Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis

DORA MATZKE AND ERIC-JAN WAGENMAKERS
University of Amsterdam, Amsterdam, The Netherlands

A growing number of researchers use descriptive distributions such as the ex-Gaussian and the shifted Wald to summarize response time data for speeded two-choice tasks. Some of these researchers also assume that the parameters of these distributions uniquely correspond to specific cognitive processes. We studied the validity of this cognitive interpretation by relating the parameters of the ex-Gaussian and shifted Wald distributions to those of the Ratcliff diffusion model, a successful model whose parameters have well-established cognitive interpretations. In a simulation study, we fitted the ex-Gaussian and shifted Wald distributions to data generated from the diffusion model by systematically varying its parameters across a wide range of plausible values. In an empirical study, the two descriptive distributions were fitted to published data that featured manipulations of task difficulty, response caution, and a priori bias. The results clearly demonstrate that the ex-Gaussian and shifted Wald parameters do not correspond uniquely to parameters of the diffusion model. We conclude that researchers should resist the temptation to interpret changes in the ex-Gaussian and shifted Wald parameters in terms of cognitive processes. Supporting materials may be downloaded from http://pbr.psychonomic-journals .org/content/supplemental.

The analysis of response times (RTs) has a long history in cognitive psychology (e.g., Hohle, 1965; Luce, 1986; Ratcliff & McKoon, 2008; Townsend & Ashby, 1983). To draw inferences about mental processes, researchers originally relied on measures of central tendency such as the mean or median RT. As it became clear that these measures may lose important information (e.g., Heathcote, Popiel, & Mewhort, 1991), a growing number of researchers started to use mathematical and statistical models that can accommodate not just mean RT, but also the shapes of entire RT distributions.

Primary among the statistical models that facilitate the analysis of RT distributions are the ex-Gaussian and the shifted Wald. Changes in the parameters of these distributions may be used to summarize the effects of experimental manipulations. For instance, Leth-Steensen, King Elbaz, and Douglas (2000) found that children with ADHD differed from age-matched controls specifically in the ex-Gaussian parameter that captures the tail of the RT distribution.

Although the ex-Gaussian and shifted Wald distributions are sometimes used as purely descriptive tools (see, e.g., Wagenmakers, van der Maas, Dolan, & Grasman, 2008), many researchers go one step farther and assume that changes in the parameters of these distributions map onto changes in specific cognitive processes. For instance, Kieffaber et al. (2006) argued that changes in the Gaussian component of the ex-Gaussian distribution reflect changes

in attentional cognitive processes, whereas changes in the exponential component reflect changes in intentional cognitive processes. The purpose of our study is to examine whether this mapping from parameters to processes is warranted. To this end, we attempt to link the parameters of the descriptive distributions to those of the Ratcliff diffusion model (Ratcliff, 1978). The diffusion model provides a theoretical account of performance in speeded two-choice tasks and has been successfully applied across a wide range of paradigms. Most importantly, the parameters of the diffusion model correspond to well-defined psychological processes, such as the rate of information accumulation (influenced by task difficulty or participant ability), response caution, a priori bias, and the time taken by processes unrelated to decision making (e.g., encoding and motor processes). The association between the diffusion model parameters and the psychological processes they are supposed to represent has been confirmed in numerous experiments (e.g., Voss, Rothermund, & Voss, 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008).

The outline of this article is as follows. In the first and second sections, we describe the ex-Gaussian and shifted Wald distributions, respectively. In the third section, we briefly explain the diffusion model and previous research that has studied the association between the diffusion model parameters and the parameters of the two descriptive distributions. The fourth section describes the simulation study, in which we systematically varied the param-

D. Matzke, d.matzke@uva.nl

eters of the diffusion model to study the corresponding changes in the parameters of the descriptive distributions. In the fifth section, we apply the descriptive distributions to recently published lexical decision data that feature manipulations of task difficulty, response caution, and a priori bias (Wagenmakers, Ratcliff, et al., 2008). The sixth section concludes our investigation.

The Ex-Gaussian Distribution

The ex-Gaussian distribution results from the convolution of a Gaussian and an exponential distribution and can be described by three parameters: μ and σ, the mean and standard deviation of the Gaussian component, and τ, the mean of the exponential component. Roughly, μ and σ reflect the leading edge and τ reflects the tail of the distribution. The ex-Gaussian distribution has a positively skewed unimodal shape and generally produces an excellent fit to empirical RT distributions. Figure 1 shows changes in the ex-Gaussian distribution as a result of changes in the ex-Gaussian parameters μ, σ, and τ. The probability density function of the ex-Gaussian is given by

$$f(x \mid \mu, \sigma, \tau) = \frac{1}{\tau\sqrt{2\pi}} \exp\left(\frac{\sigma^2}{2\tau^2} - \frac{x-\mu}{\tau} \right)$$
$$\cdot \int_{-\infty}^{[(x-\mu)/\sigma]-(\sigma/\tau)} \exp\left(-\frac{y^2}{2} \right) dy, \quad (1)$$

and its mean and variance are

$$E(x) = \mu + \tau \quad (2)$$

and

$$\mathrm{Var}(x) = \sigma^2 + \tau^2. \quad (3)$$

Originally, the ex-Gaussian distribution was thought to represent the durations of two successive components of cognitive processing. In particular, Hohle (1965) suggested that the exponential component represents "the decision and perceptual portion of an RT," whereas the Gaussian component reflects "the time required for organization and execution of the motor response" (p. 384). Although it may be tempting to associate these particu-

lar cognitive processes with the ex-Gaussian parameters, Hohle's interpretation of the ex-Gaussian parameters has been frequently challenged.

First, there is disagreement as to which processing mechanism should be attributed to the two ex-Gaussian components. McGill (1963) and McGill and Gibbon (1965), for example, suggested that residual motor latency corresponds to the exponential component of the ex-Gaussian, not to the Gaussian component. This interpretation is diametrically opposed to that of Hohle (1965).

Second, the rationale underlying Hohle's (1965) interpretation of the two ex-Gaussian components has been criticized. Hohle based his interpretation on the finding that the mean of the Gaussian component, μ, was somewhat more sensitive to manipulations of foreperiod duration, but the mean of the exponential component, τ, was more sensitive to manipulations of stimulus intensity. Because foreperiod duration was assumed to influence motor response time, and stimulus intensity was assumed to influence decision time, Hohle concluded that the Gaussian component must reflect residual time and the exponential component the decision portion of the RT.

Luce (1986, pp. 100–102), however, argued that Hohle's (1965) data are equally consistent with many other decompositions of RT, and that the influence of foreperiod duration can in fact be attributed to either component of RT. Also, experimental results often fail to support the differential sensitivity of the ex-Gaussian parameters to manipulations that are assumed to influence the decision component of RT. For example, the manipulation of word frequency, which is assumed to affect the decision portion of RT, commonly influences the τ as well as the μ parameter (see, e.g., Andrews & Heathcote, 2001; Plourde & Besner, 1997; Yap & Balota, 2007; Yap, Balota, Cortese, & Watson, 2006). Table 1 summarizes the effects that experimental manipulations have on the ex-Gaussian parameters, on the basis of a literature review of 54 studies.[1]

Third, and most importantly, it has been argued that the ex-Gaussian distribution lacks a plausible theoretical basis, and that consequently it is unable to account for the psychological mechanisms that drive performance (see, e.g., Heathcote et al., 1991; Luce, 1986). Specifically, the
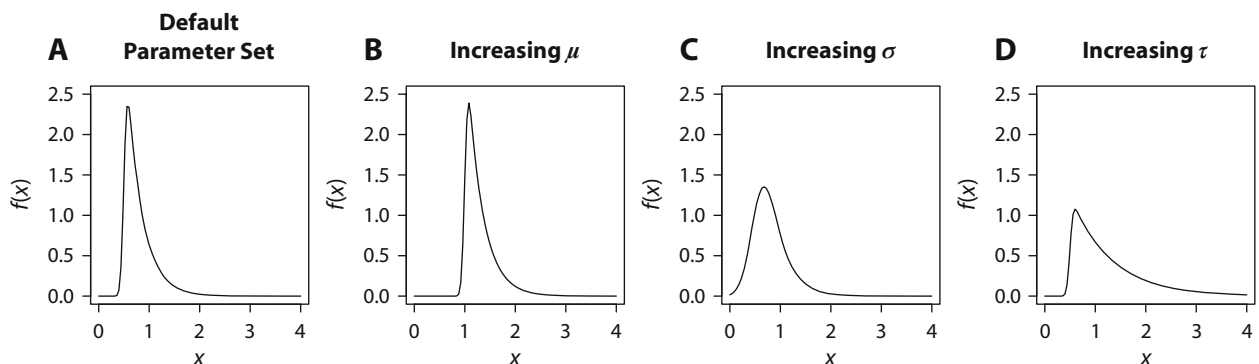

Figure 1. Changes in the ex-Gaussian distribution as a result of changes in the ex-Gaussian parameters μ, σ, and τ. The parameter sets used to generate the distributions are (A) $\mu = 0.5$, $\sigma = 0.05$, $\tau = 0.3$ (default parameter set); (B) $\mu = 1$, $\sigma = 0.05$, $\tau = 0.3$ (increasing μ); (C) $\mu = 0.5$, $\sigma = 0.2$, $\tau = 0.3$ (increasing σ); and (D) $\mu = 0.5$, $\sigma = 0.05$, $\tau = 0.8$ (increasing τ).

Table 1
The Effects of Experimental Manipulations on the Ex-Gaussian Parameters

Experimental Manipulation	N	μ	σ	τ
Word frequency (high or low)	18	16	7	12
Flanker task condition (no flanker, neutral, congruent, and incongruent)	14	13	7	4
Age (children, young adults, and older adults)	11	8	8	7
Length of study list	13	6	2	10
Number of stimulus presentations	9	1	0	8
Stimulus quality (clear or degraded)	9	8	3	8
Stroop task condition (neutral, congruent, and incongruent)	8	8	4	6
Study position of probe items	12	12	1	9
Local/global task[a] condition (neutral, congruent, and incongruent)	6	4	2	1
Output position of recalled items	5	0	0	5
Animacy of stimulus	4	1	1	2
Length of retention interval	4	1	0	4
Nonword type (pseudohomophone, legal, and illegal)	4	4	0	4
Cue-to-target stimulus onset asynchrony	3	3	0	0
Interstimulus interval	3	3	3	3
Speed–accuracy instruction	2	2	0	2

Note—Table 1 summarizes the results of an extensive literature review, covering 54 applications of the ex-Gaussian distribution. The summary is created by selecting the most frequently used experimental manipulations encountered in the literature and tallying how often, out of N attempts, the manipulations influenced each ex-Gaussian parameter. The criterion that researchers used to evaluate whether a given experimental manipulation influenced the ex-Gaussian parameters varied across the experiments. The criterion either was one of $p < .1$, $p < .05$, or $p < .001$ or was based on visual inspection of the changes in parameter values.   [a]An example of a congruent stimulus in the local/global task is the letter H, constructed with small Hs. An example of an incongruent stimulus is the letter H, constructed with small Zs. An example of a neutral stimulus is a circle, constructed with small Hs.

Gaussian component necessarily assigns positive probability to negative RTs, a conceptual inadequacy that highlights the fact that the ex-Gaussian distribution can never correspond to a plausible cognitive process model.

As a consequence of its problematic theoretical underpinning, some researchers have adopted a cautious attitude and have warned against the cognitive interpretation of the ex-Gaussian parameters. As Heathcote et al. (1991) stated, "Although the ex-Gaussian model describes RT data successfully, it does so without the benefit of an underlying theory" (p. 346). Consistent with this view, the ex-Gaussian distribution has sometimes been used as an economical three-parameter summary of RT data and as a tool to evaluate the predictions of competing cognitive models beyond the level of mean RT (e.g., Heathcote et al., 1991; Hockley, 1982, 1984; Ratcliff, 1978, 1993; Ratcliff & Murdock, 1976).

Other researchers, however, have not been so cautious and have persisted with the substantive interpretation of the ex-Gaussian parameters. Rohrer and Wixted (1994), for example, interpreted the Gaussian component as "a brief initiation that precedes retrieval" and the exponential component as "an ongoing search" (pp. 512–513). Balota and Spieler (1999) related the Gaussian component to "more stimulus driven automatic (nonanalytic) processes" and the exponential component to "more central attention demanding (analytic) processes" (p. 34). Kieffaber et al. (2006) interpreted μ in terms of attentional and τ in terms of intentional cognitive processes (p. 348). As a final example, Gordon and Carson (1990) argued that the "lumped sensory input/motor output component" of RT has a Gaussian distribution, and that the "decisional phase" of RT has an exponential distribution (p. 150; see

also Madden et al., 1999; Possamaï, 1991; and Rotello & Zeng, 2008, for similar interpretations). Table 2 gives an overview of the cognitive interpretations attributed to the ex-Gaussian parameters;[2] since the σ parameter is rarely given a cognitive interpretation, it is omitted from the overview. As can be seen in Table 2, there is some consistency in the cognitive interpretation of the ex-Gaussian parameters: Lower-order processes are generally ascribed to μ, and higher-order processes to τ. Note, however, that the precise interpretation of the ex-Gaussian parameters varies considerably across researchers.

To summarize, the ex-Gaussian distribution provides a description of empirical RT data that is accurate but lacks a plausible theoretical rationale. Despite this limitation, the ex-Gaussian parameters have often been interpreted in terms of underlying cognitive processes. The following section introduces the shifted Wald distribution, a descriptive distribution that has the potential to provide parameters that are theoretically more meaningful.

The Shifted Wald Distribution

The Wald (1947) distribution represents the density of the first passage times of a Wiener diffusion process toward a single absorbing boundary (see Figure 2). This distribution can be characterized by two parameters: γ, reflecting the drift rate of the diffusion process, and α, reflecting the separation between the starting point of the diffusion process and an absorbing barrier. In the RT context, the Wald distribution is often supplemented with a positive parameter θ that shifts the entire RT distribution. The shifted Wald has a positively skewed unimodal shape that generally produces an excellent fit to empirical RT distributions. Figure 3 shows changes in the shifted Wald distribution as

Table 2
Cognitive Interpretations Attributed to the Ex-Gaussian Parameters

Authors	μ	τ
Balota and Spieler (1999)	stimulus driven automatic (nonanalytic) processes	central attention demanding (analytic) processes
Blough (1988, 1989)	component of RT unrelated to stimulus variables (e.g., neural transmission and motor response)	momentary probability of target detection/ search component of RT
Epstein et al. (2006), Leth-Steensen et al. (2000)	—	attentional lapses
Gholson and Hohle (1968a, 1968b)	—	response choice latency/response competition
Gordon and Carson (1990), Hohle (1965), Madden et al. (1999), Possamaï (1991), Rotello and Zeng (2008)	duration of residual processes (e.g., sensory and motor processes)	durations of the decisional phase of RT
Kieffaber et al. (2006)	attentional cognitive processes	intentional cognitive processes
Penner-Wilger, Leth-Steensen, and Lefevre (2002)	retrieval processes	nonretrieval/procedure use
Rohrer (1996, 2002), Rohrer and Wixted (1994), Wixted, Ghadisha, and Vera (1997), Wixted and Rohrer (1993)	initial pause preceding the retrieval of the first response	mean recall latency/ongoing memory search
Schmiedek, Oberauer, Wilhelm, Süß, and Wittmann (2007)	—	higher cognitive functioning (e.g., working memory and reasoning)
Spieler, Balota, and Faust (1996)	—	more central processing component

Note—A dash indicates that the parameter is not given any cognitive interpretation.

a result of changes in the parameters α, θ, and γ. The probability density function of the shifted Wald is given by

$$f(x \mid \alpha, \theta, \gamma) = \frac{\alpha}{\sqrt{2\pi(x-\theta)^3}}$$
$$\cdot \exp\left\{-\frac{[\alpha - \gamma(x-\theta)]^2}{2(x-\theta)}\right\}, \qquad (4)$$

where $x > \theta$, and its mean and variance are

$$E(x) = \theta + \alpha/\gamma \qquad (5)$$

and

$$\text{Var}(x) = \alpha/\gamma^3. \qquad (6)$$

For further discussion and applications of the Wald distribution, see Burbeck and Luce (1982), Luce (1986), and Emerson (1970). For discussion and application of a more general version of a single-boundary diffusion process, see Smith (1995).

The cognitive interpretation of the shifted Wald parameters is straightforward (see, e.g., Heathcote, 2004; Luce, 1986; Schwarz, 2001, 2002). Participants are assumed to accumulate noisy information until a predefined threshold amount is reached and a response is initiated. Drift rate γ quantifies task difficulty or participant ability, response criterion α quantifies response caution, and the shift parameter θ quantifies the time needed for nondecision processes.

Although the shifted Wald distribution has a sound theoretical basis, the cognitive interpretation of its parameters has rarely been subject to empirical validation. The shifted Wald model may be particularly suited for paradigms in which there is likely only a single response boundary. Such paradigms may include simple RT tasks (Luce, 1986, pp. 51–57), go/no-go tasks (Heathcote, 2004; Schwarz, 2001), or tasks that involve saccadic eye movements that result in very few errors (Carpenter & Williams, 1995). It is not clear whether the cognitive inter-

pretation of the shifted Wald parameters still holds when the distribution is applied to data from a paradigm that clearly involves two response alternatives.

In summary, both the ex-Gaussian and shifted Wald distributions provide excellent tools to summarize RT distributions. However, the cognitive interpretation of their parameters is unclear. The ex-Gaussian distribution lacks an adequate theoretical basis, and the substantive interpretation of its parameters has been repeatedly questioned. Although the shifted Wald distribution is theoretically better justified, it is currently unclear whether the substantive interpretation of its parameters carries over from one-boundary paradigms to two-boundary paradigms.

The Ratcliff Diffusion Model

The diffusion model (Ratcliff, 1978; for reviews, see Ratcliff & McKoon, 2008; Wagenmakers, 2009) is a prominent cognitive process model of speeded two-choice decisions. The diffusion model assumes that noisy information is accumulated over time from a starting point toward one of two response boundaries (see Figure 4). A response
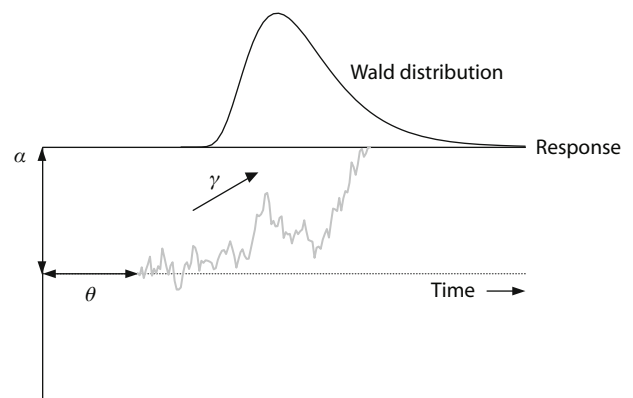

Figure 2. The shifted Wald model of RT and its parameters. See the text for details.
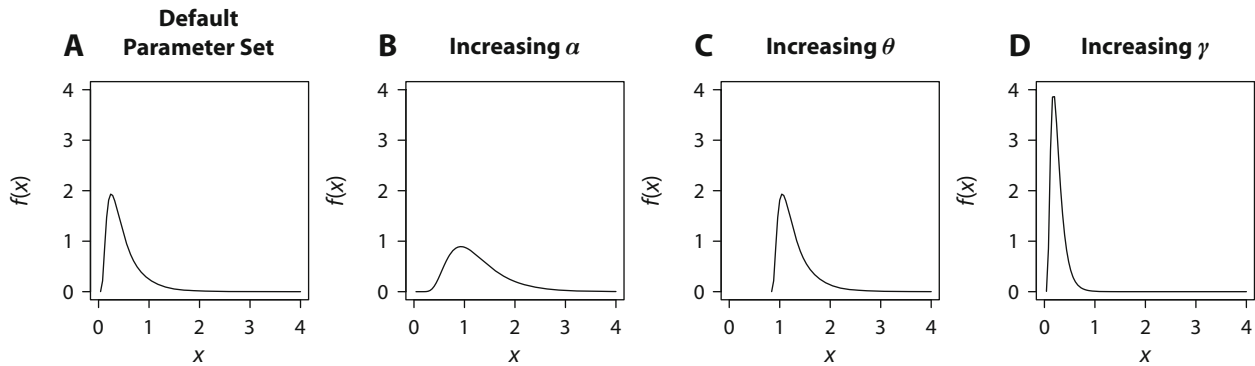
Figure 3. Changes in the shifted Wald distribution as a result of changes in the shifted Wald parameters α, θ, and γ. The parameter sets used to generate the distributions are (A) $\alpha = 1$, $\theta = 0$, $\gamma = 2$ (default parameter set); (B) $\alpha = 2.5$, $\theta = 0$, $\gamma = 2$ (increasing α); (C) $\alpha = 1$, $\theta = 0.8$, $\gamma = 2$ (increasing θ); and (D) $\alpha = 1$, $\theta = 0$, $\gamma = 3.8$ (increasing γ).

is initiated when one of the two response boundaries is reached. The diffusion model has been successfully applied to a wide range of experimental paradigms, including brightness discrimination, letter identification, lexical decision, recognition memory, and signal detection (see, e.g., Ratcliff, 1978, 2002; Ratcliff, Gomez, & McKoon, 2004; Ratcliff & Rouder, 2000; Ratcliff, Thapar, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2001, 2003, 2004; Thapar, Ratcliff, & McKoon, 2003; Wagenmakers, Ratcliff, et al., 2008). The diffusion model generally provides an excellent fit to all aspects of the observed RT data, including response accuracy and the RT distributions of both correct and error responses. As indicated by Ratcliff and McKoon (2008, p. 918), ". . . the class of diffusion models has as near to provided a solution to simple decision making as is possible in behavioral science."

One of the major strengths of the diffusion model is its ability to provide parameter estimates that can be interpreted in terms of the cognitive components underlying the decision process (see, e.g., Voss et al., 2004). The central parameters of the model are drift rate v, boundary separation a, starting point z, and nondecision time T_{er}. Drift rate v represents the mean rate of information accumulation and is determined by the quality of information that is extracted from the stimulus. Drift rate can be influenced either by individual differences in the quality of information processing or by stimulus characteristics that reflect task difficulty. Boundary separation a quantifies the distance between the two response boundaries and represents response caution. Large values of a indicate that more information must be accumulated before a decision can be made. Boundary separation is usually manipulated via speed–accuracy instructions. Starting point z represents participants' a priori bias for one of the two response alternatives. Starting point is usually manipulated either by varying the proportions of stimuli associated with the upper and lower response boundaries or by payoff manipulations. Both a and z are assumed to be under the subjective control of participants. Nondecision time T_{er} quantifies the duration of processes that are unrelated to the decision process, including stimulus encoding and response execution.

In addition to these key parameters, the diffusion model features parameters that describe how the key parameter values fluctuate from one trial to the next. Specifically, the model assumes across-trial variability in drift rate (according to a normal distribution with variance η), starting point (according to a uniform distribution with range s_z), and nondecision time (according to a uniform distribution with range s_t).[3]

To summarize, the diffusion model provides a general theoretical account of decision making in speeded two-choice tasks. Previous research has shown that the parameters of the model correspond to the psychological processes that they are assumed to represent (e.g., Ratcliff & McKoon, 2008; Voss et al., 2004; Wagenmakers, 2009). Therefore, the diffusion model can be used to judge whether the cognitive interpretation of the ex-Gaussian and shifted Wald parameters is warranted when these descriptive distributions are applied to data from speeded two-choice tasks. Ideally, the parameters of the ex-Gaussian and shifted Wald distributions would correspond uniquely with the parameters of the diffusion model. For instance, one would hope that, say, a change in drift rate in the diffusion model would correspond to
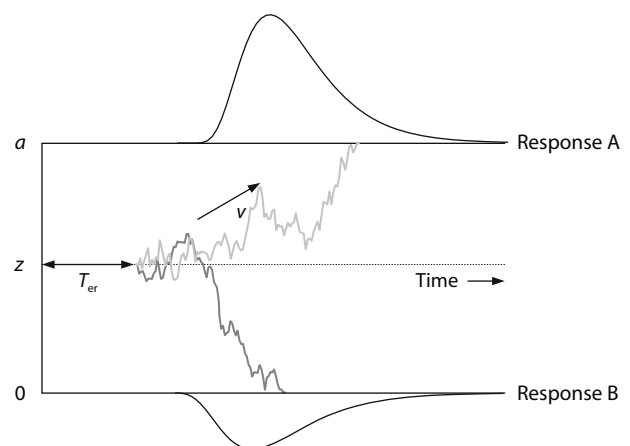

Figure 4. The diffusion model and its key parameters. See the text for details.

a change in the τ parameter in the ex-Gaussian and the γ parameter in the shifted Wald.

Links between the ex-Gaussian and diffusion model parameters. Several attempts have been made to relate the ex-Gaussian parameters to those of the diffusion model. For instance, Schmiedek, Oberauer, Wilhelm, Süß, and Wittmann (2007) showed that both the ex-Gaussian parameter τ and the diffusion model parameter v correlated strongly with people's higher cognitive functions, such as working memory and reasoning. In addition, Schmiedek et al. demonstrated by simulations that the relation between τ and higher cognitive functions could be fully explained in terms of individual differences in drift rate v. Schmiedek et al. concluded that τ is associated with v.

Other researchers have adopted a different approach, using simulations to examine changes in the ex-Gaussian parameters as a result of changes in the diffusion model drift rate and boundary separation parameters. The results have typically shown that (1) an increase in drift rate mainly causes a decrease in τ and (2) an increase in boundary separation mainly causes an increase in μ (see, e.g., Spieler, 2001; Spieler, Balota, & Faust, 2000; Yap et al., 2006).

Finally, Ratcliff (1978) used the ex-Gaussian μ and τ parameters to fit the diffusion model to data obtained from various experimental paradigms, such as the study–test paradigm (e.g., Ratcliff & Murdock, 1976), the Sternberg paradigm (e.g., Sternberg, 1966), and the continuous recognition memory paradigm (e.g., Okada, 1971). Contrary to the simulation results above, Ratcliff found that μ and τ are both sensitive to changes in drift rate and boundary separation. In particular, the results indicated that (1) increases in drift rate and starting point cause decreases in both μ and τ, and (2) an increase in boundary separation causes increases in both μ and τ.

The results above are therefore far from conclusive. Some studies (e.g., Spieler, 2001) have reported that τ and μ are selectively influenced by drift rate v and boundary separation a, respectively. Other studies (e.g., Ratcliff, 1978), however, have shown that τ and μ are sensitive to changes in a variety of diffusion model parameters. In addition, the previous work has examined only a limited range of values for the diffusion model parameters. A comprehensive investigation will require that the diffusion model parameters be manipulated on a realistic and sufficiently large range.

Links between the shifted Wald and diffusion model parameters. To the best of our knowledge, no one has yet attempted to relate the shifted Wald parameters to those of the diffusion model. Nevertheless, both the shifted Wald distribution and the diffusion model conceptualize the decision process as a gradual process of information accumulation. In fact, the shifted Wald can be thought of as a single-boundary diffusion process (cf. Figures 2 and 4). On the basis of the conceptual similarities between the two models, one might expect that (1) an increase in drift rate v would mainly cause an increase in γ, (2) an increase in boundary separation a would mainly cause an increase in α, (3) an increase in starting point z

would mainly cause a decrease in α, and (4) an increase in nondecison time T_{er} would mainly cause an increase in θ.

Empirical evidence for some of these relations has been reported by Schwarz (2001). Schwarz (2001) used a go/no-go digit comparison task and manipulated numerical distance (1 or 4) and the prior probability of go trials (.5 or .75). Shifted Wald analyses by Heathcote (2004) confirmed that the manipulation of numerical distance selectively influenced the Wald drift rate γ and that the manipulation of prior probability selectively influenced the Wald response criterion α. As expected, the Wald nondecision time θ was not influenced by either of these two manipulations. These results are encouraging, but it remains unclear to what extent the parameters of the shifted Wald will correspond to those of the diffusion model when data are obtained in a speeded task that clearly features two response alternatives.

Validation of the Ex-Gaussian and Shifted Wald Parameters Using Diffusion Model Simulations

In this section, we investigate the associations between parameters from the ex-Gaussian and shifted Wald distributions and from the diffusion model. To this end, we simulated data from the diffusion model by systematically varying its parameter values. Next, we fitted both the ex-Gaussian and the shifted Wald distributions to the simulated data sets.

Diffusion model simulations and model fitting. In each simulation, we generated data by manipulating a particular diffusion model parameter from a minimum to a maximum value while keeping the other parameters constant on their average values. Realistic parameter values were based on an extensive literature survey that covered 23 diffusion model applications. Table 3 shows the minimum, maximum, and mean values of the parameters used in the simulations. Note that some of the estimates from which the minimum, maximum, and mean values are derived result from parameter estimation with theoretically motivated constraints on some of the diffusion model parameters. For the histograms of the diffusion model parameter values found in the literature, the reader is referred to Appendix A.[4]

For the manipulation of boundary separation a, starting point z was assumed to be equidistant from the two response boundaries, so that $z = a/2$. Furthermore, the manipulation of starting point z was carried out with respect to the mean value of boundary separation a by using the minimum and maximum values of the z/a ratios, the so-called *bias parameters*, found in the literature. A z/a ratio of .5 indicates that starting point z is equidistant from the two response boundaries. Similarly, the manipulation of the parameter for trial-to-trial variability in starting point, s_z, was carried out with respect to the mean value of a by using the minimum and maximum values of the s_z/a ratios. Each parameter was manipulated in 1,000 steps of equal size, resulting in 1,000 data sets per parameter. In order to obtain relatively noise-free parameter estimates, each data set contained 10,000 RTs. The simulations were carried out using the Diffusion Model Analysis Toolbox (Vandekerckhove & Tuerlinckx, 2007, 2008).

<div style="text-align:center">

Table 3
Minimum, Maximum, and Mean Values
of the Diffusion Model Parameters Used in the Simulations

</div>

Diffusion Model Parameter	Minimum	Maximum	Mean
Drift rate v	0.0	0.586	0.223
Boundary separation a	0.056	0.393	0.125
Starting point z	0.028	0.182	0.063
Nondecision time T_{er}	0.206	0.942	0.435
Trial-to-trial variability in drift rate η	0.0	0.329	0.133
Trial-to-trial variability in starting point s_z	0.0	0.169	0.037
Trial-to-trial variability in nondecision time s_t	0.0	0.630	0.183
Bias z/a	0.272	0.782	–
s_z/a	0.0	0.900	–

Next, the ex-Gaussian and shifted Wald distributions were fitted to the simulated data sets using maximum likelihood estimation (see, e.g., Myung, 2003). Extreme parameter estimates (i.e., 15 ex-Gaussian and 8 shifted Wald estimates) were removed from the analyses. Note that the descriptive distributions were fitted to the RTs of correct responses only.

Simulation results. Figures 5 and 6 show the changes in the ex-Gaussian and shifted Wald parameters as a function of changes in the diffusion model parameters. Table 4 gives a summary of the associations between the two sets of parameters. In this section, we present only the results related to the manipulation of the key diffusion model parameters: drift rate v, boundary separation a, starting point z, and nondecision time T_{er}. Because the across-trial variability parameters cannot be interpreted in terms of cognitive processes, the results related to these parameters are presented in Appendix B.

Ex-Gaussian parameters. With respect to drift rate v, Figure 5A shows that the three ex-Gaussian parameters all decrease as v increases. The decreases in both μ and σ are, however, extremely small. In fact, changes in v are primarily reflected in τ. Also, τ continues to decrease until extreme values of v, whereas μ and σ level off at intermediate values of v.

Turning to boundary separation a, Figure 5B shows that the three ex-Gaussian parameters all increase as a increases. Although τ increases more than either μ or σ, the increase in μ is also substantial. Note that τ changes substantially more as a function of a than as a function of any other diffusion model parameter.

With respect to starting point z, Figure 5C shows that the three ex-Gaussian parameters all decrease as z increases. However, the decreases in both σ and τ are negligible. Also, τ seems relatively constant for low values of z. Changes in starting point z are thus primarily reflected in μ.

Turning to nondecision time T_{er}, Figure 5D shows that μ increases as T_{er} increases. In contrast, both the τ and the σ parameters are unaffected by T_{er}. Note that μ changes substantially more as a function of T_{er} than as a function of any other diffusion model parameter.

To summarize, the results of the simulations indicate that the ex-Gaussian parameters do not correspond uniquely to the parameters of the diffusion model. The μ parameter is substantially influenced by boundary separation a, starting point z, and nondecision time T_{er}. The

σ parameter is not influenced substantially by any of the key diffusion model parameters, and τ is substantially influenced by both drift rate v and boundary separation a.

Shifted Wald parameters. With respect to drift rate v, Figure 6A shows that both α and γ increase as v increases. In contrast, θ seems to decrease with increasing v. However, the decrease in θ is extremely small. In fact, changes in v are primarily reflected in α and γ. Note that the three shifted Wald parameters all level off for high values of v and that α and θ are also relatively constant for low values of v.

Turning to boundary separation a, Figure 6B shows that γ decreases and θ increases with increasing a. Quite unexpectedly, the response criterion α parameter *de*creases as a *in*creases. The changes in all three parameters are substantial. Note that the rates of decrease in both α and γ slow down as a increases. In fact, α levels off at intermediate values of a.

With respect to starting point z, Figure 6C shows that both α and θ decrease as z increases. In contrast, γ decreases for lower values and increases for higher values of z. The point of reversal in γ seems to correspond to the point at which z is equidistant to the two response boundaries. However, the changes in both θ and γ are extremely small. Changes in z are thus primarily reflected in α.

Turning to nondecision time T_{er}, Figure 6D shows that γ is unresponsive to changes in T_{er}. In contrast, both α and θ increase as T_{er} increases. However, the increase in α is negligible and limited to high values of T_{er}. Changes in T_{er} are thus primarily reflected in θ. In fact, θ changes substantially more as a function of T_{er} than as a function of any other diffusion model parameter.

To summarize, the results of the simulations indicate that the shifted Wald parameters likewise do not correspond uniquely to the parameters of the diffusion model. The α parameter is substantially influenced by drift rate v, boundary separation a, and starting point z. Remarkably, α decreases as a increases. The θ parameter is substantially influenced by both boundary separation a and nondecision time T_{er}. Finally, the γ parameter is substantially influenced by both v and a.

Validation of the Ex-Gaussian and Shifted Wald Parameters Using Experimental Manipulations

In this section, we present a concrete empirical illustration of the simulation results reported above by examin-
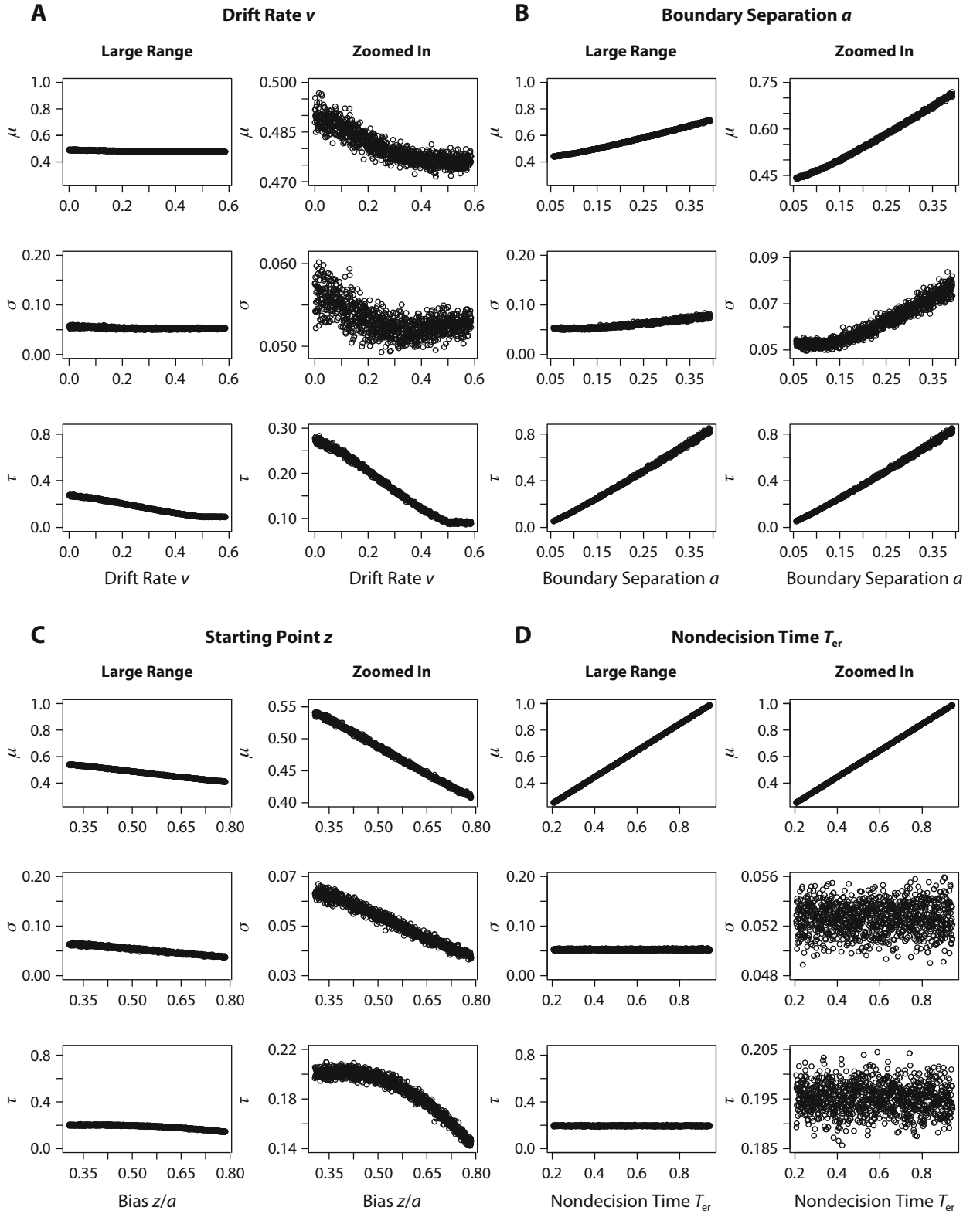
Figure 5. Changes in the ex-Gaussian parameters μ, σ, and τ as a function of systematic changes in the diffusion model parameters drift rate v (A), boundary separation a (B), starting point z (C), and nondecision time T_{er} (D). The left-hand figures in each panel plot the results on scales ranging from the minimum to the maximum values of the ex-Gaussian parameters found across all simulations. The right-hand figures in each panel plot the same results on scales ranging from the minimum to the maximum values of the ex-Gaussian parameters found for manipulations of the given diffusion model parameter.
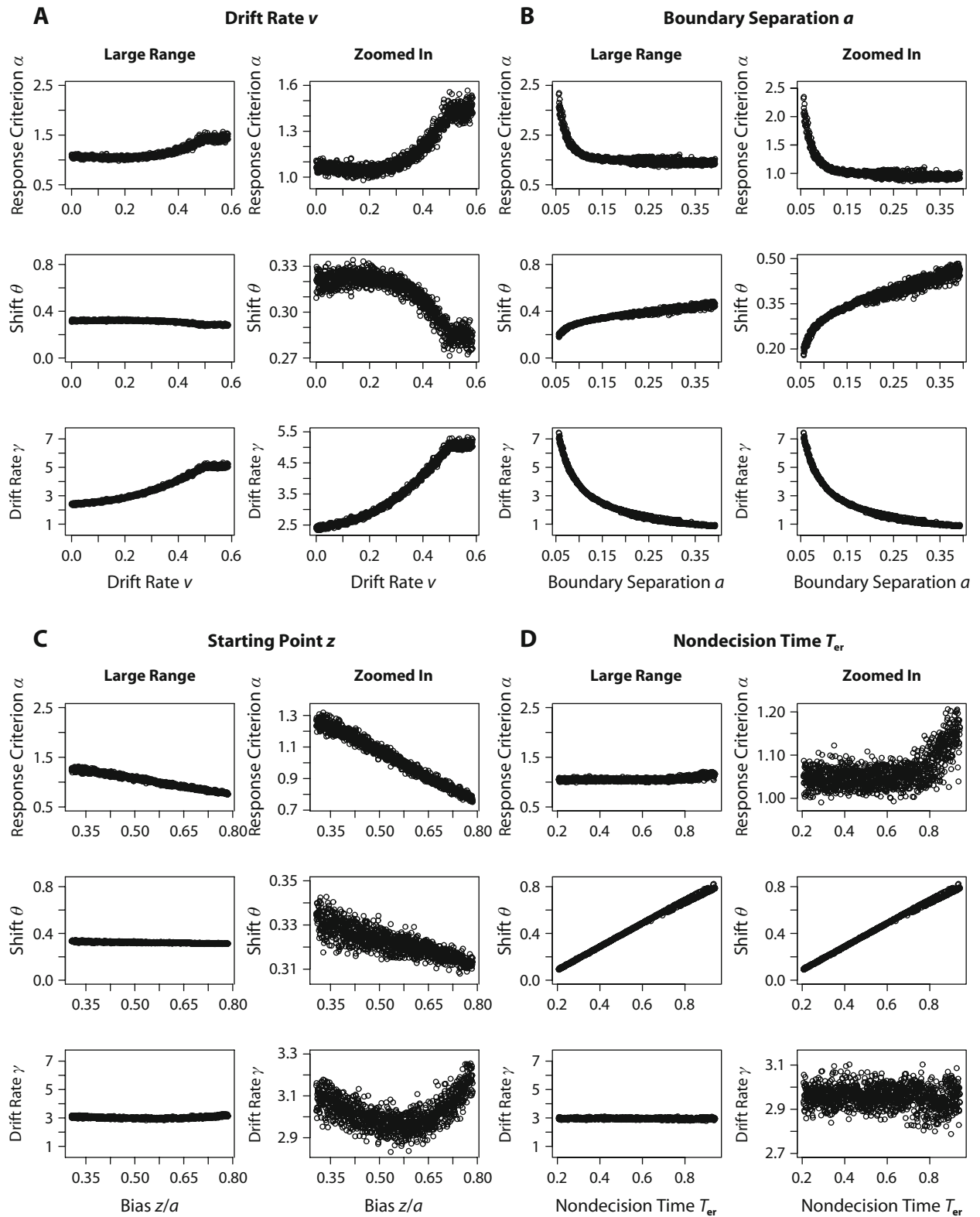
Figure 6. Changes in the shifted Wald parameters α, θ, and γ as a function of systematic changes in the diffusion model parameters drift rate v (A), boundary separation a (B), starting point z (C), and nondecision time T_{er} (D). The left-hand figures in each panel plot the results on scales ranging from the minimum to the maximum values of the shifted Wald parameters found across all simulations. The right-hand figures in each panel plot the same results on scales ranging from the minimum to the maximum values of the shifted Wald parameters found for manipulations of the given diffusion model parameter.

Table 4
The Associations Between Parameters of the Ex-Gaussian
and Shifted Wald Distributions and Parameters
of the Diffusion Model

		Diffusion Model Parameters			
		v	a	z	T_{er}
Ex-Gaussian parameters	μ	−	+ +	− −	+ +
	σ	−	+	−	×
	τ	− −	+ +	−	×
Shifted Wald parameters	α	+ +	− −	− −	+
	θ	−	+ +	−	+ +
	γ	+ +	− −	−/+	×

Note—+ +, substantial positive association; +, weak positive association; − −, substantial negative association; −, weak negative association; ×, no association; v, drift rate; a, boundary separation; z, starting point; T_{er}, nondecision time.

ing how the parameters of the descriptive distributions relate to experimentally induced changes in the diffusion model parameters. Specifically, we investigate how the ex-Gaussian and shifted Wald parameters respond to experimental manipulations that selectively affect the key parameters of the diffusion model (i.e., drift rate v, boundary separation a, and starting point z). To this end, we fitted the ex-Gaussian and shifted Wald distributions to data sets obtained from two lexical decision experiments.

The lexical decision data. To separately estimate the effects of lexical processing from the effects of strategic threshold adjustments, Wagenmakers, Ratcliff, et al. (2008) applied the diffusion model to the data of two lexical decision experiments. In the first experiment ($N = 15$), task difficulty was manipulated on three levels by varying word frequency (high, low, and very low frequency), and response caution was manipulated on two levels by instructions and feedback that emphasized either response speed or response accuracy. The resulting 3 (word frequency) × 2 (speed–accuracy instruction) cells of the experimental design each contained 160 trials per participant. The diffusion model was able to account for the effects of the manipulations with only two parameters free to vary across conditions. The effects of word frequency were entirely accounted for by changes in drift rate v, with higher word frequency associated with higher values of v. In contrast, the effects of the speed–accuracy instructions were entirely accounted for by changes in boundary separation a, with speed instructions associated with lower values of a.

In the second experiment ($N = 19$), in addition to the task difficulty manipulation, participants' a priori bias was manipulated on two levels by varying the proportion of word versus nonword stimuli in a list (i.e., 75% words or 75% nonwords). The resulting 3 (word frequency) × 2 (word–nonword proportion) cells of the experimental design each contained 160 trials per participant. As in Experiment 1, the effects of word frequency were entirely accounted for by changes in drift rate v. In contrast, the effects of the proportion manipulation were accounted for by changes in starting point z, with the 75%-word condition associated with higher values of z (in the model-

ing, the upper and lower boundaries were associated with "word" and "nonword" responses, respectively).

We fitted this data set using the ex-Gaussian and shifted Wald distributions and examined how their parameters relate to the experimental manipulation of drift rate v, boundary separation a, and starting point z. We expected that the pattern of association between the two sets of parameters would largely follow the pattern found in our simulations. The empirical results, however, are unlikely to precisely mirror the results of the simulations. Although the effects of the experimental manipulations were adequately accounted for by changes in the above-mentioned diffusion model parameters, the manipulations might not have had *completely* selective influence on these parameters. Hence, changes in the ex-Gaussian and shifted Wald parameters as a function of the experimental manipulations might reflect slight changes in diffusion model parameters other than the intended ones.

Hierarchical Bayesian modeling. We used hierarchical Bayesian modeling (see, e.g., Farrell & Ludwig, 2008; Gelman & Hill, 2007; Rouder, Lu, Speckman, Sun, & Jiang, 2005; Rouder, Sun, Speckman, Lu, & Zhou, 2003; Shiffrin, Lee, Kim, & Wagenmakers, 2008) to fit the ex-Gaussian and shifted Wald distributions to the lexical decision data. We used a hierarchical Bayesian approach to fit the descriptive distributions because the individual participant data obtained from the lexical decision tasks were considerably noisier than the synthetic data used in the previous section. As shown by Farrell and Ludwig (2008) and Rouder et al. (2005), hierarchical Bayesian methods reduce the variability in the recovered parameters and produce more accurate parameter estimates than single-level maximum likelihood estimation.

The hierarchical Bayesian approach assumes that the parameters of individual participants are drawn from group-level distributions that specify how the individual parameters are distributed in the population. The group-level distributions thus define the between-subjects variations of the parameters and can themselves be characterized by a set of parameters. For example, suppose that the RT data of each participant are assumed to come from an ex-Gaussian distribution, but with different values of μ, σ, and τ. The individual participant parameters μ_i, σ_i, and τ_i might in turn be assumed to come from normal distributions with means m, s, and t and variances s_m^2, s_s^2, and s_t^2, respectively. The benefits of hierarchical modeling arise from using the group-level distributions as priors to adjust extreme individual parameter estimates to more moderate values. In summary, hierarchical Bayesian modeling involves a

> . . . tension between fitting each subject as well as possible (optimal choice of individual parameters) and fitting the group as a whole. . . . This tension results in a movement of the individual parameters toward the group mean, a desirable characteristic given that we do not desire to overfit the data, and fit the noise in each individual's data. (Shiffrin et al., 2008, p. 1261)

$m \sim \text{Uniform}(0.25, 0.99)$

$s_m \sim \text{Uniform}(0, 0.214)$

$s \sim \text{Uniform}(0.02, 0.18)$

$s_s \sim \text{Uniform}(0, 0.046)$

$t \sim \text{Uniform}(0.05, 0.85)$

$s_t \sim \text{Uniform}(0, 0.231)$

$\mu_i \sim \text{Gaussian}(m, s_m^2)$

$\sigma_i \sim \text{Gaussian}(s, s_s^2)$

$\tau_i \sim \text{Gaussian}(t, s_t^2)$

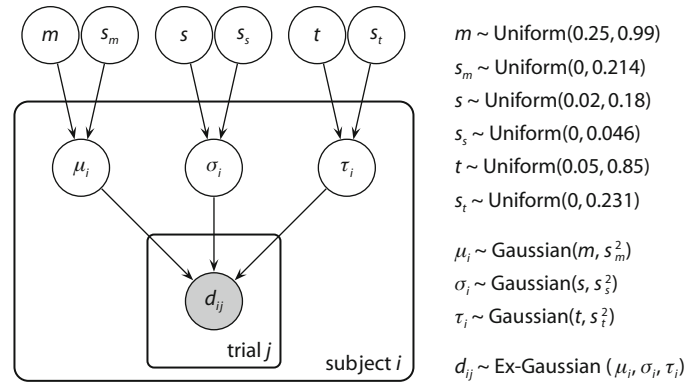$d_{ij} \sim \text{Ex-Gaussian}(\mu_i, \sigma_i, \tau_i)$

Figure 7. Graphical model of the hierarchical ex-Gaussian analysis. Note that the ranges of the uniform prior distributions for the group means are based on the minimum and maximum values of the corresponding ex-Gaussian parameters found in the simulation study reported above.

Figures 7 and 8 show graphical models for the hierarchical ex-Gaussian and shifted Wald analyses reported in this section. The nodes represent variables of interest, and the graph structure is used to indicate dependencies between the variables, with children depending on their parents. We use the convention of representing unobserved variables without shading and observed variables with shading (see, e.g., Lee, 2008). Figure 7 shows that the ex-Gaussian parameters μ_i, σ_i, and τ_i vary from participant to participant and are assumed to be drawn from group-level normal distributions with means m, s, and t, respectively. Similarly, Figure 8 shows that the shifted Wald parameters α_i, θ_i, and γ_i vary from participant to participant and are assumed to be drawn from group-level normal distributions with means a, h, and g, respectively.

The ranges of the uniform prior distributions for the group means are based on the minimum and maximum values of the corresponding ex-Gaussian and shifted Wald parameters found in the simulation study reported above.

The uniform prior distributions for the group standard deviations range from 0 to the standard deviations of the uniform priors for the corresponding group means. For example, the uniform prior for the ex-Gaussian group mean m ranges from 0.25 to 0.99; values of μ more extreme than this did not occur in our earlier diffusion model simulation study. The uniform prior for the associated group standard deviation s_m ranges from 0 to 0.214. The latter value is the maximum standard deviation for a unimodal distribution on m—that is, the standard deviation for a uniform distribution on m [i.e., $(0.99 - 0.25)/\sqrt{12} \approx 0.214$].

The starting values for the hierarchical Bayesian analysis were based on the individual parameter estimates.[5] At the beginning of each sampling run, the first 1,000 trials of the Markov chain Monte Carlo chains were discarded. Each analysis was based on 10,000 recorded samples. We used the WinBUGS program (Lunn, Thomas, Best, & Spiegelhalter, 2000) for parameter estimation.[6] Note that, similar to the procedure of Wagenmakers, Ratcliff, et al. (2008), the descriptive distributions were fitted only to

$a \sim \text{Uniform}(0.67, 2.35)$

$s_a \sim \text{Uniform}(0, 0.485)$

$h \sim \text{Uniform}(0, 0.82)$

$s_h \sim \text{Uniform}(0, 0.237)$

$g \sim \text{Uniform}(0.85, 7.43)$

$s_g \sim \text{Uniform}(0, 1.899)$

$\alpha_i \sim \text{Gaussian}(a, s_a^2)$

$\theta_i \sim \text{Gaussian}(h, s_h^2)$

$\gamma_i \sim \text{Gaussian}(g, s_g^2)$

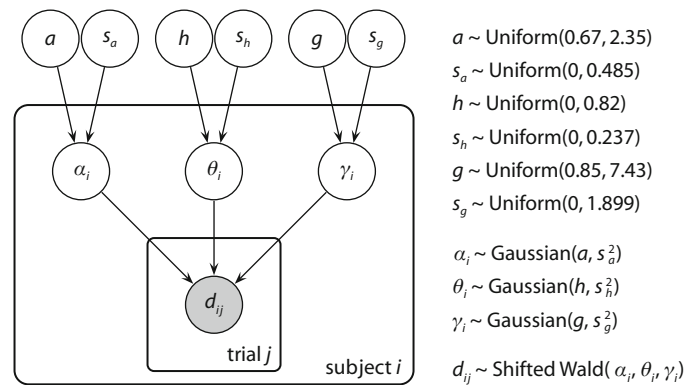$d_{ij} \sim \text{Shifted Wald}(\alpha_i, \theta_i, \gamma_i)$

Figure 8. Graphical model for the hierarchical shifted Wald analysis. Note that the ranges of the uniform prior distributions for the group means are based on the minimum and maximum values of the corresponding shifted Wald parameters found in the simulation study reported above.
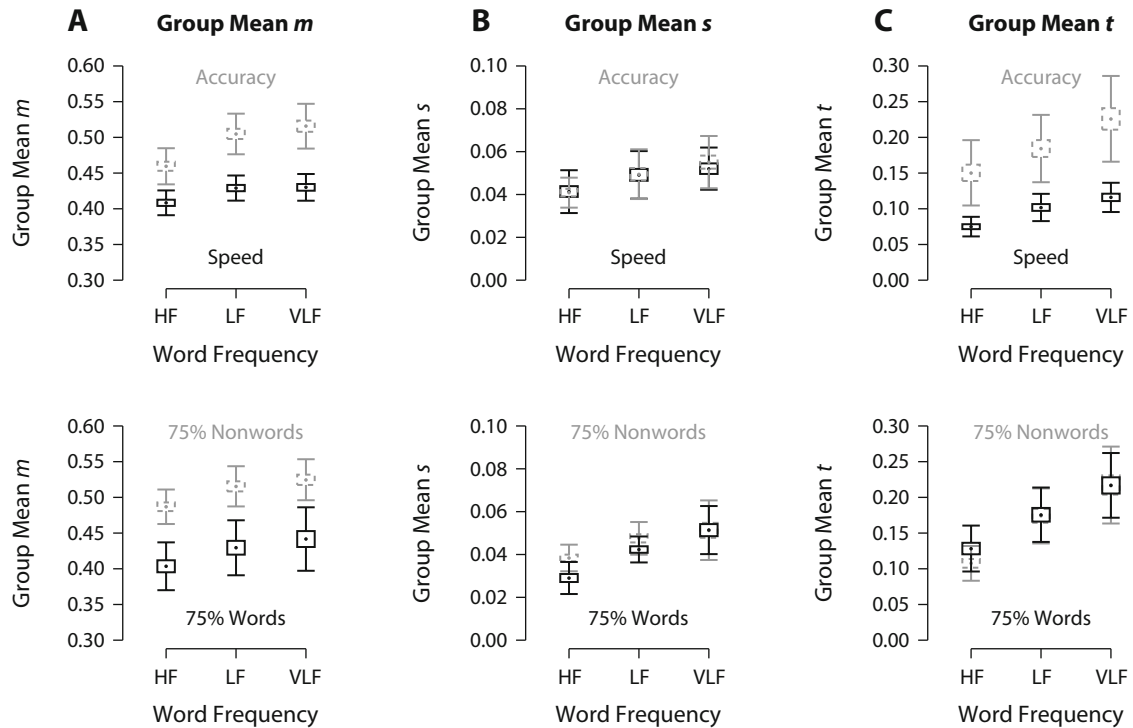
Figure 9. Box plots of the posterior distributions for the ex-Gaussian group means *m*, *s*, and *t*, derived separately for each condition of the two lexical decision experiments of Wagenmakers, Ratcliff, et al. (2008). HF, high-frequency words; LF, low-frequency words; VLF, very-low-frequency words.

the RTs of correct word responses that were slower than 300 msec and faster than 2,500 msec.

Results. Figure 9 shows box plots of the posterior distributions for the ex-Gaussian group mean parameters *m*, *s*, and *t*. Figure 10 shows box plots of the posterior distributions for the shifted Wald group mean parameters *a*, *h*, and *g*. Our discussion of the results is based on a visual inspection of the posterior distributions.

Ex-Gaussian parameters. With respect to *m* (i.e., the group mean for the μ parameter), Figure 9A shows that *m* increases when instructions emphasize choice accuracy. Since the effects of the speed–accuracy manipulation can be accounted for by changes in boundary separation *a*, this result suggests that μ increases with increasing boundary separation. Similarly, *m* increases when the stimuli consist of 75% nonwords. Since the effects of the proportion manipulation can be accounted for by changes in starting point *z*, this result suggests that μ increases with decreasing starting point. The effect of the word frequency manipulation is less clear—*m* increases from high-frequency to low-frequency words, but does not change considerably from low-frequency to very-low-frequency words. Since the effects of the word frequency manipulation can be accounted for by changes in drift rate *v*, this result suggests that μ increases slightly with decreasing drift rate.

With respect to *s* (i.e., the group mean for the σ parameter), Figure 9B shows that *s* is influenced by the effects of neither the speed–accuracy instructions nor the proportion manipulation. These results suggest that σ is not influenced by either boundary separation *a* or starting point *z*.

Again, the effect of the word frequency manipulation is less clear—*s* increases somewhat from high-frequency to low-frequency words but does not change considerably from low-frequency to very-low-frequency words. This finding suggests that σ increases slightly with decreasing drift rate *v*. Note, however, that the increase in σ is relatively small.

Turning to *t* (i.e., the group mean for the τ parameter), Figure 9C shows that *t* increases with decreasing word frequency and when the instructions emphasize choice accuracy. In contrast, *t* is unresponsive to the effects of the proportion manipulation. These results suggest that τ increases with increasing boundary separation *a* and with decreasing drift rate *v* but is unaffected by changes in starting point *z*.

In summary, the results above indicate that the ex-Gaussian parameters do not respond selectively to the effects of the word frequency, speed–accuracy, and proportion manipulations. Consistent with the diffusion model simulations reported above, these results suggest that the ex-Gaussian parameters do not correspond uniquely to the drift rate *v*, boundary separation *a*, and starting point *z* parameters of the diffusion model. The μ parameter is sensitive to changes in all three diffusion model parameters. Although σ seems to be influenced only by drift rate, this influence is relatively small. Finally, τ is sensitive to changes in both drift rate and boundary separation. These results indicate that changes in the two most important ex-Gaussian parameters, μ and τ, can reflect changes in a variety of diffusion model parameters.
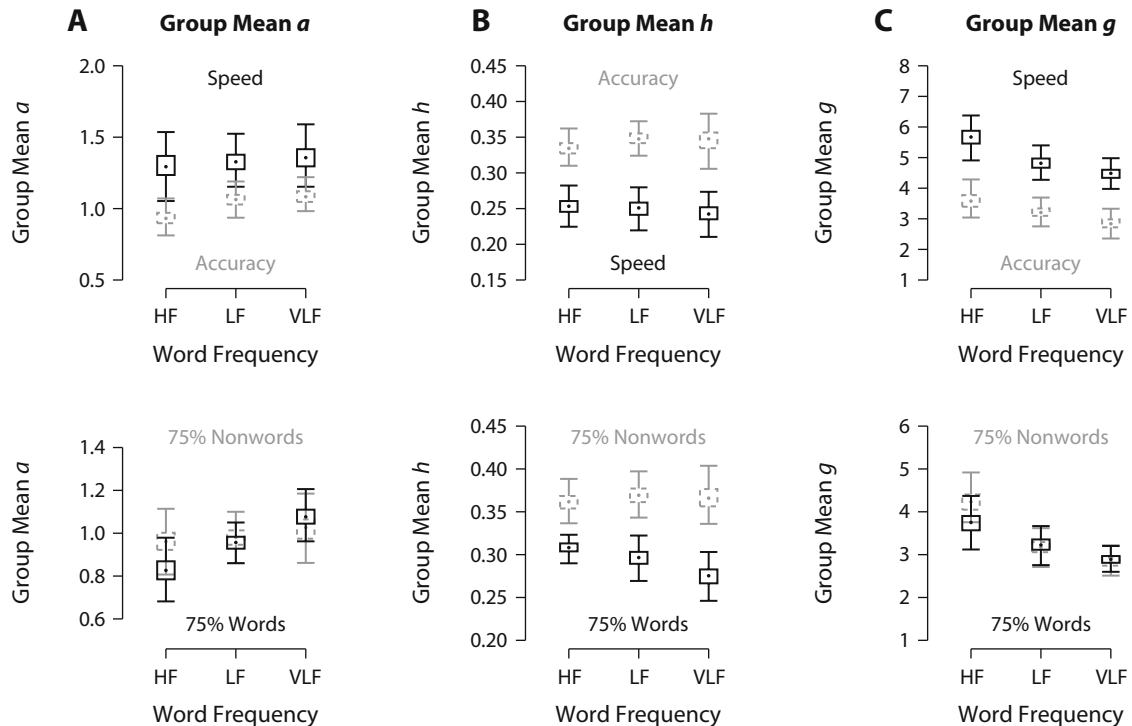
Figure 10. Box plots of the posterior distributions for the shifted Wald group means *a*, *h*, and *g*, derived separately for each condition of the two lexical decision experiments of Wagenmakers, Ratcliff, et al. (2008). HF, high-frequency words; LF, low-frequency words; VLF, very-low-frequency words.

Shifted Wald parameters. With respect to *a* (i.e., the group mean for the α parameter), Figure 10A shows that group mean *a increases* when instructions emphasize fast responding. In contrast, it seems to be unresponsive to the effects of the proportion manipulation. These results suggest that α increases with decreasing boundary separation *a* and is unaffected by changes in starting point *z*. The effect of word frequency is less clear—group mean *a* increases somewhat with decreasing word frequency when the stimuli consist of 75% words but is relatively constant in the other conditions. This result suggests that, under certain conditions, α increases slightly with decreasing drift rate.

Turning to *h* (i.e., the group mean for the θ parameter), Figure 10B shows that *h* increases when the instructions emphasize choice accuracy and when the stimuli consist of 75% nonwords. These results suggest that θ increases with increasing boundary separation *a* and with decreasing starting point *z*. Again, the effect of the word frequency manipulation is less clear—*h* seems to decrease with decreasing word frequency when the stimuli consist of 75% words but is relatively constant in the other conditions. This result suggests that, under certain conditions, θ decreases with decreasing drift rate *v*.

With respect to *g* (i.e., the group mean for the γ parameter), Figure 10C shows that *g* increases when instructions emphasize fast responding and decreases with decreasing word frequency. In contrast, it seems unresponsive to the effects of the proportion manipulation. These results suggest that γ increases with decreasing boundary separa-

tion *a*, decreases with decreasing drift rate *v*, and is unaffected by changes in starting point *z*.

To summarize, the results above indicate that the shifted Wald parameters also do not respond selectively to the effects of the word frequency, speed–accuracy, and proportion manipulations. Consistent with the diffusion model simulations reported above, these results suggest that the shifted Wald parameters do not correspond uniquely to the drift rate *v*, boundary separation *a*, and starting point *z* parameters of the diffusion model. The α and the γ parameters are responsive to changes in both boundary separation *a* and drift rate *v*, and θ is influenced by all three diffusion model parameters. These results indicate that changes in the shifted Wald parameters can reflect changes in a diversity of diffusion model parameters.

Discussion

The goal of this study was to examine the extent to which the ex-Gaussian and shifted Wald parameters could be associated with the kind of psychological processes that are hypothesized in the diffusion model, one of the most successful process models of speeded two-choice decision making. First, we generated synthetic data by systematically manipulating the diffusion model parameters to examine the associated changes in the parameters of the ex-Gaussian and shifted Wald distributions. Second, we investigated empirical data and studied how the ex-Gaussian and shifted Wald parameters relate to the experimental manipulation of the diffusion model parameters drift rate *v*, boundary separation *a*, and starting point *z*. The results were clear cut: In the

context of a two-choice task, the ex-Gaussian and shifted Wald parameters cannot be associated uniquely with the parameters of the diffusion model.

The ex-Gaussian distribution. Like the results of Ratcliff (1978), our simulation results demonstrated that the two most important ex-Gaussian parameters, μ and τ, are sensitive to changes in a variety of diffusion model parameters. Specifically, μ was influenced by boundary separation a, starting point z, and nondecision time T_{er}. On the other hand, τ was sensitive to changes in both drift rate v and boundary separation a. The results related to experimental manipulations of the diffusion model parameters followed a similar pattern. The only discrepancy was that μ also appeared to be influenced by the effects of the word frequency manipulation, suggesting that it was also sensitive to changes in drift rate v. This difference notwithstanding, the results indicate that the ex-Gaussian parameters do not correspond uniquely to those of the diffusion model.

The shifted Wald distribution. The results of the simulations indicated that the shifted Wald parameters also could not be uniquely associated with parameters of the diffusion model. Each of the shifted Wald parameters appeared to be sensitive to changes in a diversity of diffusion model parameters. The α parameter was sensitive to changes in drift rate v, boundary separation a, and starting point z. Surprisingly, α decreased as boundary separation increased. The θ parameter was affected by both boundary separation a and nondecision time T_{er}. Finally, the γ parameter was substantially influenced by both drift rate v and boundary separation a. The results related to experimental manipulations of the diffusion model parameters largely followed the same pattern. However, α was unresponsive to the effects of the proportion manipulation, suggesting that this parameter was unaffected by changes in starting point z. Also, α seemed to increase, rather than decrease, with decreasing word frequency, suggesting that it increased with decreasing drift rate v. Finally, θ was responsive to the effects of the word frequency and proportion manipulations, suggesting that it was also sensitive to changes in v and z. These differences notwithstanding, the results with experimental data also indicate that the shifted Wald parameters do not correspond uniquely to parameters of the diffusion model.

The finding that neither the simulation nor the experimental results support the interpretation of the shifted Wald parameters in terms of the psychological processes of participant ability/task difficulty, response caution, and nondecision time is disappointing and comes somewhat as a surprise. First, in view of the conceptual similarities of the shifted Wald distribution and the diffusion model, one might expect some correspondence between the two sets of parameters and their underlying cognitive processes. Yet none of our predictions derived from the theoretical similarities of the two models was supported by the results. Second, our results indicate that the differential sensitivity of the shifted Wald parameters found in the go/no-go task (Heathcote, 2004) does not generalize to tasks that involve two response boundaries. Instead, our results suggest that when the shifted Wald is applied to paradigms that involve more than a single response boundary, the

cognitive interpretation of the shifted Wald parameters no longer holds. It must further be noted that the results of Gomez, Ratcliff, and Perea (2007) suggest that the cognitive interpretation of the shifted Wald parameters might be problematic even when the distribution is applied to one-choice tasks. In particular, Gomez et al. showed that an adequate model of the go/no-go task must feature two response boundaries: one associated with the go response and another associated with the implicit choice not to respond (i.e., the no-go decision).

The effects of error rate, parameter correlations, and parameter combinations. Our results strongly suggest that the ex-Gaussian and shifted Wald parameters should not be interpreted in terms of the cognitive processes assumed by the diffusion model. Nevertheless, some issues warrant further discussion.

First, our method of data generation resulted in error rates ranging from 10% to 15% across the simulations. However, the shifted Wald distribution may perhaps be appropriate for paradigms that result in very few errors, such as tasks that involve saccadic eye movements (Carpenter & Williams, 1995). We therefore investigated how the ex-Gaussian and shifted Wald parameters change as a function of manipulation of the diffusion model parameters in data sets with lower (0.9%–4%) as well as with higher (19%–28%) error rates. Regardless of whether the error rate was low or high, the results clearly indicated that the ex-Gaussian and shifted Wald parameters cannot be associated uniquely with parameters of the diffusion model.[7]

Second, we generated data by manipulating each diffusion model parameter separately while holding the other parameters constant on their average values. Although this approach yields clear-cut and comprehensible results, it ignores the possible associations among the diffusion model parameters. We therefore investigated how changes in the ex-Gaussian and shifted Wald parameters relate to changes in the diffusion model parameters when we take into account the correlations between the latter parameters. The simulations indicated that using parameter sets with realistic parameter associations yields results that are noisier but qualitatively similar to those reported in the present article.

Finally, our results indicate that the individual ex-Gaussian and shifted Wald parameters cannot be mapped uniquely onto the parameters of the diffusion model. However, the parameters of the descriptive distributions need not be considered in isolation. Unlike the individual parameters, certain (nonlinear) combinations of the ex-Gaussian or shifted Wald parameters might map uniquely onto parameters of the diffusion model. This possibility awaits further investigation.

A common problem? Neither the ex-Gaussian nor the shifted Wald parameters appear to correspond to the psychological processes hypothesized by the diffusion model. A possible reason for this unfortunate result may be that neither of the two distributions takes into account response accuracy. Without any knowledge of response accuracy, it is very difficult to distinguish between effects of task difficulty (or participant ability) and effects of response caution. For example, does a decrease in RT come about because of an increase in drift rate v or a decrease in bound-

ary separation a? It is evident that in this case, a change in response accuracy is highly diagnostic; an increase in drift rate leads to fewer errors (i.e., an overall improvement), whereas a decrease in boundary separation leads to more errors (i.e., the speed–accuracy trade-off; see, e.g., Schouten & Bekker, 1967; Wickelgren, 1977). Because performance in RT tasks reflects the combined effects of task difficulty and response caution, a model that cannot separate these influences is unlikely to capture the cognitive processes that determine performance (Wagenmakers, van der Maas, & Grasman, 2007).

Conclusion. The present results indicate that—in the context of speeded two-alternative tasks—the ex-Gaussian and shifted Wald parameters should not be interpreted in terms of the cognitive processes hypothesized by the diffusion model. This does not imply that the ex-Gaussian and shifted Wald distributions should no longer be used as purely descriptive tools to economically summarize RT data and to constrain model development. Such descriptive use of the ex-Gaussian and shifted Wald distributions is perfectly legitimate and highly encouraged. What our findings do imply is that it may be ill-advised to attribute changes in the ex-Gaussian and shifted Wald parameters to changes in specific components of cognitive processing.

REFERENCES

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 514-544.

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General*, **128**, 32-55.

Blough, D. S. (1988). Quantitative relations between visual search speed and target–distractor similarity. *Perception & Psychophysics*, **43**, 57-71.

Blough, D. S. (1989). Contrast as seen in visual search reaction times. *Journal of the Experimental Analysis of Behavior*, **52**, 199-211.

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational & Graphical Statistics*, **7**, 434-455.

Burbeck, S. L., & Luce, R. D. (1982). Evidence from auditory simple reaction times for both change and level detectors. *Perception & Psychophysics*, **32**, 117-133.

Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, **377**, 59-62.

Emerson, P. L. (1970). Simple reaction time with Markovian evolution of Gaussian discriminal processes. *Psychometrika*, **35**, 99-110.

Epstein, J. N., Conners, C. K., Hervey, A. S., Tonev, S. T., Arnold, L. E., Abikoff, H. B., et al. (2006). Assessing medication effects in the MTA study using neuropsychological outcomes. *Journal of Child Psychology & Psychiatry*, **47**, 446-456.

Farrell, S., & Ludwig, C. J. H. (2008). Bayesian and maximum likelihood estimation of hierarchical response time models. *Psychonomic Bulletin & Review*, **15**, 1209-1217.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge: Cambridge University Press.

Gholson, B., & Hohle, R. H. (1968a). Choice reaction times to hues printed in conflicting hue names and nonsense words. *Journal of Experimental Psychology*, **76**, 413-418.

Gholson, B., & Hohle, R. (1968b). Verbal reaction times to hues and hue names and forms and form names. *Perception & Psychophysics*, **3**, 191-196.

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, **136**, 389-413.

Gordon, B., & Carson, K. (1990). The basis for choice reaction time slowing in Alzheimer's disease. *Brain & Cognition*, **13**, 148-166.

Heathcote, A. (2004). Fitting Wald and ex-Wald distributions to response time data: An example using functions for the S-PLUS package. *Behavior Research Methods, Instruments, & Computers*, **36**, 678-694.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, **109**, 340-347.

Hockley, W. E. (1982). Retrieval processes in continuous recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 497-512.

Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 598-615.

Hohle, R. H. (1965). Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology*, **69**, 382-386.

Kieffaber, P. D., Kappenman, E. S., Bodkins, M., Shekhar, A., O'Donnell, B. F., & Hetrick, W. P. (2006). Switch and maintenance of task set in schizophrenia. *Schizophrenia Research*, **84**, 345-358.

Lee, M. D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin & Review*, **15**, 1-15.

Leth-Steensen, C., King Elbaz, Z., & Douglas, V. I. (2000). Mean response times, variability and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica*, **104**, 167-190.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS—A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics & Computing*, **10**, 325-337.

Madden, D. J., Gottlob, L. R., Denny, L. L., Turkington, T. G., Provenzale, J. M., Hawk, T. C., & Coleman, R. E. (1999). Aging and recognition memory: Changes in regional cerebral blood flow associated with components of reaction time distributions. *Journal of Cognitive Neuroscience*, **11**, 511-520.

McGill, W. [J.] (1963). Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 309-360). New York: Wiley.

McGill, W. J., & Gibbon, J. (1965). The general-gamma distribution and reaction times. *Journal of Mathematical Psychology*, **2**, 1-18.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, **47**, 90-100.

Okada, R. (1971). Decision latencies in short-term recognition memory. *Journal of Experimental Psychology*, **90**, 27-32.

Penner-Wilger, M., Leth-Steensen, C., & LeFevre, J.-A. (2002). Decomposing the problem-size effect: A comparison of response time distributions across cultures. *Memory & Cognition*, **30**, 1160-1167.

Plourde, C. E., & Besner, D. (1997). On the locus of the word frequency effect in visual word recognition. *Canadian Journal of Experimental Psychology*, **51**, 181-194.

Possamaï, C.-A. (1991). A responding hand effect in a simple-RT precueing experiment: Evidence for a late locus of facilitation. *Acta Psychologica*, **77**, 47-63.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59-108.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, **114**, 510-532.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review*, **9**, 278-291.

RATCLIFF, R., GOMEZ, P., & MCKOON, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, **111**, 159-182.

RATCLIFF, R., & MCKOON, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, **20**, 873-922.

RATCLIFF, R., & MURDOCK, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, **83**, 190-214.

RATCLIFF, R., & ROUDER, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 127-140.

RATCLIFF, R., THAPAR, A., GOMEZ, P., & MCKOON, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology & Aging*, **19**, 278-289.

RATCLIFF, R., THAPAR, A., & MCKOON, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology & Aging*, **16**, 323-341.

RATCLIFF, R., THAPAR, A., & MCKOON, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & Psychophysics*, **65**, 523-535.

RATCLIFF, R., THAPAR, A., & MCKOON, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory & Language*, **50**, 408-424.

ROHRER, D. (1996). On the relative and absolute strength of a memory trace. *Memory & Cognition*, **24**, 188-201.

ROHRER, D. (2002). The breadth of memory search. *Memory*, **10**, 291-301.

ROHRER, D., & WIXTED, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition*, **22**, 511-524.

ROTELLO, C. M., & ZENG, M. (2008). Analysis of RT distributions in the remember–know paradigm. *Psychonomic Bulletin & Review*, **15**, 825-832.

ROUDER, J. N., LU, J., SPECKMAN, P., SUN, D., & JIANG, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review*, **12**, 195-223.

ROUDER, J. N., SUN, D., SPECKMAN, P. L., LU, J., & ZHOU, D. (2003). A hierarchical Bayesian statistical framework for response time distributions. *Psychometrika*, **68**, 589-606.

SCHMIEDEK, F., OBERAUER, K., WILHELM, O., SÜβ, 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**, 414-429.

SCHOUTEN, J. F., & BEKKER, J. A. M. (1967). Reaction time and accuracy. *Acta Psychologica*, **27**, 143-153.

SCHWARZ, W. (2001). The ex-Wald distribution as a descriptive model of response times. *Behavior Research Methods, Instruments, & Computers*, **33**, 457-469.

SCHWARZ, W. (2002). On the convolution of inverse Gaussian and exponential random variables. *Communications in Statistics: Theory & Methods*, **31**, 2113-2121.

SHIFFRIN, R. M., LEE, M. D., KIM, W., & WAGENMAKERS, E.-J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science*, **32**, 1248-1284.

SMITH, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review*, **102**, 567-593.

SPIELER, D. H. (2001). Modelling age-related changes in information processing. *European Journal of Cognitive Psychology*, **13**, 217-234.

SPIELER, D. H., BALOTA, D. A., & FAUST, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 461-479.

SPIELER, D. H., BALOTA, D. A., & FAUST, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 506-526.

STERNBERG, S. (1966). High-speed scanning in human memory. *Science*, **153**, 652-654.

THAPAR, A., RATCLIFF, R., & MCKOON, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology & Aging*, **18**, 415-429.

TOWNSEND, J. T., & ASHBY, F. G. (1983). *The stochastic modeling of elementary psychological processes.* Cambridge: Cambridge University Press.

VANDEKERCKHOVE, J., & TUERLINCKX, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, **14**, 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., & VOSS, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, **32**, 1206-1220.

WAGENMAKERS, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, **21**, 641-671.

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 & Language*, **58**, 140-159.

WAGENMAKERS, E.-J., VAN DER MAAS, H. L. J., DOLAN, C. V., & GRASMAN, R. P. P. P. (2008). EZ does it! Extensions of the EZ-diffusion model. *Psychonomic Bulletin & Review*, **15**, 1229-1235.

WAGENMAKERS, E.-J., VAN DER MAAS, H. L. J., & GRASMAN, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, **14**, 3-22.

WALD, A. (1947). *Sequential analysis.* New York: Wiley.

WICKELGREN, W. A. (1977). Speed–accuracy tradeoff and information processing dynamics. *Acta Psychologica*, **41**, 67-85.

WIXTED, J. T., GHADISHA, H., & VERA, R. (1997). Recall latency following pure- and mixed-strength lists: A direct test of the relative strength model of free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 523-538.

WIXTED, J. T., & ROHRER, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1024-1039.

YAP, M. J., & BALOTA, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **33**, 274-296.

YAP, M. J., BALOTA, D. A., CORTESE, M. J., & WATSON, J. M. (2006). Single- versus dual-process models of lexical decision performance: Insights from response time distributional analysis. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 1324-1344.

NOTES

1. An extensive overview of the effects of experimental manipulations on the ex-Gaussian parameters, including interaction effects, is available in the supplemental materials, downloadable from http://pbr.psychonomic-journals.org/content/supplemental.

2. Specific quotations are available in the supplemental materials.

3. The model also features a parameter s that quantifies the amplitude of the noise in the information accumulation process. This scaling parameter is usually fixed at 0.1, a convention that we adhere to throughout the present article.

4. The exact parameter values are available in the supplemental materials.

5. We later confirmed our results by using overdispersed starting values and multiple chains to obtain an $\hat{R}$ statistic of about 1, indicating that the chains converged to the stationary distribution (Brooks & Gelman, 1998).

6. The WinBUGS code and the lexical decision data are available in the supplemental materials.

7. The results of these additional simulations are available in the supplemental materials.

SUPPLEMENTAL MATERIALS

The online supplement for this article contains the following materials: an overview of the effects of experimental manipulations on the ex-Gaussian parameters, with quotations relating the ex-Gaussian parameters to cognitive processes; the parameter values found in earlier applications of the diffusion model; the full lexical decision data; WinBUGS codes used to implement the hierarchical Bayesian analysis; and results related to the effects of error rate. These materials may be downloaded from http://pbr.psychonomic-journals.org/content/supplemental.

APPENDIX A
THE DISTRIBUTION OF THE DIFFUSION MODEL PARAMETER VALUES

Figure A1 presents histograms of the best-fitting diffusion model parameter values and the corresponding z/a and s_z/a ratios found in 23 applications of the diffusion model. The histograms are based on the parameter values reported for each experimental condition of the 23 articles. The exact parameter values, including references, are available at http://pbr.psychonomic-journals.org/content/supplemental.
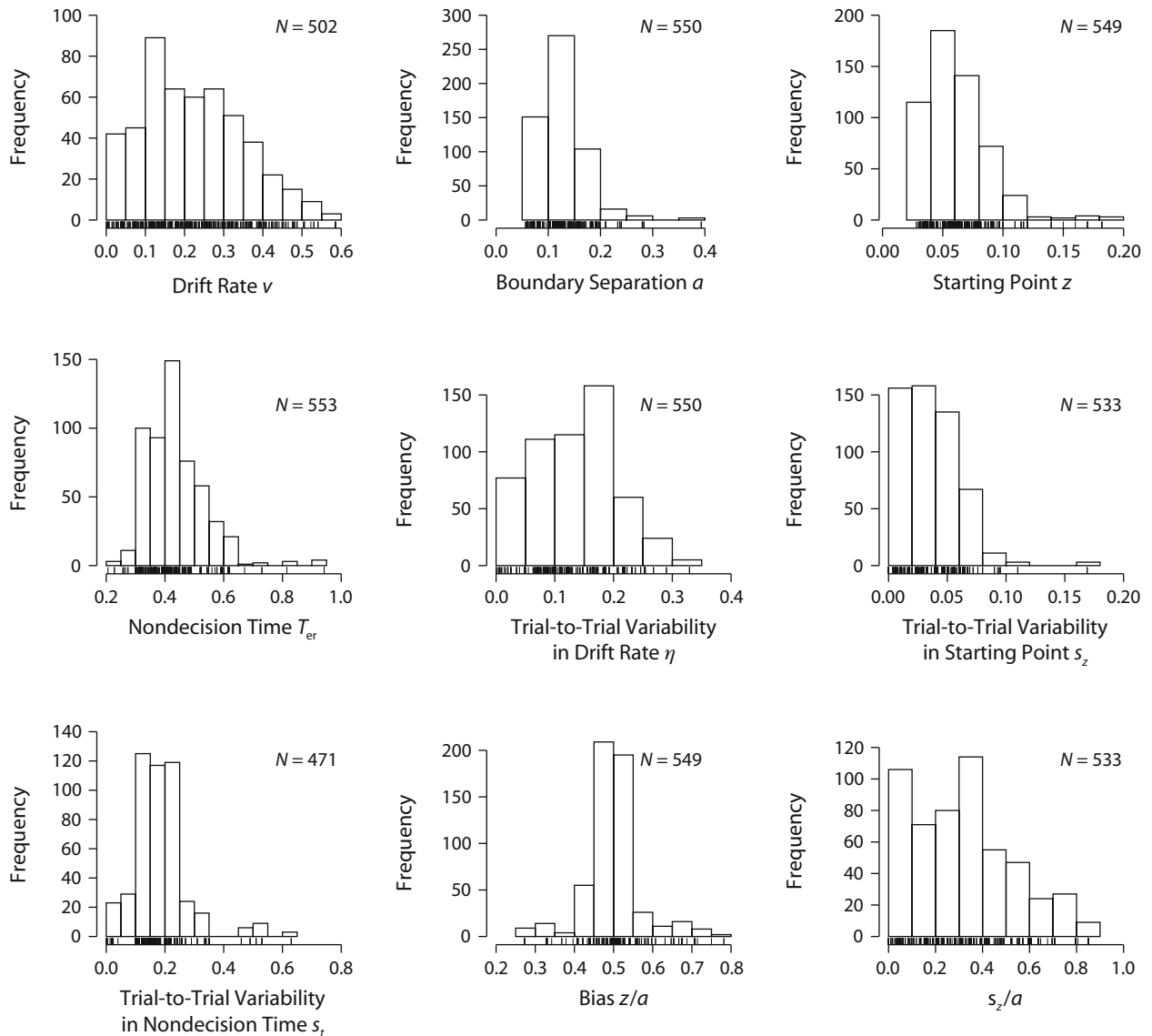

Figure A1. Histograms of the diffusion model parameter values.

APPENDIX B
RESULTS FOR THE DIFFUSION MODEL TRIAL-TO-TRIAL VARIABILITY PARAMETERS

This appendix shows how the ex-Gaussian and shifted Wald parameters change as a function of the manipulation of the trial-to-trial variability in drift rate η, the trial-to-trial variability in starting point s_z, and the trial-to-trial variability in nondecision time s_t parameters of the diffusion model. Table B1 gives a summary of the associations between the ex-Gaussian and shifted Wald parameters and the diffusion model variability parameters. Figures B1 and B2 then show the detailed changes in the ex-Gaussian and shifted Wald parameters as a function of changes in the diffusion model variability parameters.

Ex-Gaussian Parameters

With respect to trial-to-trial variability in drift rate η, Figure B1A shows that both μ and σ decrease as η increases. In, contrast, τ increases for low values of η and decreases for high values of η. However, the changes in the three ex-Gaussian parameters are all extremely small. Turning to trial-to-trial variability in starting point s_z, Figure B1B shows that both σ and the τ increase as s_z increases, but in contrast, μ decreases with increasing s_z. However, the changes in the three ex-Gaussian parameters are all negligible. With respect to trial-to-trial variability in nondecision time s_t, Figure B1C shows that both σ and τ increase as s_t increases, whereas μ decreases with increasing s_t. Note that σ is the only parameter that is substantially influenced by s_t. In fact, σ changes substantially more as function of s_t than as a function of any other diffusion model parameter.

To summarize, these results further support the conclusion that the two most important parameters of the ex-Gaussian distribution, μ and τ, do not correspond uniquely to parameters of the diffusion model. Neither of these ex-Gaussian parameters is influenced substantially by any of the variability parameters of the diffusion model. In contrast, σ seems to be uniquely associated with s_t, the parameter for trial-to-trial variability in nondecision time.

Shifted Wald Parameters

With respect to trial-to-trial variability in drift rate (η), Figure B2A shows that both α and γ decrease as η increases, but in contrast, θ increases with increasing η. However, the changes in the three shifted Wald parameters are extremely small. Turning to trial-to-trial variability in starting point (s_z), Figure B2B shows that γ is unaffected by changes in s_z, whereas α increases and θ decreases with increasing s_z. However, the changes in both α and θ are extremely small. With respect to trial-to-trial variability in nondecision time (s_t), Figure B2C shows that both α and γ increase for low and intermediate values of s_t and then decrease for high values. In contrast, θ decreases for low and intermediate values of s_t and equals 0 for high values. Although α changes more than either θ or γ, the change in θ is also substantial. Note that α changes just as much as a function of s_t as it does as a function of boundary separation a.

To summarize, these results further support the conclusion that the shifted Wald parameters do not correspond uniquely to parameters of the diffusion model. The γ parameter is not influenced substantially by any of the variability parameters of the diffusion model. In contrast, both α and θ are substantially influenced by s_t, the trial-to-trial variability in nondecision time. In addition to the influence of the key diffusion model parameters, changes in α and θ can therefore also reflect the influence of s_t.

Table B1
Associations Between Parameters of the Ex-Gaussian
and Shifted Wald Distributions and the Variability
Parameters of the Diffusion Model

		Diffusion Model Parameters		
		η	s_z	s_t
Ex-Gaussian parameters	μ	$-$	$-$	$-$
	σ	$-$	$+$	$++$
	τ	$+/-$	$+$	$+$
Shifted Wald parameters	α	$-$	$+$	$++/--$
	θ	$+$	$-$	$--$
	γ	$-$	$\times$	$+/-$

Note—$++$, substantial positive association; $+$, weak positive association; $--$, substantial negative association; $-$, weak negative association; $\times$, no association; η, trial-to-trial variability in drift rate; s_z, trial-to-trial variability in starting point; s_t, trial-to-trial variability in nondecision time.

A   **Trial-to-Trial Variability in Drift Rate η**   **B**   **Trial-to-Trial Variability in Starting Point s_z**

Large Range   **Zoomed In**   **Large Range**   **Zoomed In**

Trial-to-Trial Variability in Drift Rate η

Trial-to-Trial Variability in Starting Point s_z

C   **Trial-to-Trial Variability in Nondecision Time s_t**

Large Range   **Zoomed In**

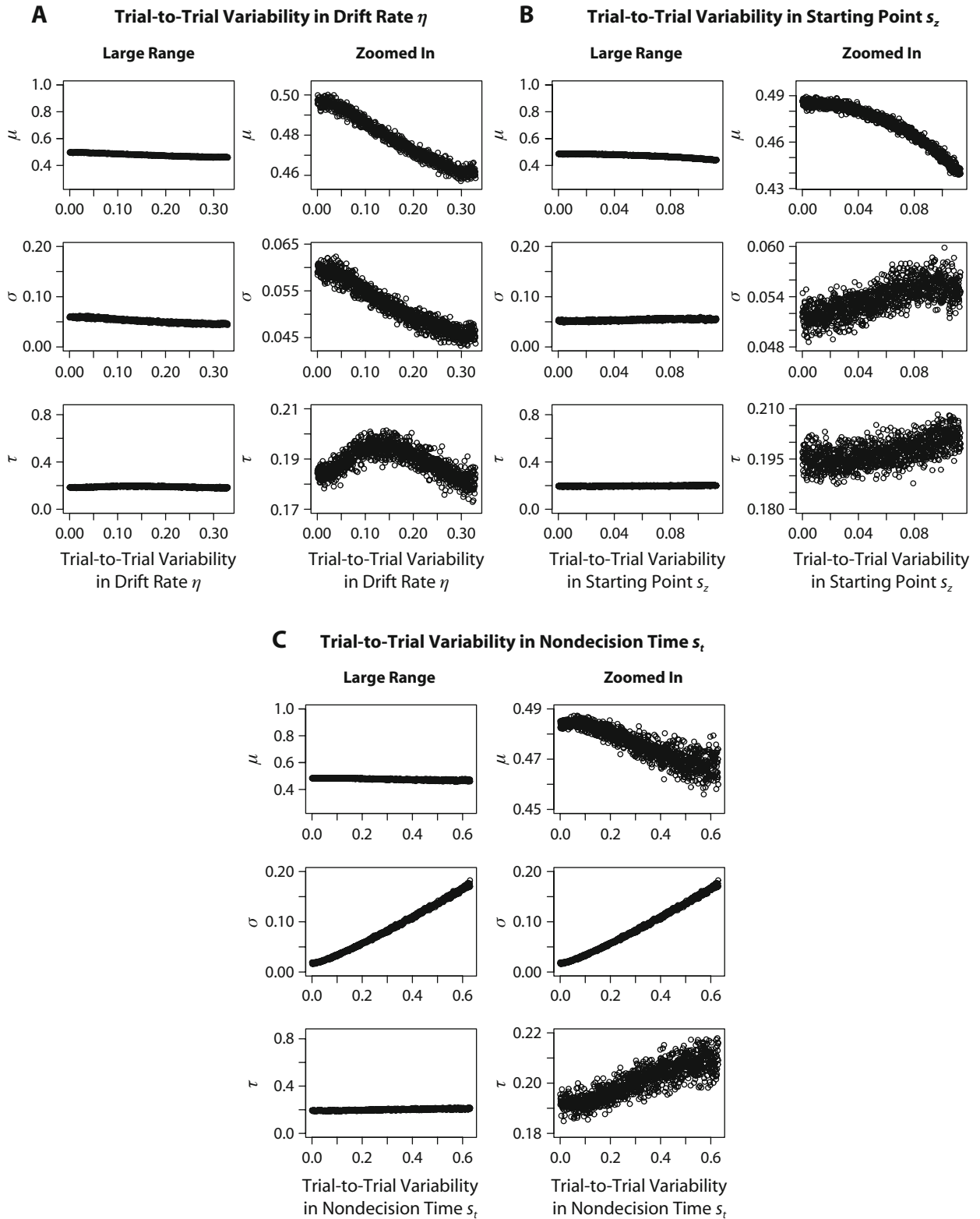Trial-to-Trial Variability in Nondecision Time s_t

Figure B1. Changes in the ex-Gaussian parameters μ, σ, and τ as a function of systematic changes in the diffusion model parameters trial-to-trial variability in drift rate η (A), trial-to-trial variability in starting point s_z (B), and trial-to-trial variability in nondecision time s_t (C). The left-hand figures in each panel plot the results on scales ranging from the minimum to the maximum values of the ex-Gaussian parameters found across all simulations. The right-hand figures in each panel plot the same results on scales ranging from the minimum to the maximum values of the ex-Gaussian parameters found for manipulations of the given diffusion model parameter.
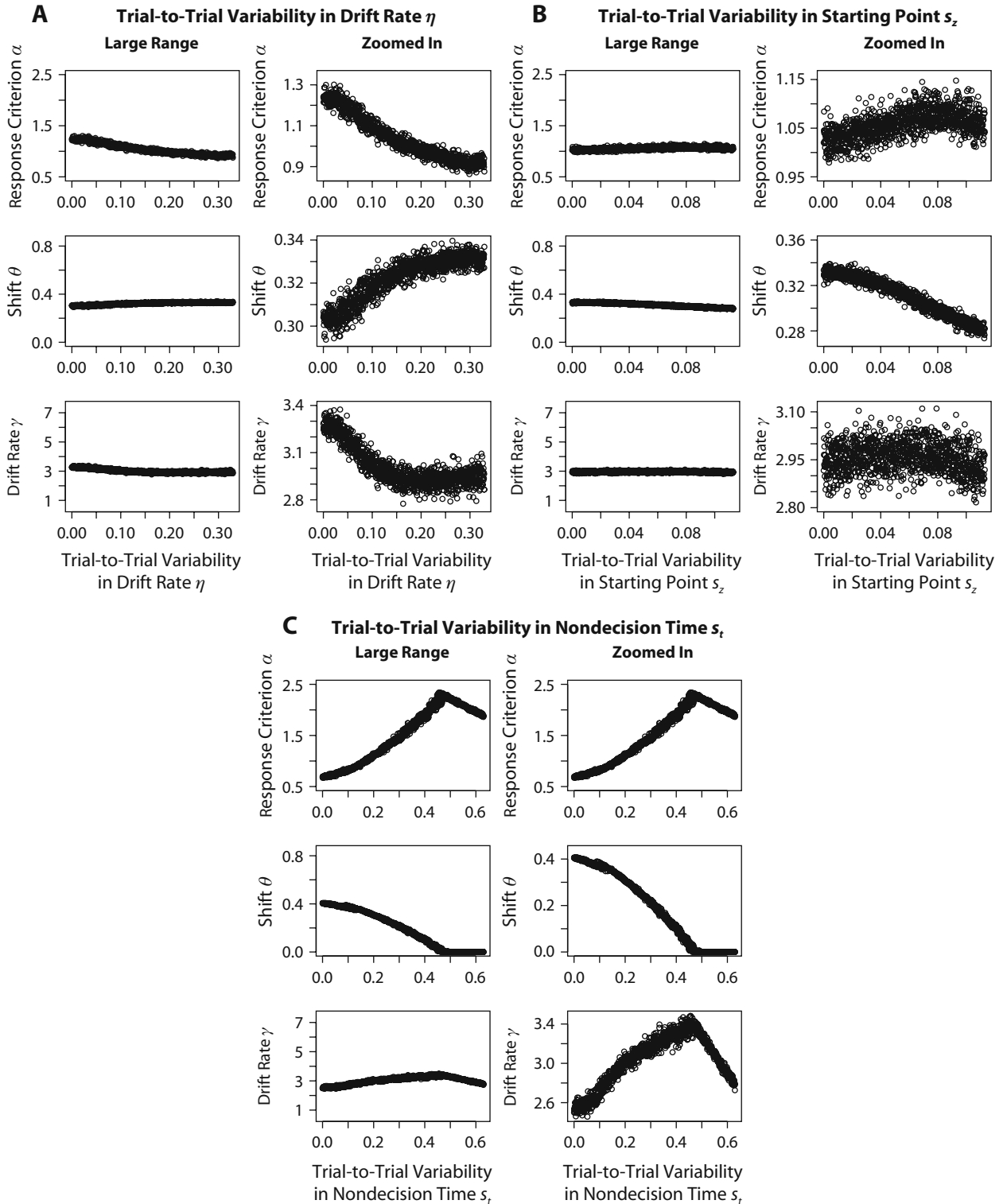
APPENDIX B (Continued)

Figure B2. Changes in the shifted Wald parameters α, θ, and γ as a function of systematic changes in the diffusion model parameters trial-to-trial variability in drift rate η (A), trial-to-trial variability in starting point s_z (B), and trial-to-trial variability in nondecision time s_t (C). The left-hand figures in each panel plot the results on scales ranging from the minimum to the maximum values of the shifted Wald parameters found across all simulations. The right-hand figures in each panel plot the same results on scales ranging from the minimum to the maximum values of the shifted Wald parameters found for manipulations of the given diffusion model parameter.