J. (2012). Reinforcement-based decision making in corticostriatal circuits: Mutual constraints by neurocomputational and diffusion models. *Neural Computation*, 24(5), 1186–1229. doi:[10.1162/NECO_a_00270](https://doi.org/10.1162/NECO_a_00270)
- Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, 111(1), 159–182. doi:[10.1037/0033-295x.111.1.159](https://doi.org/10.1037/0033-295x.111.1.159)
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. doi:[10.1162/neco.2008.12-06-420](https://doi.org/10.1162/neco.2008.12-06-420)
- Ratcliff, R., & McKoon, G. (2015). Aging effects in item and associative recognition memory for pictures and words. *Psychology and Aging*, 11. doi:[10.1037/pag0000030](https://doi.org/10.1037/pag0000030)
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9(5), 347–356. doi:[10.1111/1467-9280.00067](https://doi.org/10.1111/1467-9280.00067)
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in cognitive sciences*, 20(4), 260–281. doi:[10.1016/j.tics.2016.01.007](https://doi.org/10.1016/j.tics.2016.01.007)
- Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, 19(2), 278–282. doi:[10.1037/0882-7974.19.2.278](https://doi.org/10.1037/0882-7974.19.2.278)
- Ratcliff, R., Thapar, A., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & Psychophysics*, 65(4), 523–535. doi:[10.3758/BF03194580](https://doi.org/10.3758/BF03194580)
- Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, 50(4), 408–424. doi:[10.1016/j.jml.2003.11.002](https://doi.org/10.1016/j.jml.2003.11.002)
- Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and iq on item and associative memory. *Journal of Experimental Psychology: General*, 140(3), 464–487. doi:[10.1037/a0023810](https://doi.org/10.1037/a0023810)
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9(3), 438–481. doi:[10.3758/bf03196302](https://doi.org/10.3758/bf03196302)
- Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? a comment on theory testing. *Psychological Review*, 107(2), 358–367. doi:[10.1037/0033-295X.107.2.358](https://doi.org/10.1037/0033-295X.107.2.358)
- 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. doi:[10.1037/0096-3445.136.3.414](https://doi.org/10.1037/0096-3445.136.3.414)
- Schubert, A.-L., Hagemann, D., Voss, A., Schankin, A., & Bergmann, K. (2015). Decomposing the relationship between mental speed and mental abilities. *Intelligence*, 51, 28–46. doi:[10.1016/j.intell.2015.05.002](https://doi.org/10.1016/j.intell.2015.05.002)
- 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. doi:[10.1037/0278-7393.32.1.101](https://doi.org/10.1037/0278-7393.32.1.101)
- van Ravenzwaaij, D., Donkin, C., & Vandekerckhove, J. (2017). The ez diffusion model provides a powerful test of simple empirical effects. *Psychonomic Bulletin & Review*, 24(2), 547–556. doi:[10.3758/s13423-016-1081-y](https://doi.org/10.3758/s13423-016-1081-y)



- Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, 40(1), 61–72. doi:[10.3758/brm.40.1.61](https://doi.org/10.3758/brm.40.1.61)
- Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: A practical introduction. *Experimental Psychology*, 60(6), 385–402. doi:[10.1027/1618-3169/a000218](https://doi.org/10.1027/1618-3169/a000218)
- Voss, A., Rothermund, K., & Brandtstädter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, 44(4), 1048–1056. doi:[10.1016/j.jesp.2007.10.009](https://doi.org/10.1016/j.jesp.2007.10.009)
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, 32(7), 1206–1220. doi:[10.3758/BF03196893](https://doi.org/10.3758/BF03196893)
- Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, 39(4), 767–775. doi:[10.3758/bf03192967](https://doi.org/10.3758/bf03192967)
- Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology*, 52(1), 1–9. doi:[10.1016/j.jmp.2007.09.005](https://doi.org/10.1016/j.jmp.2007.09.005)
- Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology*, 6, 336–344. doi:[10.3389/fpsyg.2015.00336](https://doi.org/10.3389/fpsyg.2015.00336)
- Wagenmakers, E.-J. (2009). Methodological and empirical developments for the diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, 21(5), 641–671. doi:[10.1080/09541440802205067](https://doi.org/10.1080/09541440802205067)
- Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, 58(1), 140–159. doi:[10.1016/j.jml.2007.04.006](https://doi.org/10.1016/j.jml.2007.04.006)
- 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 & Review*, 14(1), 3–22. doi:[10.3758/bf03194023](https://doi.org/10.3758/bf03194023)
- White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, 54(1), 39–52. doi:[10.1016/j.jmp.2010.01.004](https://doi.org/10.1016/j.jmp.2010.01.004)
- Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). Hddm: Hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in neuroinformatics*, 7, 14–19. doi:[10.3389/fninf.2013.00014](https://doi.org/10.3389/fninf.2013.00014)

Citation

Lerche, V., & Voss, A. (2020). When accuracy rates and mean response times lead to false conclusions: A simulation study based on the diffusion model. *The Quantitative Methods for Psychology*, 16(2), 107–119. doi:[10.20982/tqmp.16.2.p107](https://doi.org/10.20982/tqmp.16.2.p107)

Copyright © 2020, Lerche and Voss. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 25/04/2019 ~ Accepted: 21/11/2019

Tables 3 and 4 follow.

Table 3 ■ Number of different parameter sets for condition B that result in the accuracy rates and mean RTs of the five cases. Example parameter sets for the different relationship patterns are given in parentheses

Case	Number of parameter sets	Condition B in comparison with condition A														
1	670	a	< (0.890)	> (1.670)	= (0.965)	= (0.935)	= (1.010)	> (1.040)	< (0.890)	> (1.040)	> (1.040)	> (1.040)	= (0.905)	< (0.890)	< (0.890)	< (0.890)
		ν	> (3.040)	< (1.600)	= (2.840)	> (3.040)	< (2.640)	= (2.680)	> (3.040)	= (2.680)	= (2.680)	< (2.600)	> (3.080)	= (3.000)	= (3.000)	= (3.000)
		t_0	> (0.425)	< (0.100)	= (0.400)	= (0.410)	= (0.380)	= (0.380)	= (0.420)	= (0.375)	< (0.375)	= (0.380)	> (0.425)	= (0.420)	= (0.420)	> (0.425)
		N	291	282	68	10	4	4	3	2	2	2	2	1	1	1
2	457	a	> (1.655)	< (0.725)	< (0.815)	= (0.980)	< (0.890)	> (1.100)	= (1.025)							
		ν	< (1.320)	> (3.040)	= (2.680)	< (2.200)	< (2.440)	< (1.960)	< (2.120)							
		t_0	< (0.100)	> (0.500)	> (0.475)	> (0.425)	> (0.450)	= (0.380)	= (0.405)							
		N	160	144	48	36	30	25	14							
3	1619	a	> (1.970)	> (1.175)	= (1.025)	< (0.890)	> (1.550)	> (1.355)	> (1.310)							
		ν	< (1.880)	> (3.120)	> (3.600)	> (4.120)	= (2.680)	> (3.040)	= (2.800)							
		t_0	< (0.100)	> (0.425)	> (0.460)	> (0.495)	< (0.315)	= (0.380)	= (0.380)							
		N	527	267	246	228	156	133	62							
4	702	a	< (0.890)	> (1.760)	= (1.025)	> (1.175)	= (0.935)	> (1.040)	> (1.040)	= (1.010)	< (0.890)					
		ν	> (3.040)	< (1.520)	= (2.680)	< (2.280)	> (3.040)	< (2.600)	= (2.680)	< (2.640)	= (3.000)					
		t_0	> (0.470)	< (0.100)	> (0.430)	= (0.380)	> (0.460)	> (0.425)	> (0.425)	> (0.430)	> (0.470)					
		N	294	243	68	66	12	7	6	4	2					
5	435	a	> (1.565)	< (0.725)	< (0.815)	= (0.965)	< (0.890)	= (1.025)	< (0.830)							
		ν	< (1.400)	> (3.040)	= (2.680)	< (2.240)	< (2.440)	< (2.120)	< (2.600)							
		t_0	< (0.100)	> (0.450)	> (0.425)	= (0.380)	= (0.400)	< (0.355)	> (0.425)							
		N	163	144	48	31	25	19	5							

Note. In the comparison condition A, parameter values are $a = 0.965$, $\nu = 2.840$, $t_0 = 0.400$.

Table 4 ■ Number of different parameter sets for condition B that result in the accuracy rates and mean RTs of the five cases. Example parameter sets for the different relationship patterns are given in parentheses. Absolute deviations up to 0.148, 0.396, and 0.050 were regarded as irrelevant for parameters a , ν , and t_0 , respectively (i.e., as identical to the standard condition). These limits are 10% (in contrast to the 5% presented in Table 3) of the range of each parameter used for the construction of the data sets)

Case	Number of parameter sets	Condition B in comparison with condition A														
		a	ν	t_0	N	a	ν	t_0	N	a	ν	t_0	N	a	ν	t_0
1	670	> (1.670)	< (0.800)	< (0.965)	= (0.875)	< (0.815)	> (1.145)	> (1.115)	= (1.100)	= (1.100)	< (2.440)	< (2.440)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		< (1.600)	> (3.360)	= (2.840)	> (3.240)	> (3.280)	= (2.480)	= (2.480)	< (0.340)	= (0.355)	< (0.350)	< (0.350)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		< (0.100)	> (0.450)	= (0.400)	= (0.425)	= (0.440)	< (0.340)	= (0.355)	< (0.340)	= (0.355)	< (0.350)	< (0.350)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		242	228	150	24	16	6	2	1	2	1	1	1	1	1	1
2	457	> (1.655)	< (0.695)	< (0.815)	= (1.100)	= (0.905)	> (1.145)	= (0.905)	= (1.100)	= (1.100)	< (2.440)	< (2.440)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		< (1.320)	> (3.240)	= (2.640)	< (1.960)	= (2.480)	< (1.880)	< (2.400)	= (0.355)	< (0.355)	< (0.350)	< (0.350)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		< (0.100)	> (0.510)	> (0.475)	= (0.375)	> (0.450)	= (0.355)	> (0.450)	= (0.355)	< (0.340)	= (0.355)	< (0.350)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		142	120	72	68	30	17	8	1	2	1	1	1	1	1	1
3	1619	= (1.085)	> (1.970)	> (1.370)	> (1.670)	> (1.280)	< (0.815)	> (1.115)	= (1.100)	= (1.100)	< (2.440)	< (2.440)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		> (3.400)	< (1.880)	= (2.680)	= (2.480)	> (3.240)	> (4.520)	> (3.480)	> (4.520)	> (4.520)	> (3.480)	> (3.360)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		> (0.450)	< (0.100)	= (0.355)	< (0.270)	= (0.405)	> (0.510)	> (0.450)	= (0.440)	= (0.440)	< (0.350)	< (0.350)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		509	391	274	198	141	80	19	7	19	7	7	7	7	7	7
4	702	< (0.815)	> (1.760)	= (0.965)	= (1.085)	> (1.235)	= (0.875)	> (1.445)	= (1.100)	= (1.100)	< (2.440)	< (2.440)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		> (3.280)	< (1.520)	= (2.800)	= (2.480)	< (2.160)	> (3.240)	= (2.480)	> (3.240)	> (3.240)	> (3.240)	> (3.240)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		> (0.490)	< (0.100)	> (0.450)	= (0.405)	= (0.355)	> (0.475)	= (0.390)	= (0.400)	= (0.400)	< (0.350)	< (0.350)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		244	210	76	74	64	24	8	2	8	2	2	2	2	2	2
5	435	> (1.565)	< (0.695)	= (1.040)	< (0.815)	< (0.755)	= (0.905)	= (1.100)	= (1.100)	= (1.100)	< (2.440)	< (2.440)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		< (1.400)	> (3.240)	< (2.080)	= (2.640)	= (2.880)	= (2.480)	< (1.960)	= (2.480)	= (2.480)	< (0.340)	< (0.340)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		< (0.100)	> (0.460)	= (0.355)	= (0.425)	> (0.450)	= (0.400)	< (0.325)	= (0.400)	= (0.400)	< (0.350)	< (0.350)	= (1.100)	= (1.100)	< (2.440)	= (0.355)
		137	120	54	36	36	30	22	22	22	22	22	22	22	22	22

Note. In the comparison condition A, parameter values are $a = 0.965$, $\nu = 2.840$, $t_0 = 0.400$.