Psychon Bull Rev (2012) 19:555–574
DOI 10.3758/s13423-012-0254-6

THEORETICAL REVIEW

Response time models of delta plots with negative-going slopes

Wolf Schwarz · Jeff Miller

Published online: 18 May 2012

Abstract Delta plots (DPs) graphically compare reaction time (RT) quantiles obtained under two experimental conditions. In some research areas (e.g., Simon effects), decreasing delta plots (nDPs) have consistently been found, indicating that the experimental effect is largest at low quantiles and decreases for higher quantiles. nDPs are unusual and intriguing: They imply that RT in the faster condition is more variable, a pattern predicted by few standard RT models. We describe and analyze five classes of well-established latency mechanisms that are consistent with nDPs—exhaustive processing models, correlated stage models, mixture models, cascade models, and parallel channels models—and discuss the implications of our analyses for the interpretation of DPs. DPs generally do not imply any specific processing model; therefore, it is more fruitful to start from a specific quantitative model and to compare the DP it predicts with empirical data.

Keywords Delta plot · RT models · Simon effect · Activation suppression model

Chronometric research in cognitive psychology often seeks to infer something about the number, nature, and temporal organization of basic information-processing components by looking at how reaction time (RT) varies across two or more experimental conditions. Classic examples of this approach using mean RT are Sternberg's (1969) additive factor method in memory scanning and Treisman and Gelade's (1980) analysis of the slope of RT versus set size functions in visual search.

Increasingly, researchers have augmented the study of mean RT with studies of RT distributions. First, in some experimental paradigms, it is possible to distinguish among different classes of models on the basis of their specific predictions concerning between-condition effects on RT distributions as well as means, allowing model classes to be tested via distributional comparisons (e.g., Meyer, Irwin, Osman, & Kounios, 1988; Miller, 1982; Pashler, 1994b; Ratcliff & Smith, 2004; Ruthruff, 1996; Sigman & Dehaene, 2005). Second, some models predict how the exact shapes of RT distributions should be affected by certain experimental manipulations, and these models can be assessed in terms of the adequacy of their fits to the observed distributions (e.g., Grice, 1972; Logan, 1992; Ratcliff & Rouder, 2000). Third, at a purely descriptive level, RT distributions contain far more information than is adequately summarized by their means, so it may be possible to find empirical differences among different experimental manipulations that all have the same effect on mean RT (e.g., Balota, Yap, Cortese, & Watson, 2008; Heathcote, Popiel, & Mewhort, 1991; Townsend & Ashby, 1983).

In this article, we consider in detail one method for describing the effects of an experimental manipulation on RT distributions, the so-called *delta plot* (for short, DP), introduced by De Jong, Liang, and Lauber (1994). As will be developed more formally in the next section, the DP is a graphical comparison between two RT distributions that focuses on the question of whether the experimental manipulation has a larger effect on the relatively fast responses or on the relatively slow ones.

W. Schwarz (✉)
Department of Psychology, University of Potsdam,
P.O. Box 60 15 53, 14415 Potsdam-Golm, Germany
e-mail: wschwarz@uni-potsdam.de

J. Miller
Department of Psychology, University of Otago,
Dunedin, New Zealand
e-mail: miller@psy.otago.ac.nz

The analysis of DPs is intriguing for two reasons. First, although most experimental manipulations have a larger effect on relatively slow responses than on relatively fast responses, a few experimental manipulations are exceptions, producing just the opposite pattern. The best-documented of these exceptions is the standard Simon task (for recent reviews, see Pratte, Rouder, Morey, & Feng, 2010; Proctor, Miles, & Baroni, 2011; van den Wildenberg et al., 2010). In this task, participants respond by pressing keys to the left or right of the midline on the basis of a relevant visual stimulus attribute such as color (e.g., respond to green stimuli with the left hand and to red ones with the right hand). Irrelevantly, the stimulus appears randomly on the left or right side of the fixation point. The standard finding is that responses are faster when the stimulus appears on the same side as the hand that is required to respond (*congruent* trials) than when it appears on the opposite side (*incongruent* trials). For the present purposes, though, the most important aspect of this congruency effect is that it tends to be largest for fast or medium responses and small or absent for slow responses, contrary to the pattern observed with many other experimental manipulations in RT tasks. For reasons that will become apparent in the next section, we will refer to this unusual pattern as the finding of *DPs with negative-going slopes* (in short, nDPs).

Second, DP analysis is intriguing because the exceptional nDPs observed with Simon tasks are not predicted by standard RT models. As recently reviewed by Pratte et al. (2010), these models typically predict that the fastest responses should be affected less than—or, at most, equally with—the slowest responses. Similarly, standard neurocomputational modeling architectures typically produce increasing rather than decreasing DPs (e.g., Davelaar, 2008, Fig. 3). Thus, the opposite pattern observed with the standard Simon effect would seem to provide very specific clues about the nature of that effect. On the other hand, the exact nature of these specific clues is not immediately obvious, which suggests that a cautious approach is needed when interpreting DPs. Rather than asking which latency mechanism is implied by a particular DP, it might be more fruitful to ask, conversely: Which latency mechanisms are able to predict (i.e., are compatible with) specific forms of DPs?

The main purpose of this article is to describe several classes of quantitative models capable of producing nDPs like those that have been observed in Simon tasks. Clearly, the proper interpretation of nDPs requires knowledge of which information-processing models are compatible with this unusual DP shape. Our emphasis is, therefore, not on models of the Simon task per se but, rather, on the broader set of models that might be considered when this shape of nDPs is observed.

This article is organized as follows. First, by way of background, we present a formal description of the DP and review the basic statistical relationships that lead to nDPs. Second, also as background, we briefly review the *activation-suppression* model of Ridderinkhof (2002a, 2002b; Ridderinkhof, van den Wildenberg, Wijnen, & Burle, 2004), which is an influential account of nDPs in the context of standard Simon effects. No previous explanations of nDPs have been developed into formal quantitative models, though, so the exact conditions under which such accounts produce nDPs are at present unclear. In fact some of the quantitative models that we suggest might be considered as possible implementations of these, so a further goal of this article is to explore some possibilities for formal versions of the activation-suppression model.

Third, in the main section of the article, we describe several formal models that are capable of producing nDPs. These models are developed from principles that are familiar within RT modeling, and some could even be viewed as rather minor variants of existing RT models. It should be emphasized that we do not attempt to develop a detailed quantitative model for all aspects of the Simon effect per se or a detailed version of the activation-suppression model. Instead, we adopt a *breadth-first* search strategy to explore more generally what types of RT models are capable of producing nDPs. Even though it is never possible to identify all such models, the interpretation of any observed data pattern must be guided by some knowledge of the range of underlying mechanisms compatible with that pattern. Fourth, in the General Discussion section, we consider in more detail the implications of this range of options for current interpretations of effects producing an nDP.

The delta plot: Definition and basic properties

We first outline how empirical DPs are typically constructed. Consider two experimental conditions in which RTs are obtained, such as, for example, congruent and incongruent trials in a conflict (Ridderinkhof et al., 2004) or context (Pratte et al., 2010) task. A typical example is the study of Simon effects under rest and exercise by Davranche and McMorris (2009). As is illustrated in the upper panel of Fig. 1, a first step is to generate a frequency distribution of these RTs, separately for each condition $i = 1,2$, where we use the convention that conditions 1 and 2 are, on average, faster (e.g., congruent) and slower (e.g., incongruent), respectively. RTs are first rank ordered for each participant and condition separately and then are collected to form bins of equal area. As is shown in Fig. 1, Davranche and McMorris used ten bins; in the group average results shown, each bin contained 20 RTs per participant. Next, the mean RT within each bin is computed and used as an estimate of the

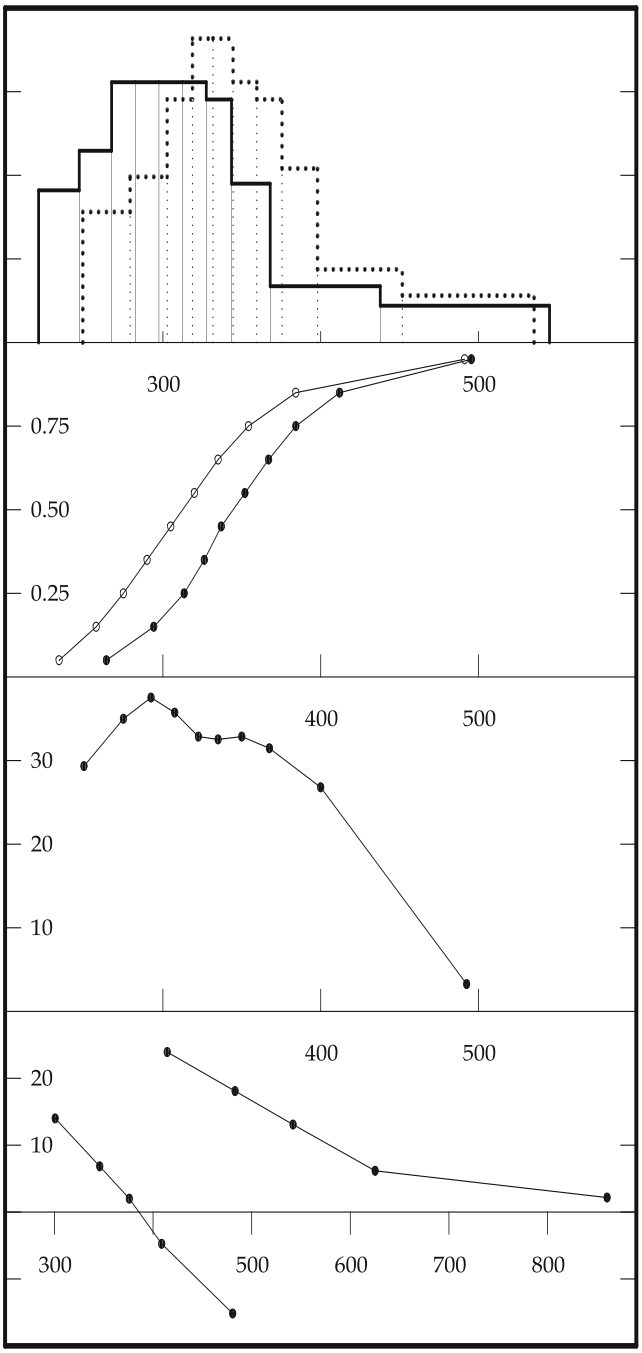

Fig. 1 *Top panel*: Binned reaction time (RT) frequency distributions for congruent (heavy solid line) and incongruent (heavy dotted line) conditions, from the data in Davranche and McMorris (2009, Fig. 4A, "rest" condition). The light solid and dotted lines indicate the bin edges. *Second panel*: Cumulative distribution functions (CDFs) for RT in the congruent (open circles) and incongruent (filled points) condition, as computed from the bins in the top panel. *Third panel*: Delta plot of quantile differences (ordinate) against quantile averages (abscissa), as computed from the CDFs in the second panel. *Bottom panel*: Delta plots of quantile differences (ordinate) against quantile averages (abscissa). The curve on the left shows data from Burle, van den Wildenberg, and Ridderinkhof (2005, Fig. 1), and the one on the right shows data from Davranche, Paleresompoulle, Pernaud, Labarelle, and Hasbroucq (2009, Fig. 1B)

corresponding RT quantile; for example, with ten bins, one obtains estimates of the 5 %, 15 %, . . . , 95 % quantiles. As is shown in the second panel of Fig. 1, the RT quantiles are then used to construct an empirical estimate of the two cumulative distribution functions (CDFs), denoted as $F_i(t)$, which, for each value of t, indicate the proportion of RTs smaller than or equal to t.

The DP, which is shown in the third panel of Fig. 1, is basically a visual comparison of the RT quantiles obtained under the two experimental conditions. In the empirical example shown in Fig. 1, the horizontal separation of the two empirical CDFs in the second panel is 30 ms for the initial bins of the two conditions and, so (by definition), is the value of the DP. This value first increases slightly to about 40 ms, which reflects an initial increase in the separation of the two empirical CDFs. Starting with the 35 % quantiles, this separation then gets smaller and eventually drops off to a value of just 5 ms for the difference of the estimate of the two 95 % quantiles. Therefore, the DP decreases over the region from about 300 to 500 ms. Recall that the pth RT quantile under condition i, $q_i(p)$, is that value of RT below which falls the proportion p of all RTs in that condition, $q_i(p) = F_i^{-1}(p)$, $i = 1,2$. More specifically, the DP is a parametric plot of the difference, or separation, of quantiles, $y(p) = q_2(p) - q_1(p)$, against the average of the quantiles, $x(p) = [q_1(p) + q_2(p)]/2$, as p is varied from zero to one. In practice, the two cumulative RT distributions from which a DP is generated are often first averaged by vincentizing the CDFs within each condition across participants (cf. Davranche & McMorris, 2009; Ratcliff, 1979). However, this aspect is not crucial for the construction of DPs per se, and in principle, a separate DP can be derived for each participant individually. Also, we notice that most recent studies (e.g., Kubo-Kawai & Kawai, 2010; Proctor, Pick, Vu, & Anderson, 2005; Vallesi, Mapelli, Schiff, Amodio, & Umiltà, 2005; Wascher, Schatz, Kuder, & Verleger, 2001; Wylie, Ridderinkhof et al., 2010) have derived DPs from correct-response RTs only but that some (e.g., Ridderinkhof, Scheres, Oosterlaan, & Sergeant, 2005) have used RTs from both correct and error trials (for a discussion of this point, see Burle, Possamaï, Vidal, Bonnet, & Hasbroucq, 2002; van den Wildenberg et al., 2010).

Although the concave form of the nDP shown in the third panel of Fig. 1 is quite typical of many reported nDPs from Simon tasks (cf. Vallesi et al., 2005, Table 2; Wylie et al., 2010, Fig. 4; Wylie, Ridderinkhof, Bashore, & van den Wildenberg, 2009, Fig. 3), nDPs with somewhat different shapes have been reported as well, as is illustrated by the empirical examples shown in the bottom panel of Fig. 1. Specifically, nDPs for Simon tasks have sometimes been reported to decrease linearly (cf. Burle, van den Wildenberg, & Ridderinkhof, 2005, Fig. 1; DeJong et al., 1994, Fig. 3) or even to be convex (they decrease, but at a decreasing rate;

cf. Davranche, Paleresompoulle, Pernaud, Labarelle, & Hasbroucq, 2009, Fig. 1). As is also illustrated in Fig. 1, decreasing DPs have sometimes been found to cross the abscissa, meaning that the largest RT quantiles are larger for the congruent than for the incongruent condition, a violation of stochastic dominance. Evidently, any evaluation of the statistical reliability of this feature must take the standard errors involved in estimating the largest quantiles into account, a point which we will address in our General Discussion section.

As can be seen in Fig. 1, for a DP to decrease as t increases, it is necessary for the early or medium quantiles of the distributions to be widely separated and for the later quantiles to differ by less and less. That is, the horizontal separation of the CDFs must get smaller and smaller as t increases. This qualitative pattern implies that the condition with the shorter RTs also shows the wider spread of RT or, more technically, that the shorter mean RT goes with the larger RT variance. Like the nDP itself, this pattern of larger mean plus smaller variance is quite remarkable because it violates the general finding of a *consistent ordering* of the RT moments—that is, that empirically, the smaller mean RT almost always goes with the smaller variance (for a detailed review, see Luce, 1986, chaps. 2, 11; Wagenmakers & Brown, 2007).

Interpretations of delta plots with negative-going slopes

The literature offers no clear consensus about exactly what nDPs imply about the underlying mechanisms generating them. A sceptical view is that of Zhang and Kornblum (1997), who emphasized the statistical properties of DPs. They pointed out that these plots are essentially a variant of the more traditional quantile–quantile (QQ) plot. Specifically, they showed that a DP is linear if and only if the two RT distributions being compared belong to a location-and-scale (LS) family—that is, if $F_i(t) = F[(t - \mu_i)/\sigma_i]$, where F is the standard form of the LS-family. In that case, the slope of the DP is independent of the location parameters μ_i and depends only on the spreads (i.e., the σ_i's) of the two RT distributions involved. Specifically, the slope is equal to the ratio of the difference of the scale parameters over the mean scale parameters—that is, equal to $(\sigma_2 - \sigma_1)/[(\sigma_1 + \sigma_2)/2]$. Therefore, if the DP is linear, a decrease simply reflects the fact that the condition with the shorter RT has the larger spread of RTs. In view of these results, Zhang and Kornblum concluded that the DP is a simple graphical display that "reflects the statistical properties of the pair of RT distributions and not necessarily functional hypotheses concerning processing mechanisms" (p. 1551). Although statistically impeccable, this conclusion does not in itself give any positive clues about the kinds of RT mechanisms that might produce the statistical properties necessary for nDPs. Thus, Zhang and Kornblum's point of view nicely highlights the major concern of the present article, which is to explore the types of processing mechanisms compatible with nDPs. It should also be noted that Zhang and Kornblum's results, important as they are, refer to the case in which the two RT distributions belong to one LS family and, thus, generate linear DPs. However, as is shown in Fig. 1, many empirical DPs in context tasks appear decidedly nonlinear, and the implication of Zhang and Kornblum's results for this case are not clear.

Various less sceptical and more neutral descriptions of nDPs essentially present them as convenient and interesting data summaries. For example, Kubo-Kawai and Kawai (2010) stated that "the Simon effect decreased as a function of the [increasing RT] bins" (p. 457). Similarly, Vallesi et al. (2005) summarized their DP analysis by stating that "the distribution for corresponding responses lies to the left of the non-corresponding distribution only in the faster part, indicating that the correspondence effect is present only when RTs are fast, and then vanishes" (p. B38). These summaries are little more than restatements of the observed CDFs—namely, that the congruency effect is relatively large for fast responses but relatively small or absent for slow ones. In view of Zhang and Kornblum's (1997) result, this seems rather uncontroversial, as long as it is taken only as a purely descriptive claim about how the congruency effect varies with RT.

More optimistically, DPs are sometimes seen to provide supporting evidence for specific information-processing models. In their original study demonstrating nDPs in the context of the Simon effect, for example, DeJong et al. (1994) claimed that their observed nDPs "almost certainly provide a reliable estimate of the actual time course of these effects" (p. 733). They argued that nDPs suggest a mechanism for inhibiting irrelevant position information that starts out weak at the beginning of the trial and strengthens over time. This idea was subsequently developed into the activation-suppression hypothesis (Burle et al., 2002; Ridderinkhof 2002a, 2002b; Ridderinkhof et al., 2004; Wylie, van den Wildenberg, et al., 2009). According to this hypothesis, in a conflict task, the automatic activation due to the irrelevant stimulus feature is initially strong but, over time, is selectively inhibited. The buildup of this inhibition is gradual and slow, so the effect of the automatic activation diminishes over the course of the trial. According to Ridderinkhof (2002a, Fig. 5), the concept of gradual selective response inhibition translates into a concave DP like the one shown in the third panel of Fig. 1, which initially (i.e., for short RTs) increases and then, due to the inhibition, levels off for medium RTs and decreases, possibly to zero or even negative values for long RTs. Many studies by Ridderinkhof and colleagues and by others have since confirmed this prediction under various conditions, especially

with the standard Simon task (for recent reviews, see Pratte et al., 2010; Proctor et al., 2011; van den Wildenberg et al., 2010). Therefore, decreasing DPs have come to be seen as a critical test of the activation-suppression hypothesis, and they are often considered to be a key signature of a selective response inhibition characterized by a slow buildup. Similarly, Burle et al. (2002; e.g., p. 333) endorsed the idea that nDPs indicate increasing suppression, which they interpreted as a sign of increasing executive control. Before accepting nDPs as strong evidence for the activation-suppression hypothesis, however, it seems worthwhile to consider what other mechanisms could also generate this observed data pattern.

Five classes of models predicting delta plots with negative-going slopes

Exhaustive models and their delta plots

We first consider certain types of exhaustive processing models. Such models include terms representing the maximum of two random variables, and in this section, we will show that these can produce nDPs of the form shown in Fig. 1. In RT research, such models can arise in systems requiring exhaustive processing of different items or channels that are being processed in parallel (e.g., Townsend, 1984; Townsend & Ashby, 1983). In general, if the start of a later process must await the completion of two or more parallel prior processes (e.g., because the information from all of the prior processes is needed as input to the later process), the starting time for the later process can be modeled as the maximum of the completion times of the prior processes.

One familiar example of an exhaustive processing model arises within the context of the psychological refractory period paradigm. In this paradigm, the stimuli, S_1 and S_2, for two separate tasks are presented sequentially, separated by a brief stimulus onset asynchrony (SOA), and the participant must make separate responses, R_1 and R_2, to the two stimuli. According to standard central bottleneck models for this paradigm (e.g., Pashler, 1994a), each task is accomplished by carrying out a sequence of three processing stages, A_i, B_i, and C_i, where the letter denotes the stage and the subscript denotes the task. These models explain the standard finding that RT_2 decreases as SOA increases by assuming that central processing of the second task (i.e., stage B_2) cannot begin until first-task central processing and second-task precentral processing have both completed, which happens at time $\max(A_1 + B_1 - SOA, A_2)$, defining the onset of S_2 as time zero. Thus, $RT_2 = \max(A_1 + B_1 - SOA, A_2) + B_2 + C_2$. Essentially this same exhaustive processing model can also be used to model precue utilization tasks (Schwarz & Ischebeck, 2001); other examples of maximum-based models arise within the context of stochastic PERT models (e.g., Fisher & Goldstein, 1983; Miller, 1993; Schweickert, 1982; Schweickert & Townsend, 1989) and within multichannel models in which processing waits for the completion of the slowest channel (e.g., Colonius & Ellermeier, 1997; Colonius & Vorberg, 1994).

Figure 2 illustrates an exhaustive processing model that might be applicable to Simon tasks, in order to show how these models can predict nDPs. The upper sequence of stages depicts performance on congruent trials, with the three stages corresponding to stimulus encoding, response activation, and response execution, respectively. The lower sequence of stages depicts performance on incongruent trials. It is assumed that the encoding stage detects the mismatch between the irrelevant position information and the relevant stimulus dimension (e.g., Proctor et al., 2011, pp. 248–250). Because of that mismatch, it is necessary not only to activate the correct response using the relevant dimension (stage B_c), but also to actively inhibit the activation of the incorrect response produced by the irrelevant dimension (stage B_i). Inhibition of the incorrect response might be necessary, for example, to resolve response conflict within a winner-take-all response system. Activation of the correct response and inhibition of the incorrect one can go on in parallel, and response execution (i.e., stage C in Fig. 2) cannot begin until both are completed. Thus, the RTs on congruent and incongruent trials are, respectively,

$$RT_c = A + B_c + C$$

and

$$RT_i = A + \max(B_c, B_i) + C.$$

Intuitively, it makes sense that RT_i should tend to be lengthened relative to RT_c, especially at the faster end of the distribution. When B_c is relatively fast, there is a larger

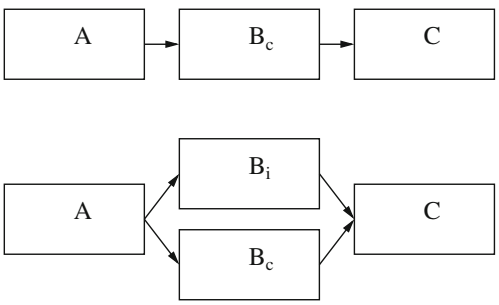

Fig. 2 *Top panel*: Serial processing model based on three stages: stimulus encoding (A), response activation (B_c), and response execution (C). *Bottom panel*: Exhaustive processing model: On incongruent trials, the mismatch between relevant and irrelevant stimulus dimension is detected at the encoding stage, A. Response execution C cannot begin until the correct response is activated (B_c) and the incorrect response is successfully inhibited (B_i); the latter two processes go on in parallel

cost of waiting for B_i to be done too. When B_c is relatively slow, however, there is a smaller cost of waiting for B_i; in fact, if $B_c > B_i$, there is no cost at all.

Figure 3 shows an example of RT distributions and DPs predicted by this model. RTs were assumed to come from a model in which A + C was modeled as an ex-Gaussian random variable, with μ = 220 ms, σ = 10 ms, and τ = 10 ms in both conditions; B_c was modeled as an inverse Gaussian variable with drift and variance parameter μ = 1, σ^2 = 100, and barrier a = 100 in both conditions. In the incongruent condition max(B_c,B_i) was modeled as the maximum of B_c and the random variable B_i, also an inverse Gaussian variable, with μ = 1.135, σ^2 = 31.3 and a = 100, a total of seven parameters. In line with the intuitive arguments given above, Fig. 3 shows that the exhaustive model is able to predict the RT histograms and concave nDPs of the type shown in panel 3 of Fig. 1, which for the parameters chosen rises from about 30 to 38 ms and then drops to a value close to zero at higher RT quantiles. This predicted pattern seems to be in reasonably good agreement with the empirical DP shown in the third panel of Fig. 1.

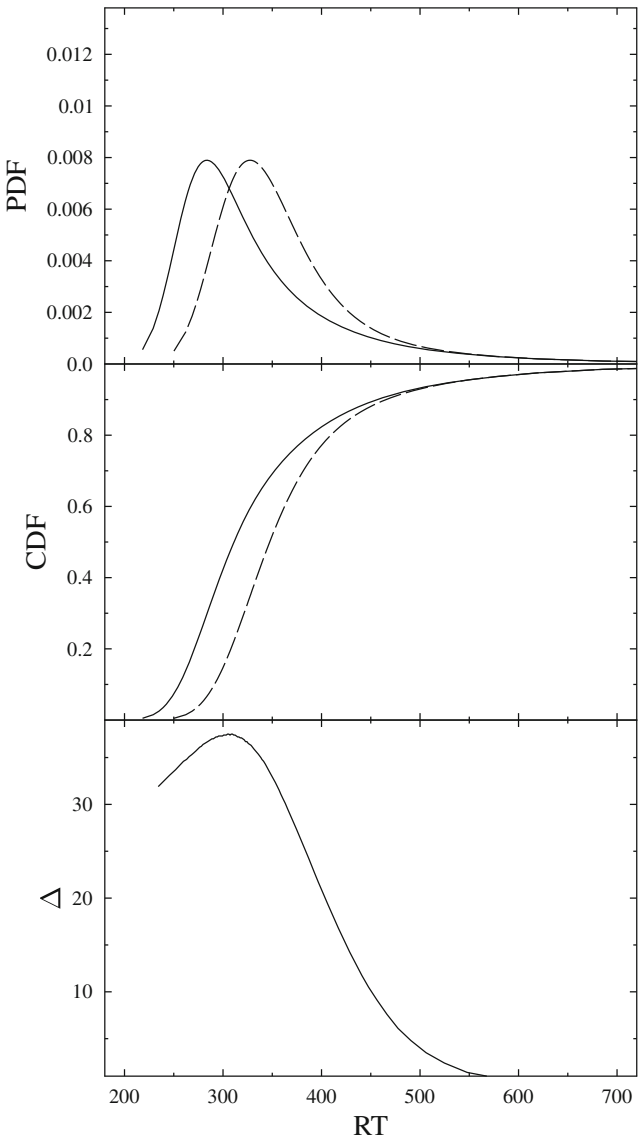

Fig. 3 *Top panel*: Reaction time (RT) frequency distributions for congruent (solid line) and incongruent (broken line) conditions, as predicted by the exhaustive model shown in the bottom part of Fig. 2. The distributions have mean RTs of 450 and 472 ms and standard deviations of 64 and 60 ms, respectively. *Middle panel*: Cumulative distribution functions (CDFs) for RT in the congruent (solid line) and incongruent (broken line) conditions. *Bottom panel*: Delta plot of quantile differences (ordinate) against quantile averages (abscissa)

Stage models and their delta plots

Stage models have a long history in RT modeling (e.g., Donders, 1868/1969; Sternberg, 1969). In these models, processing is carried out by a sequence of stages, so the total RT is simply the sum of the times needed by the individual stages, as is shown for three serial stages in the top panel of Fig. 2. Such models are easily capable of producing positive-going DPs (Pratte et al., 2010). More relevant for the present concerns, however, is that they can also produce nDPs, as is illustrated in this section.

Consider the first two stages, A and B, of a simple serial RT model (as in the top panel of Fig. 2). Let t_A and t_B be random variables representing the durations of these two stages, ignoring stage C for the moment. The total time needed for the first two stages is $t_A + t_B$, and its mean RT is of course $\mu_A + \mu_B$, where μ_A and μ_B are the means of the individual random variables. More critically, the variance of $t_A + t_B$ is

$$\mathrm{Var}(t_A + t_B) = \sigma_A^2 + \sigma_B^2 + 2\,\sigma_A\,\sigma_B\,\rho_{AB}, \qquad (1)$$

where σ_A and σ_B are the standard deviations of the two variables and ρ_{AB} is their correlation.

For our purposes, the key fact about this stage model is that the variance of the total RT increases with the correlation between the two stage durations, ρ_{AB}. This implies that, in any comparison between two conditions, the condition with the smaller mean RT could easily have the larger variance if it happens to have a larger correlation in the finishing times of its two stages. This is evidently the case when the correlation is positive for the congruent and negative for the incongruent conditions, but the requirement can also be met when both correlations are positive or when both are negative. The possibility of correlations among the times needed for different stages, while sometimes ignored in RT modeling, has been examined seriously by many (e.g., Colonius, 1986; Dzhafarov, 1992; Dzhafarov & Cortese, 1996; Pieters, 1983; Schwarz, 1994; Taylor, 1976; Van der

Heijden, Schreuder, Maris, & Neerincx, 1984). Thus, it is not unreasonable to consider models with correlated stage times as candidates for explaining nDPs. After showing numerically how such models can produce the nDP pattern, we discuss some possible explanations both of correlations among stage times and of changes in these correlations across conditions.

To provide a concrete illustration of how nDPs can arise from stage models, we adopted the ex-Gaussian model (cf. Luce, 1986, chap. 3.2.1) for the distributions in both of the two conditions being compared (e.g., congruent and incongruent). For this illustration, we used an exponential component for stage C shown in the top panel of Fig. 2, and this component had a mean of 50 ms in both conditions and was independent of the normal component in both. The ex-Gaussian's normal component in each of the conditions was conceived of as itself representing the sum of two normally distributed stage times, A and B. In the faster condition, the two stage times had means of $\mu_A = \mu_B =$ 200 ms and standard deviations of $\sigma_A = \sigma_B = 20$ ms. The correlation of these two stage times was $\rho_{AB} = 0.5$, so the overall standard deviation of the normal component was $\sqrt{20^2 + 20^2 + 2 \times 20 \times 20 \times 0.5} = 35$ ms. In the slower condition, the two stage times had means of $\mu_A = 200$ ms and $\mu_B = 250$ ms and standard deviations of $\sigma_A = 20$ ms and $\sigma_B = 25$ ms. The correlation of these two stage times was $\rho_{AB} = -0.5$, so the overall standard deviation of the normal component was $\sqrt{20^2 + 25^2 - 2 \times 20 \times 25 \times 0.5} = 23$ ms. Given that the normal standard deviations were all chosen in fixed proportion to their mean (a factor of 0.1), this specific version of the model has five parameters. Thus, including the exponential components, the overall mean and standard deviation of RT are 450 and 61 ms in the congruent condition, as opposed to 500 and 55 ms in the incongruent one, so the faster condition has the larger variance.

The distributions and DP predicted from ex-Gaussian distributions with the normal components involving the positively and negatively correlated stage durations are shown in Fig. 4. The nDP data pattern shown in Fig. 1 is nicely reproduced, with a larger separation of the CDFs at the low end of the RT distribution, a larger variance in the faster condition, and an nDP. This is, of course, only one simple illustration of how decreased correlations among stage times can reduce overall RT variance and produce nDPs; more elaborate illustrations (e.g., involving more stages) can easily be constructed analogously.

It is possible to speculate about various possible sources of correlation between different stage times and about several factors that might cause the correlations to change across conditions, thereby potentially producing both a larger RT variance in the faster condition and an nDP. As one rather generic example, tasks or processes competing for the same limited-capacity resources tend to be negatively correlated (e.g., Navon & Gopher, 1979, 1980; Pashler, 1994b). Thus, if the stages used in a faster task suffered less from resource competition than did the stages used in a slower task—perhaps because they required fewer resources in the first place—then the faster-task stages would naturally have a larger (or less negative) correlation, resulting in more RT variance.

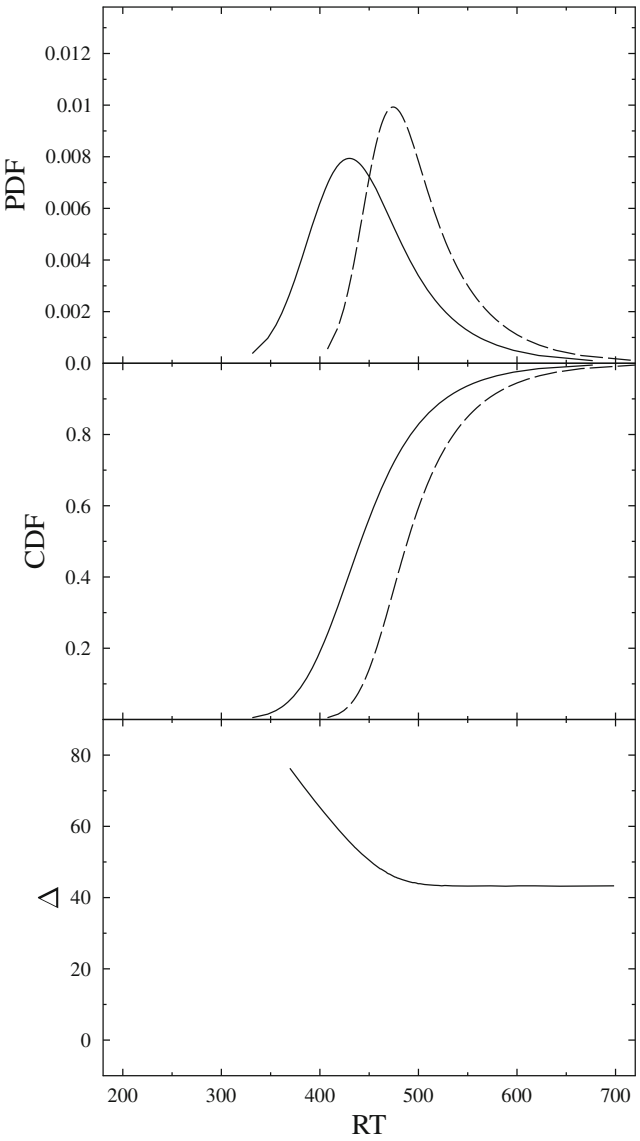

Fig. 4 *Top panel*: Reaction time (RT) frequency distributions for congruent (solid line) and incongruent (broken line) conditions, as predicted by the serial model shown in the top part of Fig. 2. The distributions have mean RTs of 450 and 500 ms and standard deviations of 61 and 55 ms, respectively. *Middle panel*: Cumulatiave distribution functions (CDFs) for RT in the congruent (solid line) and incongruent (broken line) conditions. *Bottom panel*: Delta plot of quantile differences (ordinate) against quantile averages (abscissa)

Several more specific sources of correlation between different stage times have been suggested. For example, in their analysis of temporal preparation effects, Schwarz (1994, p. 514) and Los and Schut (2008) suggested that

nonspecific motor preparation can take place in parallel with perceptual analysis (for further discussion, see Leonhard, Bratzke, Schröter, & Ulrich, 2012). According to this model, trials with slower perceptual analysis would provide more opportunity for motor preparation—ultimately producing faster motor execution—so the durations of the perceptual and motor stages would tend to be negatively correlated across trials. If the inverse relationship were stronger for incongruent trials than for congruent ones, the between-stage correlation could be more negative for incongruent trials than for congruent ones, leading to nDPs. This might happen, for example, if perceptual times were more variable on incongruent trials or if the degree of nonspecific motor preparation was less variable on these trials. The reason is that in this model, assuming perfect parallel motor preparation, the correlation (r) of the perceptual (P) and residual motor ($M - P$) stage durations would be $r(P, M - P) = -1/\sqrt{1 + \mathrm{Var(M)}/\mathrm{Var(P)}}$, which becomes more negative as Var(P) increases or as Var(M) decreases.

Within the Simon task, pre-trial preparatory differences would be another natural source of correlations that could differ between the congruent and incongruent conditions. Suppose that when the trial starts, the participant has already selectively preactivated a code for either "left" or "right," perhaps based on the stimulus location or the response made on the previous trial (e.g., Hommel, Proctor, & Vu, 2004; Kornblum, 1969; Notebaert & Soetens, 2003; Notebaert, Soetens, & Melis, 2001). Suppose further that this preactivation influences both the time needed for an early stimulus localization stage and the time needed for a later response selection stage, with both of these stages operating faster when the appropriate output code has been selectively preactivated. In the congruent condition, where both stages need to output the same code, both of these stages would operate relatively quickly when the correct codes were preactivated and relatively slowly when the incorrect codes were preactivated, which implies a positive correlation across many congruent trials between the stage times. In the incongruent condition, where the two stages need to output opposite codes, the stage outputting the preactivated code would operate relatively quickly, and the other stage would operate relatively slowly, which implies a negative correlation between the stage times.

It is true that many models of conflict tasks do not meet the rigid assumptions of classical serial stage organization. For example, the activation-suppression model (Ridderinkhof, 2002a) assumes parallel response activation from relevant and irrelevant stimulus features. However, we note that additive models as defined above (i.e., $\mathrm{RT} = t_A + t_B$) are not limited to strictly serial latency mechanisms in which stage $n + 1$ starts only if and when stage n is completely finished (cf. Atkinson, Holmgren, & Juola, 1969; Townsend & Ashby, 1983). In the simplest case, consider a parallel-exhaustive model in which two processes start simultaneously and both need to be finished to produce the overt response, $\mathrm{RT} = \max(t_A,t_B)$. Even for this strictly parallel model, the total RT still consists of two additive components: the time from the start until the faster process finishes, $T_1 = \min(t_A,t_B)$, plus the remaining time from there until the slower process finishes, $T_2 = \max(t_A,t_B) - \min(t_A,t_B)$. That is, when imagined as an evolving process that unfolds in time, then in any actual realization of $T = \max(A,B)$, the first event is that the faster subprocess finishes. This happens after $T_1 = \min(t_A,t_B)$ time units, which may be interpreted as the duration of a serial stage 1. Furthermore, the time from there until the slower process also finishes, $T_2 = \max(t_A,t_B) - \min(t_A,t_B)$, may be seen as the duration of a serial stage 2. Overall, then, $\mathrm{RT} = T_1 + T_2$. The term $\min(t_A,t_B)$ appears with different signs in the two components T_1 and T_2; therefore, these components tend to be negatively correlated, even when the process durations t_A and t_B themselves are independent. As a consequence, if the congruency effect prolongs one of the processes, the stage correlation will become more negative under many constellations, leading to the characteristic nDP pattern of an increased mean RT yet a smaller RT variance for incongruent trials.

Mixture models and their delta plots

We next consider mixture models (for general background and review, see Luce, 1986, chap. 7; Yantis, Meyer, & Smith, 1991). The basic assumption of mixture models is that on each trial, the participant is in one particular discrete internal state that characterizes the processing mode on that trial. Across trials, this state is probabilistically selected from a small set of possible states. Here, we focus on binary mixture models, which assume that on each trial, only two such internal states, or processing modes, are available to the participant.

Mixture models provide a fairly general formal structure (see Luce, 1986, chap. 7), and they occur in many substantive contexts under somewhat different labels. For example, the dual-route model of spatial compatibility effects (e.g., Van Duren & Sanders, 1988, Fig. 3b) holds that the information conveyed by the stimulus is processed via two parallel routes: one controlled and relatively slow, the other fast and automatic. On each individual trial, the overt response is generated via either one or the other of the two processing paths, with complementary probabilities of p and $1 - p$. Formally identical models have also been proposed in many other contexts—for example, to account for processing differences following compatible versus incompatible trials (Wylie et al., 2010) and for the effects of set on speed

versus accuracy (Wylie, van den Wildenberg, et al., 2009; Yellott, 1971).

The mean RT according to the binary mixture model is (e.g., Yantis et al., 1991, Eq. 3)

$$\mathrm{E}[\mathrm{RT}_i] = p_i\mu_\mathrm{a} + (1 - p_i)\mu_\mathrm{b},$$

and the RT variance is (e.g., Yantis et al., 1991, Eq. 5)

$$\mathrm{Var}[\mathrm{RT}_i] = p_i\sigma_\mathrm{a}^2 + (1 - p_i)\sigma_\mathrm{b}^2 + p_i(1 - p_i)(\mu_\mathrm{b} - \mu_\mathrm{a})^2,$$

where μ_a, μ_b, and σ_a^2 σ_b^2 are the means and variances, respectively, of the two base states.

The inset of the upper panel of Fig. 5 shows two base densities that characterize the distributions of RTs as they would be observed under each "pure" state. We have chosen for illustrative purposes two inverse Gaussian distributions (both with barriers $a = 100$ and variance parameter $\sigma^2 = 1$), differing only in their drift rates ($\mu = 0.35$ vs. $\mu = 0.25$), but could just as well have chosen, for example, a pair of lognormal or even normal distributions to a similar effect. The mean RTs are $\mu_\mathrm{a} = 286$ ms and $\mu_\mathrm{b} = 400$ ms, and the associated standard deviations are $\sigma_\mathrm{a} = 48$ ms and $\sigma_\mathrm{b} = 80$ ms. As a functional interpretation, for example, the one state (a) could represent "fast and automatic" processing, and the other state (b) "slow and controlled" processing; for example, there might actually be no cost on incompatible trials, only benefits for compatible ones. Broadly, these distributions exhibit the qualitative features (means, spread, skewness) often associated with RTs. Note, in particular, that for these base distributions, the means and variances are ordered consistently: The distribution with the smaller mean also has the smaller variance. Correspondingly, the DP formed from comparing these two types of processing, not shown here, would be strictly increasing. The main part of the upper panel shows two mixture distributions, corresponding to two experimental conditions $i = 1,2$. These were generated from the two inset base distributions, with mixture proportions of $p_1 = .4$ in the faster condition and $p_2 = .05$ in the slower condition, giving a total of six model parameters. In the faster condition, the overall mean RT is 354 ms, and its standard deviation is 89 ms; in the slower condition, mean RT is 394 ms, and its standard deviation is 83 ms. Thus, even though the means and variances of the two base distributions shown in the upper panel are consistently ordered, the basic mixture mechanism assumed by this class of models can produce an inconsistent moment ordering. The DP in the lower panel of Fig. 5 is generated from the corresponding CDFs in the middle panel. Evidently, after a short initial increase, the long tail of the DP is decreasing. That is, the binary mixture model can produce nDPs, even when the base distributions (inset within top panel) show a consistent ordering of means and variances and an increasing DP (not shown). A very similar DP emerges if the controlled route can be used only on incompatible trials (i.e., $p_2 = 0$), which would mean functionally that there is actually no cost for incompatible trials but only benefits for compatible ones.

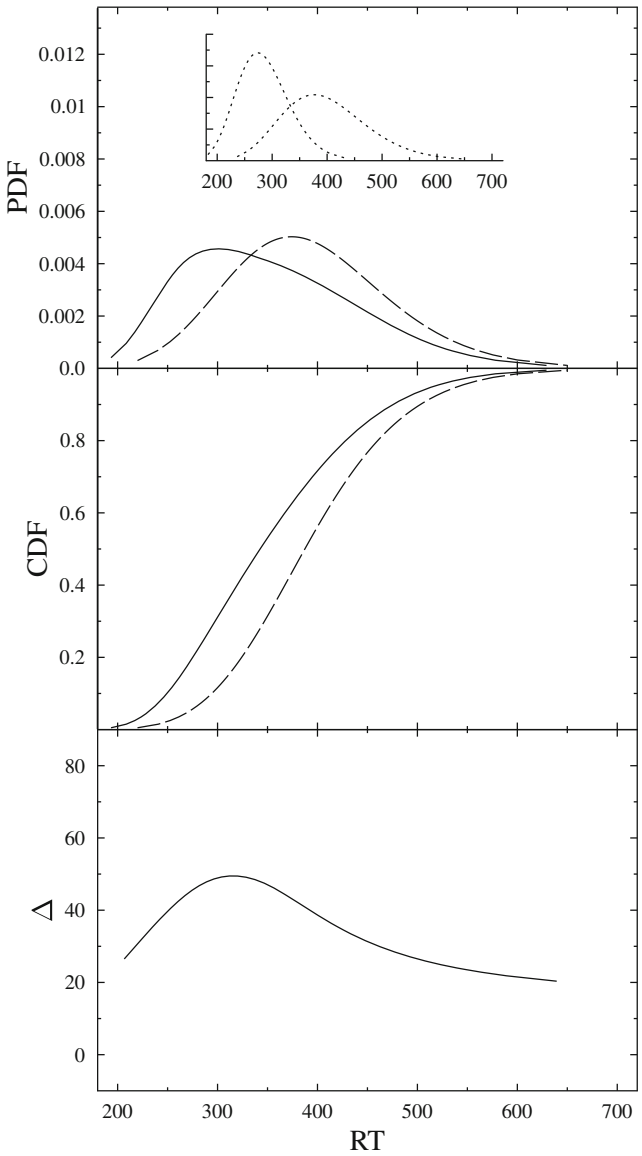

Fig. 5 *Top panel*: The inset shows base reaction time (RT) frequency distributions for a faster and a slower processing state, with means of 286 and 400 ms and standard deviations of 48 and 80 ms, respectively. The RT frequency distribution for the congruent condition (solid line) is a mixture distribution, with weights of $p_1 = .40$ for the faster base distribution, and $1 - p_1 = .60$ for the slower base distribution, giving a mean RT of 354 ms and an RT standard deviation of 89 ms. The RT frequency distribution for the incongruent condition (broken line) is a mixture distribution, with weights of $p_2 = .05$ for the faster base distribution, and $1 - p_2 = .95$ for the slower base distribution, giving a mean RT of 394 ms and an RT standard deviation of 83 ms. *Middle panel*: Cumulatiave distribution functions (CDFs) for RT in the congruent (solid line) and incongruent (broken line) conditions. *Bottom panel*: Delta plot of quantile differences (ordinate) against quantile averages (abscissa)

As was explained above, nDPs reflect an inconsistent ordering of means and variances—the faster RT distribution having the larger spread. Thus, an obvious question to ask is whether, or under which conditions, the binary mixture model predicts this inconsistent moment ordering. Consider the RTs (RT_i) in conditions $i = 1,2$ as functions of the mixture parameters p_i. To fix ideas, assume (as shown in Fig. 5) that $\mu_a < \mu_b$ and that $\sigma_a < \sigma_b$; that is, RTs produced from one state (a) are shorter, on average, and also less variable than RTs from the other state (b). If RT_1 refers to the condition with the shorter mean RT, it can be shown (see the Appendix) that the pattern of inconsistent moment ordering

$$E[RT_1] < E[RT_2] \text{ and } Var[RT_1] > Var[RT_2]$$

arises if (and only if) the following condition is satisfied:

$$p_1 + p_2 < 1 - \frac{\sigma_b^2 - \sigma_a^2}{(\mu_b - \mu_a)^2}.$$

This result indicates some boundary conditions. For example, the faster state may not be more frequent overall than the slower one, because the condition cannot be satisfied if $p_1 + p_2 > 1$. Also, the mean RTs associated with the two states need to differ sufficiently, relative to the difference in variances, because the condition cannot be met if $(\mu_b - \mu_a)^2 < \sigma_b^2 - \sigma_a^2$. For the underlying cognition, that implies that one state must be clearly slower than the other and be used more often. Both conditions together tend to produce the pattern shown in Fig. 5, in which, at the lower end, the CDF of the faster condition is well separated from that of the slower condition, whereas the CDFs converge at the upper end. In sum, then, binary mixture models provide yet another plausible class of processing models capable of accounting for nDPs.

Cascade models and their delta plots

McClelland (1979; for elaborations, see Miller, van der Ham, & Sanders, 1995; Schwarz, 2003) proposed a cascade model that retains the basic notion of discrete, functionally distinct processing stages of serial RT models but differs from them in two important ways. First, rather than being all-or-none, the output activation of each stage varies gradually as a continuous function of time. Second, all stages are concurrently active: Their activation is continuously passed on from stage to stage such that the increase of activation at a given stage is proportional to the difference between its current activation and the activation of the immediately preceding stage. Thus, the cascade model pictures the flow of information as a continuous wave of activation traveling along successive discrete processing stages, a notion that has proven especially influential in the area of cognitive psychophysiology (e.g., Forster & Corballis, 2000; Scheffers & Coles, 2000). Also, it is one of relatively few models assuming partially overlapping processing stages that are sufficiently specified to allow derivation of explicit quantitative predictions (Schweickert & Mounts, 1998; Ulrich, Mattes, & Miller, 1999).

Closer analysis of this model reveals that it is easily capable of generating nDPs. Because McClelland's (1979) original cascade model had no stochastic mechanism for generating RT distributions—but only mean RTs—we adopted the general stochastic cascade model framework described by Schwarz (2003). Within this framework, neural spikes are transferred from stage to stage, with possibly different stage-specific rates, and the continuous activation level at each stage corresponds to the hazard (or intensity) function driving the neural spike counting process at that stage. A response is given when a criterion number (k) of spikes has been collected at the final output stage.

The upper panel in Fig. 6 shows RT frequency distributions based on a cascade model with two stages, each of which is characterized by a separate spike transmission rate (cf. Schwarz, 2003, Eq. 4). A response is given as soon as $k = 20$ spikes are collected at the second stage. The figure is based on the assumption that neural spikes at the first processing stage arrive at a higher rate (α_1) on incongruent than on congruent trials but are then transmitted faster (rate α_2) on to the second stage on congruent than on incongruent trials. Such a pattern might arise, for example, if the first stage represents automatic processing and the second a more controlled, or selective, form of processing. The automatic process generates spikes favoring both response codes on incongruent trials (because there is some activation associated with each response), but it generates only spikes favoring one response code on congruent trials (because both types of activation favor the same response). If spikes at the first stage are generated indiscriminately by the pooled activation of either response code, the rate α_1 would be higher for the incongruent condition. On the other hand, controlled, selective processing might depend on the differential evidence favoring one of the responses over the other, in which case the transmission of spikes to the second stage would proceed at a higher rate (α_2) for the congruent condition. The associated CDFs in the middle panel of Fig. 6 show that for short RTs, the congruency effect is large and that it declines as RT increases, thus generating (using five free parameters) the nDP shown in the bottom panel. Thus, the basic parallel architecture and the considerable flexibility of cascade models enable them to generate nDPs, even with a relatively straightforward implementation of conflict tasks.

Parallel channels models and their delta plots

As a final class, we consider parallel channels models, which, like cascade accounts, are also based on the concept

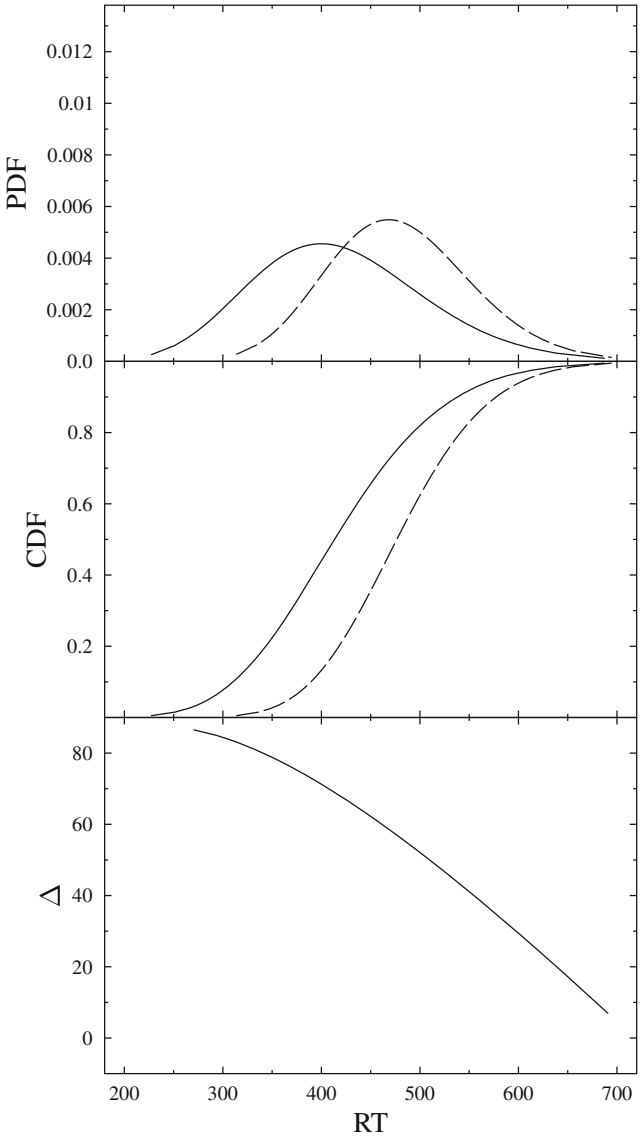

Fig. 6 *Top panel*: Reaction time (RT) frequency distributions for congruent (solid line) and incongruent (broken line) conditions, as predicted by a cascade model with two stages. In the congruent condition, the spike arrival rate at stage 1 and the spike transmission rate into stage 2 are α_1 = 1/20 and α_2 = 1/20; in the incongruent condition, they are α_1 = 1/15 and α_2 = 1/200; in both conditions, a response is given as soon as k = 20 spikes are collected at stage 2. The distributions have means of 420 and 481 ms and standard deviations of 89 and 74 ms, respectively. *Middle panel*: Cumulative distribution functions (CDFs) for RT in the congruent (solid line) and incongruent (broken line) conditions. *Bottom panel*: Delta plot of quantile differences (ordinate) against quantile averages (abscissa)

of massively parallel processing. These models assume that a criterion number of parallel processing channels must finish—or "information grains" be activated—before a motor command is initiated. More specifically, they assume that an overt response is produced as soon as the first k parallel channels have finished, from a total of n activated channels (cf. Meijers & Eijkman, 1974; Meijers, Teulings, & Eijkman, 1976; Miller & Ulrich, 2003). More formally, then, RT depends on the latency of the kth fastest channel, out of a total of n activated channels. A basic assumption within these models is that the total number (n) of activated channels varies systematically across conditions, in contrast to the constant criterion number (k) of channels that need to finish before a motor response is initiated. From these assumptions, it is intuitively clear that mean RT decreases as more channels are activated. For example, the 15th fastest channel out of a total of n = 20 activated channels will, on average, clearly be faster than the 15th fastest channel out of a total of just n = 17 channels. The critical question in the present context, however, is the following: Does this decrease in mean RT go together with a corresponding decrease in RT variance?

Figure 7 shows that the answer is "no, not necessarily." The figure is based on the assumption that, in both conditions, a response is generated as soon as k = 15 parallel channels have finished. In the incongruent condition, a total of n = 17 channels are activated; in the congruent condition, the total number of channels is n = 20. For simplicity, assume that all channels act independently and that their latency is uniformly distributed between 100 and 600 ms, for an average of 350 ms. Under these assumptions, mean RT in the incongruent condition is 517 ms, and its standard deviation is 43 ms. With n = 20 activated channels, mean RT in the congruent condition drops, as expected, to 457 ms, but the standard deviation increases by 12 % to 48 ms. Figure 7 indicates that this pattern (generated using five free parameters) translates into an nDP that is initially large and falls off for larger RT quantiles. We conclude that the class of parallel channels models is also capable of producing nDPs. We note, however, that with various forms of the basic channel latency distribution, an inconsistent congruency effect on mean and variance is more difficult, or even impossible, to obtain within this class of models, especially for channel latencies with a sharply peaked unimodal density. That is, these models do not predict nDPs in general; however, the crucial point is that under at least some conditions, they do. The uniform distribution is not a necessary condition, though, to get an nDP in this model; for example, beta or bimodal distributions can produce the effect as well. It should also be noted that the model does not assume that the overt (observed) latency is uniform, which would clearly be an implausible assumption. In contrast, uniform latencies of covert (latent) component processes actually have enjoyed some prominence; for example, Ratcliff and Tuerlinckx (2002, Fig. 1 and p. 441) assumed the nondecisional latency component T_{er} of their diffusion model to be uniformly distributed; the same assumption has been made routinely in fitting this popular model to data (e.g., Ratcliff, 2002).

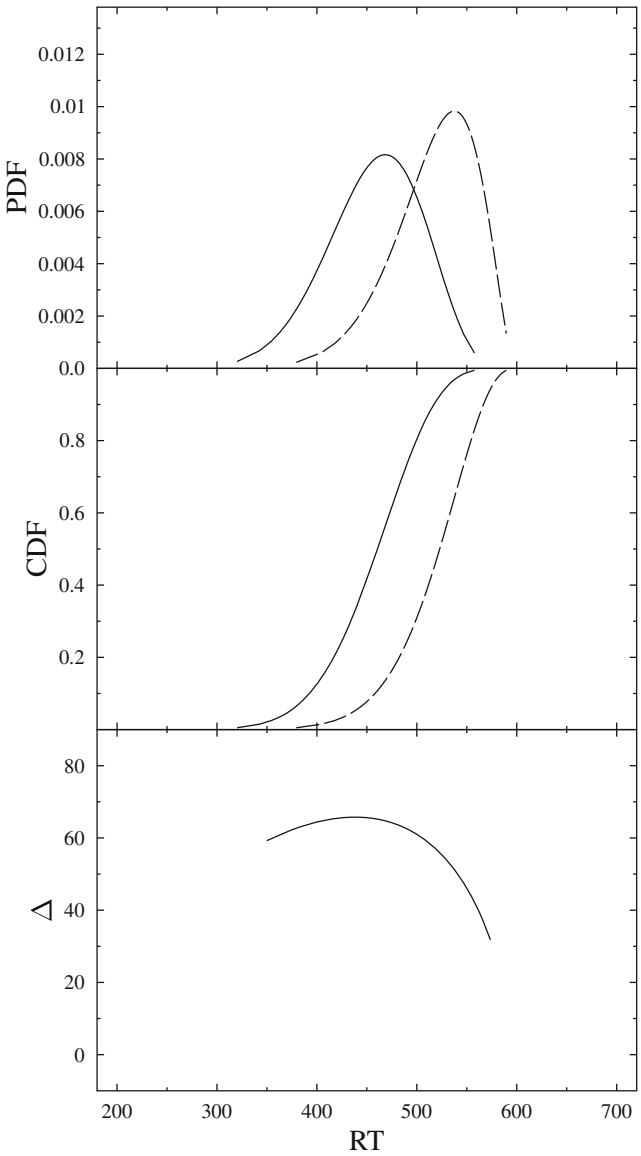

Fig. 7 *Top panel*: Reaction time (RT) frequency distributions predicted by a model in which 15 parallel channels have to be finished before an overt response is executed. In the congruent condition (solid line), 20 channels are activated; in the incongruent condition (broken line), 17 channels are activated. The finishing latency for each channel is uniform [100, 600]. The distributions have means of 457 and 517 ms and standard deviations of 48 and 43 ms, respectively. *Middle panel*: Cumulative distribution functions (CDFs) for RT in the congruent (solid line) and incongruent (broken line) conditions. *Bottom panel*: Delta plot of quantile differences (ordinate) against quantile averages (abscissa)

General discussion

Deriving DPs from RT data is one specific form of distributional analysis that, in going beyond mean RT, can potentially provide useful clues about how to formulate, test, and revise theoretical accounts of the latency mechanisms underlying those RT data. For example, the distinctive nDP signature often found with standard Simon tasks is an exceptional feature of RTs that can help to narrow down the search space for plausible candidate models. Analyzing DPs may also help to better appreciate that formally related tasks sometimes produce superficially similar results on the level of mean RT yet differ reliably in important finer dynamical aspects. For example, the Simon, Stroop, and Eriksen paradigms are all formally "conflict tasks", characterized by having both response-relevant and response-irrelevant stimulus features that vary across trials. Although all three paradigms typically produce a congruency main effect on mean RT, often of roughly comparable size, they may nevertheless differ in the basic form of DPs they generate (for a systematic review, see Pratte et al., 2010).

A necessary prerequisite to reap these potential benefits of DP analysis, however, would seem to be a better understanding of quantitative latency mechanisms that are, or are not, capable of producing nDPs. At present, there seems no clear consensus about exactly which dynamic, process-related implications nDPs actually have. Specifically, development of detailed quantitative models could help to predict the precise form of a DP and its variation across different experimental conditions. These predictions could then be compared with empirical DPs, much as, say, theoretical ROC curves from competing recognition memory models are routinely compared with data (e.g., Heathcote, 2003; Wixted, 2007). For example, an important impetus behind many DP analyses of the Simon effect to date is the view that the activation-suppression model (Ridderinkhof, 2002a, 2002b) with its gradual buildup of inhibition predicts an nDP. However, at present, this model has not been specified to a degree that would permit one to deduce the precise quantitative form of the DP that it predicts. Thus, it appears that the full potential of critical DP analyses has not yet been exploited.

The results presented in this article are a step toward this aim. They show that various well-established RT models are capable of producing nDPs—at least, under the right additional assumptions, which we specify in some detail. For example, as was noted previously by Pratte et al. (2010), serial stage models do not normally predict nDPs. We have shown that they may do so with correlated stage durations, however —as long as the between-stage correlation is larger in the faster condition. As we have argued in our discussion of the correlated stages model, this is a plausible scenario in the case of congruent versus incongruent stimuli in the Simon task. Similarly, two-state mixture models can predict nDPs if the mean RTs associated with the two states differ sufficiently and if the slower (e.g., controlled) processing state is the more frequently used "standard route". Also, we show that models in which RT depends critically on the slower of two parallel processes—such as processing the relevant stimulus feature and inhibiting activation from the irrelevant stimulus feature—are clearly consistent with

nDPs. Other models that, under at least some conditions, are able to generate nDPs include cascade models and massively parallel models in which the number of parallel channels varies across conditions.

In the introduction, we conjectured that nDPs would seem to provide rather specific clues about the nature of the effect under investigation. This view can be made more precise on the basis of our results. First, we note that from the sheer number and diversity of the different processing architectures that are all consistent with nDPs—ranging from serial to mildly and massively parallel on to cascade and mixture models—it appears unrealistic to expect that any unique model class can be inferred on the basis of an nDP. Second, a more realistic view is that certain classes of models that are incapable of generating nDPs can be ruled out as adequate accounts. For example, no candidate model of the standard Simon effect can be regarded as plausible if it is unable to generate an nDP. Of course, the ability to generate an nDP is not sufficient for a model to provide an adequate account of the Simon effect. For example, it must also make appropriate predictions about the effects of various experimental manipulations and have a certain degree of neurological plausibility.

Diffusion models

One flexible and widely studied class of models that apparently has to add implausible ad hoc assumptions to account for the nDPs often seen with Simon effects is the class of diffusion models (e.g., Luce, 1986, chap. 9; Ratcliff & Smith, 2004). These models assume that stimulus-related information is accumulated gradually over time; this accumulation is modeled as a diffusion process. A standard assumption is that within any single trial, the mean drift rate μ and the variance σ^2 of this accumulation process are constant. In this case, the time to reach a single evidence threshold a, typically postulated to be the point at which a decision is reached, has an inverse Gaussian distribution, with mean a/μ and variance $a\,\sigma^2/\mu^3$. In a simple implementation of this model, relative to congruent stimuli the presentation of incongruent stimuli might be assumed either (1) to increase the evidence threshold, a, or (2) to lower the drift rate, μ. In either case, the functions relating the mean and variance of the decision time to the underlying model parameters imply that the mean and variance would change concurrently as either parameter was varied. Therefore, this simple model does not generate the pattern of means and variances that typically underlies nDPs (see also Pratte et al., 2010, pp. 2022–2023), and indeed it generates increasing DPs.

A slightly more sophisticated diffusion model for the Simon task could assume that the relevant (e.g., color) and irrelevant (e.g., location) stimulus features induce separate, possibly time-dependent, drift components [say, $\mu_r(t) \geq 0$, and $\mu_i(t) \geq 0$]. For congruent stimuli, these two components activate the same response and, thus, add, whereas for incongruent stimuli, they activate opposite responses and, thus, subtract (see Schwarz & Ischebeck, 2003). As one interpretation of the activation-suppression hypothesis, assume that the drift component induced by the relevant stimulus feature is constant [$\mu_r(t) = \mu_r$], whereas that induced by the irrelevant stimulus feature becomes gradually inhibited over time; for example, $\mu_i(t) = \mu_i \cdot e^{-\lambda t}$. Thus, for congruent stimuli, the net drift rate would start as $\mu_r + \mu_i$ and decline to μ_r. For incongruent stimuli, it would start as $\mu_r - \mu_i$ and increase to μ_r. That is, for congruent stimuli, the asymptotic drift rate μ_r is approached from above; for incongruent stimuli, μ_r is approached from below. This means that the net drift rate at any point in time would be smaller for incongruent stimuli than for congruent stimuli, and this condition leads again to an increase of both mean and variance for the decision times. Overall, then, we conclude that standard implementations of diffusion models, including those with time-dependent diffusion rates, do not predict nDPs.

Of course, this does not rule out the possibility that nDPs could be produced by more complex diffusion model variants, possibly with congruent and incongruent conditions differing in threshold and drift at a time, with suitably chosen time- and state-dependent drift and variance terms or with across-trial parameter randomization schemes (Ratcliff, 1978). For example, even the standard diffusion model with a single barrier, leading to the inverse Gaussian distribution, can produce nDPs. To get this result, assume that relative to the congruent condition, for the incongruent condition the variance parameter σ^2 is smaller (but μ, a remain unchanged), or, formally equivalently, that both μ and a are larger (but σ remains unchanged). It is not apparent to us, though, why that should be so. The standard notions (cf. Luce, 1986, chap. 9) about diffusion models are that the incongruent condition has smaller drift and/or higher barriers and that, in a randomized trial design, the choice of barriers does not depend on the actual stimulus presented.

Similarly, the standard diffusion model would produce nDPs within the framework of the mixture model, which can be seen as a very simple form of parameter randomization. That is, if RTs under both states underlying the mixture model were inverse Gaussians differing only in their drift rates, the state with the higher drift rate would have the shorter mean RT and also the smaller RT variance. Indeed, this is precisely the model from which Fig. 5 was generated: The inset in its top panel shows two inverse Gaussian RT distributions differing only in their drift parameters. The bottom panel shows that if, under the congruent and incongruent conditions, these two RT distributions are mixed in the proportions used to construct Fig. 5, the diffusion-based mixture model generates an nDP. In our view, this observation does not strongly support diffusion-based accounts of

nDPs, though. Instead, it simply indicates that the randomization scheme underlying the mixture model provides a fairly flexible mechanism that works similarly for nearly any "well-behaved" unimodal distribution, which includes, of course, the inverse Gaussian.

Time course of effects

In their original study reporting nDPs in the context of the Simon effect, DeJong et al. (1994) claimed that their observed nDPs "almost certainly provide a reliable estimate of the actual time course of these effects" (p. 733). In view of the diversity of models that potentially predict nDPs, our results further strengthen previous conclusions (e.g., Zhang & Kornblum, 1997) that even if the separation of two RT CDFs is strictly limited to the few earliest quantiles, there is still no assurance that the effect generating this pattern necessarily operates at an early stage of processing within any trial. To illustrate this important point further, using the Simon task as an example, consider a standard serial model (cf. Sternberg, 1969, and Fig. 2) with two stages, RT = $t_A + t_B$, such as the popular ex-Gaussian model used in Fig. 4. Let the notation RT = $t_A + t_B$ indicate that, on each trial, the stage whose duration is t_A precedes the stage whose duration is t_B.[1] Suppose that in a Simon task, the congruency effect observed selectively influences the stage with latency t_A, prolonging it stochastically from t_A on congruent trials to t_A' on incongruent trials. Now, the distribution of $t_A + t_B$ is evidently the same as that of $t_B + t_A$ (i.e., when stage t_B precedes the affected stage t_A) and so, therefore, are two DPs, one involving a plot of $t_A' + t_B$ versus $t_A + t_B$, and one involving a plot of $t_B + t_A'$ versus $t_B + t_A$. Thus, if an nDP is observed, it might just as well have resulted from an effect on the *later* stage (i.e., from the underlying model RT = $t_B + t_A$). Indeed, in a serial model with, say, 100 stages, an nDP might show a congruency effect that is strictly limited to the 10 % shortest RTs, and yet the effect may, on each trial, exclusively arise at the very last of these 100 stages. For serial models, it is, in principle, impossible to distinguish between "early stage" and "late stage" chronometric interpretations of an observed effect on the basis of the associated DP.

The activation-suppression model

As was noted above, the activation-suppression model (Ridderinkhof, 2002a, 2002b) has been a strong impetus behind many extant DP analyses, even though the empirical evidence supporting this model extends beyond DP analyses (for a recent review, see van den Wildenberg et al., 2010). Could any of the nDP-consistent models described in the preceding sections fit in with the basic assumptions of that model?

Neither mixture models, correlated serial-stages models, cascade models, nor parallel channels models seem to correspond at all well with the assumptions underlying the activation-suppression model. According to mixture models, the response on each trial is made from within one of several distinct states, but the activation-suppression model postulates no such mutually exclusive states. The models with correlated stages are strictly serial, in contrast to the inherently parallel architecture of the activation-suppression model. Furthermore, the correlated-stage models include nothing that might be construed as suppression over time of the influence of the irrelevant attribute. Within cascade and parallel channels models, there is only accumulation of activation (e.g., a summation of neural spikes), but no explicit mechanism for inhibition or suppression. Thus, it seems more natural to view these models as alternatives to the activation-suppression model, however formalized, rather than as possible implementations of it.

On the other hand, maximum-based exhaustive processing models do correspond in some ways to the assumptions underlying the activation-suppression model, so these could possibly be viewed as potential formalizations of it. For example, in both exhaustive processing models and the activation-suppression model, nDPs emerge from the parallel operation of two processes. One might even view the two processes within the exhaustive processing model as corresponding to activation of responses by the relevant information and suppression of the irrelevant information, which would make for an even closer correspondence of this model to the activation-suppression model. On the other hand, the activation-suppression model posits a continuously decreasing influence of one channel on the other, whereas the exhaustive model could be said to involve a more all-or-none influence of the suppression process on the activation process. Specifically, suppression has an effect when it is the slower process, by increasing the maximum, but no effect at all when it is the faster process.

Statistical analyses of nDPs

An important benefit of an explicit formulation of models capable of producing nDPs is that they provide a more detailed characterization of DPs and, thereby, allow for a more efficient and powerful analysis of empirical DPs. For example, even for the standard Simon paradigm involving colored patches and lateralized manual responses from healthy participants, there is considerable variation in the results reported. Burle et al. (2002, Fig. 1), Kubo-Kawai and Kawai (2010, Fig. 1), Pratte et al. (2010, Fig. 2D), and Proctor et al. (2005, Figs. 2 and 4) found strictly decreasing DPs, whereas Davranche and McMorris (2009, Fig. 4), Ridderinkhof et al. (2004, Fig.

[1] For example, t_A might be the normal and t_B the exponential component in the ex-Gaussian model.

2), Vallesi et al. (2005, Table 1), and Wylie at al. (2010, Fig. 4) reported an increasing-then-decreasing pattern. Similarly, some studies (Burle et al., 2002, Fig. 1; Kubo-Kawai & Kawai, 2010, Fig. 1; Vallesi et al., 2005, Fig. 1; Wylie, Ridderinkhof, et al., 2009, Fig. 3B; Wylie et al., 2010, Fig. 4) reported nDPs crossing the abscissa at the highest quantiles (thus violating stochastic dominance), whereas many others (e.g., Burle et al., 2002, Fig. 3; Davranche & McMorris, 2009, Fig. 1; Davranche et al., 2009, Fig. 1; DeJong et al., 1994, Fig. 3; Pratte et al., 2010, Fig. 3B) did not.

How can these different DP patterns be reconciled? One informative approach is to try to relate them to subtle stimulus- or design-related differences among these studies (Proctor et al., 2011). In addition, to address the question of how reliable and significant these apparent discrepancies really are, it is also clearly relevant to consider the standard errors (*SE*s) of the plotting positions of a DP and the factors influencing them. To investigate these *SE*s, we simulated the complete data generation and data analysis process following the experimental protocol of Davranche and McMorris (2009) precisely as shown in Fig. 1. In these simulations, we used the ex-Gaussian model as a descriptive model that would produce the type of densities, CDFs, and nDPs shown in the top three panels of Fig. 1. Across simulated experiments, we computed the means and standard deviations of the obtained horizontal and vertical locations of each point on the resulting DPs. The means are depicted as the points in Fig. 8, and the standard deviations, which represent *SE*s in this context, were used to compute the horizontal and vertical error bars (i.e., ±1.96 × *SE*).

Davranche and McMorris (2009) collected in their study n = 200 RTs per condition for each of 12 participants. Under these conditions, with the procedure illustrated in Fig. 1, the estimates of the points along the DP are essentially unbiased in all cases shown in Fig. 8, as is indicated by the fact that the points lie directly on the theoretical curves. The *SE*s, however, can become quite sizable, especially at the right end of a DP, as is shown by the error bars. These *SE*s can also be assessed analytically. Consider, for example, the estimates of the 95 % quantile in the two conditions; for a single participant, their *SE*s are[2] 12.3 ms (congruent RTs) and 9.2 ms (incongruent RTs). The plotting position of the corresponding point on the DP of 1 participant is the average (x-coordinate) and difference (y-coordinate) of these quantiles, and so their *SE*s are,[3] respectively, 7.7 and 15.4 ms. As is shown in the left panel of Fig. 8 (filled circles), the *SE* taken from the 95 % quantiles of 12 independent participants, all operating under exactly the same parameters, are then 2.2. and 4.5 ms. This gives a 95 % confidence ellipse with a vertical extension of about 9 ms and a horizontal extension of about 4.4 ms. Evidently, the outermost points along a DP will be particularly sensitive to the various procedures used for removing outliers (see, e.g., Wascher et al., p. 735; Wylie, Ridderinkhof, et al., 2009, p. 2062; Wylie, van den Wildenberg, et al., 2009, p. 1847; for analytical background on such procedures, see Ulrich & Miller, 1994).

When only four bins are used (e.g., Proctor et al., 2005) rather than ten, different points along the underlying DP are estimated, as indicated by the open circles in the left panel of Fig. 8. For example, the rightmost open circle in the left panel of Fig. 8 refers to estimates at the 87.5 % quantile. Figure 8 shows that the *SE* of comparable points along the DP is roughly similar but tends to be slightly larger than with ten bins. With fewer bins, more data per bin enter into each quantile estimate, but the arithmetical mean of the largest quarter of a sample is necessarily a cruder estimate of the 87.5 % quantile, as compared with an estimate obtained with narrower bins. Figure 8 also illustrates that with just a few bins, it is clearly much harder to extract shape information about the curvature of a DP. For all choices of the number of bins, increasing the number of trials per participant by a factor of, for example, four will cut

[2] The estimate of the pth quantile (x_p) of a distribution has a squared standard error equal to

$$SE^2(\hat{x}_p) = \frac{p(1-p)}{nf^2(x_p)}$$

where n is the sample size and $f(x_p)$ is the underlying density, evaluated at x_p (cf. Kendall & Stuart, 1977, chap. 10.10, Eq. 10.29). For example, the p = 0.95 quantile of the ex-Gaussian density with μ = 283, σ = 20, τ = 40 (used in Fig. 8) is equal to $x_{0.95}$ = 407.8; at this point, the density equals 0.00125. Inserting into the above formula with n = 200 and p = 0.95 gives $SE(\hat{x}_{0.95})$ = 12.32. Of course, in actual applications, one replaces the estimated value $\hat{x}_p$ for the true quantile x_p.

[3] Let t_p^c and t_p^i be the estimates of the pth RT quantile for a single participant in the congruent and incongruent conditions, respectively. These estimates are independent because they come from different conditions, and they have individual variances as given in Footnote 1. The plotting position $(x_p|y_p)$ on the DP is then $x_p = (t_p^c + t_p^i)/2$ and $y_p = (t_p^i - t_p^c)$. Therefore,

$$\mathrm{Var}(x_p) = \frac{1}{4}\left[\mathrm{Var}\left(t_p^c\right) + \mathrm{Var}\left(t_p^i\right)\right]$$

and

$$\mathrm{Var}(y_p) = \mathrm{Var}\left(t_p^c\right) + \mathrm{Var}\left(t_p^i\right),$$

and finally,

$$\mathrm{Cov}(x_p, y_p) = \frac{1}{2}\left[\mathrm{Var}\left(t_p^i\right) - \mathrm{Var}\left(t_p^c\right)\right].$$

These relations explain the unequal *SE*s associated with the two coordinates x_p and y_p forming a DP and their complex dependence structure, which, for nDPs, typically varies from positive to negative correlation as p increases. By way of contrast, we note that the coordinates of QQ-plots (e.g., Marden, 2004) are independent of one another, which makes it easier to derive and interpret the statistical properties of such plots.

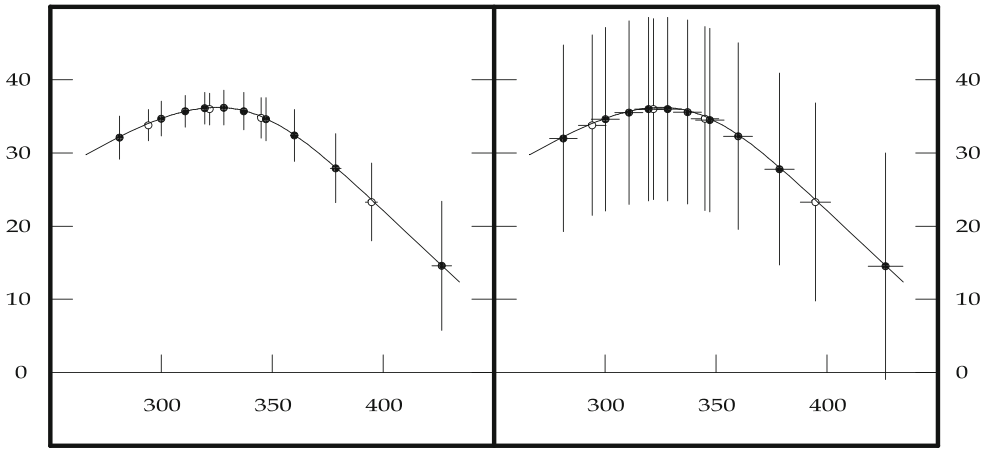

Fig. 8 *Left panel*: Solid line shows a population delta plot, using the ex-Gaussian model with parameters τ =40, μ = 283, σ = 20 in the congruent condition and τ = 30, μ = 325, σ = 25 in the incongruent condition, giving a delta plot (DP) comparable to that in the third panel of Fig. 1. Filled points refer to DP estimates from ten bins, at the 5 %, 15 %, . . . ,95 % quantiles, using 12 participants, each with 200 trials in each condition, as in the study of Davranche and McMorris (2009). Open circles refer to DP estimates from four bins, at the 12.5 %, 37.5 %, 62.5 %, and 87.5 % quantiles, also using 12 participants, each with 200 trials per condition. Mean DP plotting positions and standard errors (*SE*s) were obtained from simulating the model 100,000 times and applying the procedure shown in Fig. 1. The length of the vertical and horizontal error bars indicates ±1.96 × *SE*. These simulations assumed that all 12 participants operate under exactly the same set of parameter values as that described above. *Right panel*: Same as left panel, except that these simulations assumed that the 12 participants operate under parameter values drawn independently from normal distributions with means equal to the fixed parameter values used in the left panel and a standard deviation equal to 5 % of that value

the resultant *SE*s in half and has essentially the same effect as multiplying the number of participants by 4 within this simulation.

Of course, the assumption that all participants operate under exactly the same population parameters seems quite unrealistic, but the right panel of Fig. 8 shows that it is critical. For the simulations depicted in this panel, it was assumed that each parameter value for each of the 12 participants is drawn from a normal distribution that has its mean set at the fixed parameter value used for the left panel of Fig. 8 and that a standard deviation equal to 5 % of each mean parameter value is introduced to model between-participants variability—an amount that would seem to be rather modest. The pattern arising under this more realistic scenario shows that even this small jitter in parameter values between participants increases the *SE*s quite considerably. For example, the 95 % confidence ellipse of the last point on the DP has then a vertical main axis of about 16 ms and a horizontal minor axis of about 7.8 ms and actually extends down to the negative half-plane where stochastic dominance is violated. These *SE*s are, in fact, larger than the *SE*s along the DP of a single fixed participant, illustrating that inter-participant variability can easily outweigh the reduction of the *SE* achieved by Vincentizing. Another important effect of jitter in parameter values between participants—not illustrated in the figure—is that the *SE* then depends to a much lesser degree on the number of trials run per participant. For example, running 800 rather than 200 trials would reduce the vertical and horizontal *SE* of the rightmost data point in the right panel of Fig. 8 only to 14.0 and 7.0 ms, respectively. For a factor of four increase in trials, this represents a rather minuscule reduction in *SE* and indicates that in the presence of even modest parameter jitter, it does not pay to sample each participant extensively. In contrast, running four times as many participants does essentially cut the *SE*s in half.

The results shown in Fig. 8 suggest that it is often unclear whether, for example, a single final negative data point on an empirical DP indicates a true violation of stochastic dominance or simply reflects statistical noise. Similarly, it is often unclear whether a DP has significant concave or convex curvature, especially when as few as four data points are obtained on the DP. Clearly, model-based techniques of DP analysis could assist in resolving whether differences in the form of reported DPs are systematic or simply within the error margin. Another possibility for addressing the issue of statistical noise is to draw on the statistical properties of QQ plots, which are a more thoroughly investigated and better understood data format (for a review, see, e.g., Marden, 2004). As was pointed out by Zhang and Kornblum (1997), a DP is a linear 1–1 transform of a QQ plot; given the extensive literature on their statistical properties, QQ plots may provide more efficient statistical techniques for assessing quantile differences between conditions.

A related point that our results can help to clarify concerns the question of whether one and the same model can be compatible with qualitatively different forms of DPs. For example, Pratte et al. (2010) took the view that models

based on the activation-suppression hypothesis will *always* predict nDPs and, therefore, cannot account for tasks in which nDPs are not routinely found: "Because each of these [inhibition] models explains negatively sloped delta plots, however, they cannot account for the Stroop congruency effect, in which there is a positively sloped delta-plot pattern" (p. 2016; for a similar view, see Proctor et al., 2011, p. 263). Our results indicate that for many models, this implication is too strong. In particular, many of the models we describe as nDP consistent can also predict positive-going DPs with some combinations of parameters. For example, the two-state mixture model will usually predict an nDP under the conditions specified in the Appendix, but it will produce positive-going DPs under various other conditions. Thus, it is certainly possible that the same model might account for both Simon and Stroop congruency effects via differences in parameter values between those two tasks.

In summary, our results confirm previous conclusions (De Jong et al., 1994; Pratte et al., 2010; Proctor et al., 2011; Ridderinkhof, 2002a, 2002b) that DPs can be a valuable data-analytic tool, extending routine analyses of mean RT to a finer distributional level and providing more stringent criteria that any candidate model must pass. In order to take full advantage of the added distributional information provided by such analyses, however, it is necessary to have explicit models that can predict the quantitative form that a DP will take under various experimental conditions. The model classes presented in this article provide examples of the classes that may be considered when attempting to implement such explicit models. Indeed, such models must ultimately be evaluated with respect to their ability to produce realistic RT distributions (CDFs and QQ plots), and good fits to full RT distributions will, as a by-product, naturally imply good fits to DPs as well.

Acknowledgment Work on this manuscript started while the second author was on sabbatical leave at the University of Tübingen, Germany; it was finished while the first author was on sabbatical leave at the University of Otago, Dunedin, New Zealand. We would like to thank both institutions for their hospitality and support.

Appendix Inconsistent moment ordering with mixture models

The binary mixture model predicts that under condition i = 1,2

$$\mathrm{E}[\mathrm{RT}_i] = p_i\mu_a + (1-p_i)\mu_b$$

and the RT variance is

$$\mathrm{Var}[\mathrm{RT}_i] = p_i\sigma_a^2 + (1-p_i)\sigma_b^2 + p_i(1-p_i)(\mu_b-\mu_a)^2$$

We assume that the states themselves show a consistent ordering of RT means and variances, in which case, say, $\mu_a < \mu_b$ and $\sigma_a^2 < \sigma_b^2$. Denote without loss of generality as RT_1 and RT_2 the RT in the condition with faster and slower means, respectively. Because $\mu_a < \mu_b$, this convention implies that $p_1 > p_2$. Inserting for $\mathrm{Var}[\mathrm{RT}_i]$ and elementary simplifications then show that

$$\mathrm{Var}[\mathrm{RT}_1] - \mathrm{Var}[\mathrm{RT}_2] = (p_1-p_2)\left\{(\sigma_a^2-\sigma_b^2) + [1-(p_1+p_2)](\mu_b-\mu_a)^2\right\}.$$

Because $p_1 > p_2$, the sign of $\mathrm{Var}[\mathrm{RT}_1] - \mathrm{Var}[\mathrm{RT}_2]$ is equal to the sign of the expression within the {. . .} brackets. This implies that

$$\mathrm{Var}[\mathrm{RT}_1]\left\{\begin{matrix}>\\=\\<\end{matrix}\right\}\mathrm{Var}[\mathrm{RT}_2] \quad \text{iff} \quad 1-\frac{\sigma_b^2-\sigma_a^2}{(\mu_b-\mu_b)^2}\left\{\begin{matrix}>\\=\\<\end{matrix}\right\}p_1+p_2$$

In particular, $\mathrm{E}[\mathrm{RT}_1]$ will be smaller than $\mathrm{E}[\mathrm{RT}_2]$ but $\mathrm{Var}[\mathrm{RT}_1]$ will be larger than $\mathrm{Var}[\mathrm{RT}_2]$ if, relative to their difference in variances, the state means μ_a and μ_b differ sufficiently and the faster state is used less frequently overall.

References

Atkinson, R. C., Holmgren, J., & Juola, J. F. (1969). Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics, 6*, 321–326. doi:10.3758/BF03212784

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59*, 495–523. doi:10.1016/j.jml.2007.10.004

Burle, B., Possamaï, C.-A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: An electromyographic and distributional analysis. *Psychological Research, 66*, 324–336. doi:10.1007/s00426-002-0105-6

Burle, B., van den Wildenberg, W., & Ridderinkhof, K. R. (2005). Dynamics of facilitation and interference in cue-priming and Simon tasks. *European Journal of Cognitive Psychology, 17*, 619–641. doi:10.1080/09541440540000121

Colonius, H. (1986). Measuring channel dependence in separate activation models. *Perception & Psychophysics, 40*, 251–255. doi:10.3758/BF03211504

Colonius, H., & Ellermeier, W. (1997). Distribution inequalities for parallel models of reaction time with an application to auditory profile analysis. *Journal of Mathematical Psychology, 41*, 19–27. doi:10.1006/jmps.1997.1144

Colonius, H., & Vorberg, D. (1994). Distribution inequalities for parallel models with unlimited capacity. *Journal of Mathematical Psychology, 38*, 35–58. doi:10.1006/jmps.1994.1002

Davelaar, E. J. (2008). A computational study of conflict-monitoring at two levels of processing: Reaction time distributional analyses and hemodynamic responses. *Brain Research, 1202*, 109–119. doi:10.1016/j.brainres.2007.06.068

Davranche, K., & McMorris, T. (2009). Specific effects of acute moderate exercise on cognitive control. *Brain and Cognition, 69*, 565–570. doi:10.1016/j.bandc.2008.12.001

Davranche, K., Paleresompoulle, D., Pernaud, R., Labarelle, J., & Hasbroucq, T. (2009). Decision making in elite white-water

athletes paddling on a kayak ergometer. *Journal of Sport & Exercise Psychology, 31*, 554–565.

De Jong, R., Liang, C. C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus–response correspondence. *Journal of Experimental Psychology. Human Perception and Performance, 20*, 731–750. doi:10.1037/0096-1523.20.4.731

Donders, F. C. (1969). Over de snelheid van psychische processen [On the speed of mental processes] (W. G. Koster, Trans.). In W. G. Koster (Ed.), *Attention and performance II* (pp. 412–431). Amsterdam: North-Holland. (Original work published 1868) doi:10.1016/0001-6918(69)90065-1

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology, 36*, 235–268. doi:10.1016/0022-2496(92)90038-9

Dzhafarov, E. N., & Cortese, J. M. (1996). Empirical recovery of response time decomposition rules: I. Sample-level decomposition tests. *Journal of Mathematical Psychology, 40*, 185–202. doi:10.1006/jmps.1996.0020

Fisher, D. L., & Goldstein, W. M. (1983). Stochastic PERT networks as models of cognition: Derivation of the mean, variance, and distribution of reaction time using order-of-processing (OP) diagrams. *Journal of Mathematical Psychology, 27*, 121–151. doi:10.1016/0022-2496(83)90040-8

Forster, B., & Corballis, M. C. (2000). Interhemispheric transfer of colour and shape information in the presence and absence of the corpus callosum. *Neuropsychologia, 38*, 32–45. doi:10.1016/S0028-3932(99)00050-0

Grice, G. R. (1972). Application of a variable criterion model to auditory reaction time as a function of the type of catch trial. *Perception & Psychophysics, 12*, 103–107. doi:10.3758/BF03212853

Heathcote, A. (2003). Item recognition memory and the receiver operating characteristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1210–1230. doi:10.1037/0278-7393.29.6.1210

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response-time distributions: An example using the Stroop task. *Psychological Bulletin, 109*, 340–347. doi:10.1037//0033-2909.109.2.340

Hommel, B., Proctor, R. W., & Vu, K.-P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research, 68*, 1–17. doi:10.1007/s00426-003-0132-y

Kendall, M. G., & Stuart, A. (1977). *The advanced theory of statistics (Vol. I)* (4th ed.). London: Griffin.

Kornblum, S. (1969). Sequential determinants of information processing in serial and discrete choice reaction time. *Psychological Review, 76*, 113–131. doi:10.1037/h0027245

Kubo-Kawai, N., & Kawai, N. (2010). Elimination of the enhanced Simon effect for older adults in a three-choice situation: Ageing and the Simon effect in a go/no-go Simon task. *Quarterly Journal of Experimental Psychology, 63*, 452–464. doi:10.1080/17470210902990829

Leonhard, T., Bratzke, D., Schröter, H., & Ulrich, R. (2012). Time-course analysis of temporal preparation on central processes. *Psychological Research, 76*, 236–251. doi:10.1007/s00426-011-0364-1

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 883–914. doi:10.1037//0278-7393.18.5.883

Los, S. A., & Schut, M. L. J. (2008). The effective time course of preparation. *Cognitive Psychology, 57*, 20–55. doi:10.1016/j.cogpsych.2007.11.001

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

Marden, J. I. (2004). Positions and QQ plots. *Statistical Science, 19*, 606–614. doi:10.1214/088342304000000512

McClelland, J. L. (1979). On the time relations of mental processes: A framework for analyzing processes in cascade. *Psychological Review, 86*, 287–330. doi:10.1037/0033-295X.86.4.287

Meijers, L. M. M., & Eijkman, E. G. J. (1974). The motor system in simple reaction time experiments. *Acta Psychologica, 38*, 367–377. doi:10.1016/0001-6918(74)90041-9

Meijers, L. M. M., Teulings, J., & Eijkman, E. G. J. (1976). Model of the electromyographic activity during brief isometric contractions. *Biological Cybernetics, 25*, 7–16.

Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed–accuracy decomposition. *Psychological Review, 95*, 183–237. doi:10.1037//0033-295X.95.2.183

Miller, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14*, 247–279. doi:10.1016/0010-0285(82)90010-X

Miller, J. O. (1993). A queue-series model for reaction time, with discrete-stage and continuous-flow models as special cases. *Psychological Review, 106*, 702–715. doi:10.1037//0033-295X.100.4.702

Miller, J. O., & Ulrich, R. (2003). Simple reaction time and statistical facilitation: A parallel grains model. *Cognitive Psychology, 46*, 101–151. doi:10.1016/S0010-0285(02)00517-0

Miller, J. O., van der Ham, F., & Sanders, A. F. (1995). Overlapping stage models and reaction time additivity: Effects of the activation equation. *Acta Psychologica, 90*, 11–28. doi:10.1016/0001-6918(95)00028-S

Navon, D., & Gopher, D. (1979). On the economy of the human information processing system. *Psychological Review, 86*, 214–255.

Navon, D., & Gopher, D. (1980). Task difficulty, resources, and dual-task performance. In R. S. Nickerson (Ed.), *Attention and performance VII* (pp. 297–315). Hillsdale: Erlbaum.

Notebaert, W., & Soetens, E. (2003). The influence of irrelevant stimulus changes on stimulus and response repetition effects. *Acta Psychologica, 112*, 143–156. doi:10.1016/S0001-6918(02)00080-X

Notebaert, W., Soetens, E., & Melis, A. (2001). Sequential analysis of a Simon task: Evidence for an attention-shift account. *Psychological Research, 65*, 170–184. doi:10.1007/s004260000054

Pashler, H. E. (1994a). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244. doi:10.1037//0033-2909.116.2.220

Pashler, H. E. (1994b). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology. Human Perception and Performance, 20*, 330–342. doi:10.1037//0096-1523.20.2.330

Pieters, J. P. M. (1983). Sternberg's additive factor method and underlying psychological processes: Some theoretical considerations. *Psychological Bulletin, 93*, 411–426. doi:10.1037//0033-2909.93.3.411

Pratte, M. S., Rouder, J. N., Morey, R. D., & Feng, C. (2010). Exploring the differences in distributional properties between Stroop and Simon effects using delta plots. *Attention, Perception, & Psychophysics, 72*, 2013–2025. doi:10.3758/APP.72.7.2013

Proctor, R. W., Miles, J. D., & Baroni, G. (2011). Reaction time distribution analysis of spatial correspondence effects. *Psychonomic Bulletin & Review, 18*, 242–266. doi:10.3758/s13423-011-0053-5

Proctor, R. W., Pick, D. F., Vu, K. P. L., & Anderson, R. E. (2005). The enhanced Simon effect for older adults is reduced when the irrelevant location information is conveyed by an accessory stimulus. *Acta Psychologica, 119*, 21–40. doi:10.1016/j.actpsy.2004.10.014

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108. doi:10.1037//0033-295X.85.2.59

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446–461. doi:10.1037//0033-2909.86.3.446

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. doi:10.3758/BF03196283

Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology. Human Perception and Performance, 26*, 127–140. doi:10.1037//0096-1523.26.1.127

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*, 333–367. doi:10.1037/0033-295X.111.2.333

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaching to dealing with contaminant reaction and parameter variability. *Psychonomic Bulletin & Review, 9*, 438–481. doi:10.3758/BF03196302

Ridderinkhof, K. R. (2002a). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention & Performance XIX* (pp. 494–519). Oxford: Oxford University Press.

Ridderinkhof, K. R. (2002b). Micro- and macro-adjustments of task set: Activation and suppression in conflict tasks. *Psychological Research, 66*, 312–323. doi:10.1007/s00426-002-0104-7

Ridderinkhof, K. R., Scheres, A., Oosterlaan, J., & Sergeant, J. A. (2005). Delta plots in the study of individual differences: New tools reveal response inhibition deficits in AD/HD that are eliminated by Methylphenidate treatment. *Journal of Abnormal Psychology, 114*, 197–215. doi:10.1037/0021-843X.114.2.197

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Wijnen, J., & Burle, B. (2004). Response inhibition in conflict tasks is revealed in delta plots. In M. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 369–377). New York: Guilford.

Ruthruff, E. D. (1996). A test of the deadline model for speed–accuracy tradeoffs. *Perception & Psychophysics, 58*, 56–64. doi:10.3758/BF03205475

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology. Human Perception and Performance, 26*, 141–151. doi:10.1037//0096-1523.26.1.141

Schwarz, W. (1994). Diffusion, superposition, and the redundant-targets effect. *Journal of Mathematical Psychology, 38*, 504–520. doi:10.1006/jmps.1994.1036

Schwarz, W. (2003). Stochastic cascade processes as a model of multistage concurrent information processing. *Acta Psychologica, 113*, 231–261. doi:10.1016/S0001-6918(03)00032-5

Schwarz, W., & Ischebeck, A. (2001). On the interpretation of response time vs onset asynchrony functions: Application to dual-task and precue-utilization paradigms. *Journal of Mathematical Psychology, 45*, 452–479. doi:10.1006/jmps.2000.1336

Schwarz, W., & Ischebeck, A. (2003). On the relative speed account of number-size interference in comparative judgments of numerals. *Journal of Experimental Psychology. Human Perception and Performance, 29*, 507–522. doi:10.1037/0096-1523.29.3.507

Schweickert, R. (1982). The bias of an estimate of coupled slack in stochastic PERT networks. *Journal of Mathematical Psychology, 26*, 1–12. doi:10.1016/0022-2496(82)90032-3

Schweickert, R., & Mounts, J. (1998). Additive effects of factors on reaction time and evoked potentials in continuous-flow models. In C. E. Dowling, F. S. Roberts, & P. Theuns (Eds.), *Recent progress in mathematical psychology: Psychophysics, knowledge, representation, cognition, and measurement* (pp. 311–327). Hillsdale: Erlbaum.

Schweickert, R., & Townsend, J. T. (1989). A trichotomy: Interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (PERT) networks. *Journal of Mathematical Psychology, 33*, 328–347. doi:10.1016/0022-2496(89)90013-8

Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the mind's bottleneck. *PLoS Biology, 3*, 334–349.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315. doi:10.1016/0001-6918(69)90055-9

Taylor, D. A. (1976). Stage analysis of reaction time. *Psychological Bulletin, 83*, 161–191. doi:10.1037//0033-2909.83.2.161

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology, 28*, 363–400. doi:10.1016/0022-2496(84)90007-5

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

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136. doi:10.1016/0010-0285(80)90005-5

Ulrich, R., Mattes, S., & Miller, J. O. (1999). Donders's assumption of pure insertion: An evaluation on the basis of response dynamics. *Acta Psychologica, 102*, 43–75. doi:10.1016/S0001-6918(99)00019-0

Ulrich, R., & Miller, J. O. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology. General, 123*, 34–80. doi:10.1037//0096-3445.123.1.34

Vallesi, A., Mapelli, D., Schiff, S., Amodio, P., & Umiltà, C. (2005). Horizontal and vertical Simon effect: Different underlying mechanisms? *Cognition, 96*, B33–B43. doi:10.1016/j.cognition.2004.11.009

van den Wildenberg, W. P. M., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., & Ridderinkhof, K. R. (2010). To head or to heed? Beyond the surface of selective action inhibition: A review. *Frontiers in Human Neuroscience, 5*, 1–13. doi:10.3389/fnhum.2010.00222

Van der Heijden, A. H. C., Schreuder, R., Maris, L., & Neerincx, M. (1984). Some evidence for correlated separate activation in a simple letter-detection task. *Perception & Psychophysics, 36*, 577–585.

Van Duren, L. L., & Sanders, A. F. (1988). On the robustness of the additive factors stage structure in blocked and mixed choice reaction designs. *Acta Psychologica, 69*, 83–94. doi:10.1016/0001-6918(88)90031-5

Wagenmakers, E.-J., & Brown, S. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. *Psychological Review, 114*, 830–841. doi:10.1037/0033-295X.114.3.830

Wascher, E., Schatz, U., Kuder, T., & Verleger, R. (2001). Validity and boundary conditions of automatic response activation in the Simon task. *Journal of Experimental Psychology. Human Perception and Performance, 27*, 731–751. doi:10.1037//0096-1523.27.3.731

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114*, 152–176. doi:10.1037/0033-295X.114.1.152

Wylie, S. A., Ridderinkhof, K. R., Bashore, T. R., & van den Wildenberg, W. P. M. (2009). The effect of Parkinson's disease on the dynamics of on-line and proactive cognitive control during action selection. *Journal of Cognitive Neuroscience, 22*, 2058–2073. doi:10.1162/jocn.2009.21326

Wylie, S. A., Ridderinkhof, K. R., Elias, W. J., Frysinger, R. C., Bashore, T. R., Downs, K. E., . . . van den Wildenberg, W. P. M. (2010). Subthalamic nucleus stimulation influences expression and suppression of impulsive behavior in Parkinson's disease. *Brain, 133*, 3611–3624. doi:10.1093/brain/awq239

Wylie, S. A., van den Wildenberg, W. P. M., Ridderinkhof, K. R., Bashore, T. R., Powell, V. D., Manning, C. A., & Wooten, G. F. (2009). The effect of speed–accuracy strategy on response interference control in Parkinson's disease. *Neuropsychologia, 47*, 1844–1853. doi:10.1016/j.neuropsychologia.2009.02.025

Yantis, S., Meyer, D. E., & Smith, J. E. K. (1991). Analyses of multinomial mixture distributions: New tests for stochastic models of cognition and action. *Psychological Bulletin, 110*, 350–374. doi:10.1037//0033-2909.110.2.350

Yellott, J. I., Jr. (1971). Correction for fast guessing and the speed–accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology, 8*, 159–199. doi:10.1016/0022-2496(71)90011-3

Zhang, J., & Kornblum, S. (1997). Distributional analysis and De Jong, Liang, and Lauber's (1994) dual-process model of the Simon effect. *Journal of Experimental Psychology. Human Perception and Performance, 23*, 1543–1551. doi:10.1037//0096-1523.23.5.1543