Psychological Review
2015, Vol. 122, No. 4, 830–837

0033-295X/15/$12.00 http://dx.doi.org/10.1037/a0039658

THEORETICAL NOTE

Sequential Processes and the Shapes of Reaction Time Distributions

Saul Sternberg
University of Pennsylvania

Benjamin T. Backus
State University of New York, College of Optometry

It is sometimes suggested that reaction time (RT) distributions have the same shape across conditions or groups. In this note we show that this is highly unlikely if the RT is the sum of the stochastically independent durations of 2 or more stages (sequential processes) (a) that are influenced selectively by different factors, or (b) 1 of which is influenced selectively by some factor. We provide an example of substantial shape differences in RT data from a flash-detection experiment, data that have been shown to satisfy requirement (a). Ignoring these requirements, we also note that in a large range of instances reviewed by Matzke and Wagenmakers (2009) in which the ex-Gaussian distribution was fitted to RT data from different conditions in the same experiment, most sets of distributions fail to satisfy even a weak requirement for shape invariance. In the Appendix we describe the Summation Test for selectively influenced stages with independent durations (Roberts & Sternberg, 1993), and provide an example of its application.

Keywords: reaction time (RT), stage model, selective influence, RT distribution, ex-Gaussian distribution

Under what conditions do reaction time (RT) distributions have the same shape across conditions or groups? Shape invariance of a set of RT distributions means that they differ by at most their means and time scales. Thus, the distributions of X and Y have the same shape if and only if there are constants a and b such that $Y = a + bX$. For example, one proposal about the cognitive effects of aging is the controversial General Slowing Hypothesis: With increasing age, all the operations of the central nervous system in most or all tasks become proportionally slower (Cerella, 1985; Eckert, 2011; Myerson et al., 2003a; Myerson, Hale, Zheng, Jenkins, & Widaman, 2003b; Salthouse, 1996; Sleimen-Malkoun, Temprado, & Berton, 2013; but see also Bashore et al., 2014, and Ratcliff et al., 2000). In effect, with increasing age, time runs more slowly. Rouder, Yue, Speckman, Pratte, and Province (2010) discuss other considerations that lead to shape invariance. Ratcliff and McKoon (2008) use the approximate linearity of Q-Q plots to argue that for diffusion model predictions and some data sets, the shapes of RT distributions are approximately invariant across experimental conditions and experiments (p. 895). And, according to Ratcliff and Smith (2010, p. 90), "Invariance of distribution shape is one of the most powerful constraints on models of RT distributions. . . . That the diffusion model predicts this invariance is a strong argument in support of its use in performing process decomposition of RT data."

The primary purpose of this note is to show that for a process organized in stages that have stochastically independent durations and are selectively influenced by experimental factors, it is highly unlikely that the distributions of RTs in several conditions in an experiment can have the same shape.

Stage Models

Stage models are ubiquitous in research on speeded tasks (e.g., King & Dehaene, 2014; Sanders, 1998; Schall, 2003; Schall et al., 2011; Sigman & Dehaene, 2008; Sternberg, 1998, 2001) and elsewhere (Borst & Anderson, 2015). For several sets of RT data, Roberts and Sternberg (1993) provide evidence for selectively influenced stages whose durations are stochastically independent. Even in Ratcliff's (1978) diffusion model, in which the "one-shot" decision process (Ratcliff & Tuerlinckx, 2002, p. 439) is represented by multiple activations that grow in parallel, the decision process **D** is augmented by two additional stages arranged sequentially whose durations are stochastically independent: an initial stage **E** for stimulus encoding, and a final stage **R** for response execution. In the application of the diffusion model considered by Gomez, Perea, and Ratcliff (2013), the duration of **E** in a lexical-decision task is found to be selectively influenced by the relatedness of masked primes. In the experiments considered by Ratcliff and Smith (2010), **E** delays the start of **D** by an amount that is changed by 100 ms or more by variations in stimulus noise. Because the same factor also affects **D,** its influence with respect to **E** and **D** is not selective; however, its influence is selective with respect to **E** and **R,** and **D** and **R.**

Saul Sternberg, Department of Psychology, University of Pennsylvania; Benjamin T. Backus, Graduate Center for Vision Research, State University of New York, College of Optometry.

The authors thank Sylvan Kornblum, David Krantz, Frank Norman, Richard Schweickert, Jacob Sternberg-Sher (age 9), James Townsend, and Eric-Jan Wagenmakers for helpful comments and suggestions. Benjamin T. Backus collaborated in formulating the experimental problem and collecting the data used in the example. Saul Sternberg developed the theory, performed the recent data analyses, and wrote the article.

Correspondence concerning this article should be addressed to Saul Sternberg, Suite 400A, 3401 Walnut Street, Philadelphia, PA 19104-6228. E-mail: saul@psych.upenn.edu

Two Stages With Selective Effects on Both

Because two or more stages can be concatenated and treated as a single stage, we can limit consideration to processes consisting of two stages without loss of generality. Consider a process that consists of stages **A** and **B** with durations T_A and T_B, so that the RT is $RT = T_A + T_B$. Assume that T_A and T_B are stochastically independent. We then have an *SIStage process* (a process consisting of sequential operations whose durations are stochastically independent; Roberts & Sternberg, 1993). Suppose two factors, F_j and G_k, each with two levels, $j = 1, 2$, and $k = 1, 2$, that influence the stage durations selectively, so that $T_A = T_A(F_j) = T_{Aj}$, $T_B = T_B(G_k) = T_{Bk}$, and $RT_{jk} = T_{Aj} + T_{Bk}$. Consider a 2×2 factorial experiment with the four resulting conditions, giving us RT_{11}, RT_{12}, RT_{21}, and RT_{22}. (Because an $m \times n$ experiment can be regarded as a concatenation of 2×2 experiments, we can do so without loss of generality.) Then, because convolution is associative and commutative, $RT_{11} + RT_{22}$ has the same distribution as $RT_{12} + RT_{21}$, namely, the convolution of the distributions of T_{A1}, T_{A2}, T_{B1}, and T_{B2}. Thus,

$$RT_{11} * RT_{22} = RT_{12} * RT_{21}, \tag{1}$$

where "*" represents convolution (Ashby & Townsend, 1980, p. 108). It follows that:

$$\kappa_{r11} + \kappa_{r22} = \kappa_{r12} + \kappa_{r21}, \quad (r \geq 1), \tag{2}$$

where $\kappa_{rjk} = \kappa_r(RT_{jk})$ is the *rth* cumulant of RT_{jk}. Let us assume that RT distributions are "well-behaved," in the sense that cumulants of (at least) orders $r = 1, 2, 3$, and 4 exist.[1]

Equation 2 results from three assumptions: (a) *stages,* (b) *stochastic independence,* and (c) *selective influence.* To these, let us add a fourth assumption: (d) *shape invariance*: The RT_{jk} distributions differ by at most means and scale factors. Whereas differences among means influence only the means of the RT_{jk} distributions and influence none of the cumulants above the first, differences among scale factors influence all of the cumulants and central moments above the first. Now, if the distributions of two random variables, X_1 and X_2 have the same shape, with scale factor C, then

$$\kappa_r(X_2)/\kappa_r(X_1) = C^r, \quad (r \geq 2). \tag{3}$$

Let the scale factor associated with RT_{jk} be $C_{jk} > 0$. It follows from Equation 3 that

$$\kappa_{rjk} = \kappa_{r00} C_{jk}^r, \quad (r \geq 2, \ \kappa_{r00} \neq 0). \tag{4}$$

where the $\{\kappa_{r00}\}$ are a set of constants, one for each r. With Assumption (d), Equations 2 and 4 then imply that:

$$C_{11}^r + C_{22}^r = C_{12}^r + C_{21}^r, \quad (r \geq 2). \tag{5}$$

Given that $\kappa_{200} \neq 0$ (nonzero variances) and $\kappa_{400} \neq 0$ (nonzero kurtosis values, which, among common distributions, excludes only the Gaussian) there are only three relations among the C_{jk} that satisfy Equation 5:[2]

(*i*) the C_{jk} are identical,
(*ii*) $C_{11} = C_{21}$ and $C_{22} = C_{12}$,
(*iii*) $C_{11} = C_{12}$ and $C_{22} = C_{21}$.

Given (*i*), only the mean *RT* and none of the higher cumulants can be influenced by either factor. Given (*ii*), factor *F* can influence only the mean. Given (*iii*), factor *G* can influence only the mean. Thus, for the four RT distributions to have the same shape, at least one of the two factors can cause no more than a *shift* (a change in mean only) of the RT distribution, a highly unlikely possibility.[3]

It is remarkable that whereas we have shown that the four distributions RT_{11}, RT_{12}, RT_{21}, and RT_{22} are very likely to differ, the relations among them must be such that when they are combined in pairs, as in Equation 1, those differences "cancel out."

Two Stages With a Selective Effect on One

This is sometimes assumed or concluded in applications of Ratcliff's (1978) diffusion model. Suppose the two-stage model, with one factor F_j ($j = 1, 2, \ldots$) that influences just T_A, so that

$$RT_j = T_{Aj} + T_B, \quad (j \geq 1). \tag{6}$$

We then have

$$\kappa_{rj} = \alpha_{rj} + \beta_r, \quad (j \geq 1, \ r \geq 1), \tag{7}$$

where κ_{rj}, α_{rj}, and β_r are the *rth* cumulants of RT_j, T_{Aj}, and T_B, respectively. Equation 7 follows from assumptions (a), (b), and (c), above. Addition of the *shape invariance* assumption then requires

$$\kappa_{rj} = C_j^r \kappa_{r1}, \quad (j \geq 2, \ r \geq 2), \tag{8}$$

where $C_j \neq 1$ is the scale factor that relates RT_j to RT_1. Combining Equations 7 and 8 and rearranging, we have

$$\beta_r = \frac{(\alpha_{rj} - C_j^r \alpha_{r1})}{(C_j^r - 1)}, \quad (j \geq 2, \ r \geq 2). \tag{9}$$

Thus, either T_B is a constant ($\beta_r = 0$, $r \geq 2$) or its distribution (which is uniquely determined up to its mean by the $\{\beta_r\}$, $r \geq 2$) is restricted by properties of the distribution of T_A, and may vary with the level of the factor F_j that is assumed to influence only T_A, a contradiction.

[1] In what follows, two well-known properties of cumulants (κ_r) of order r are used: For stochastically independent random variables X and Y, $\kappa_r(X + Y) = \kappa_r(X * Y) = \kappa_r(X) + \kappa_r(Y)$; also, $\kappa_r(CX) = C^r\kappa_r(X)$. Note that $\kappa_{rjk} = M_{rjk}$, for $1 \leq r \leq 3$ and $\kappa_{4jk} = M_{4jk} - 3M_{2jk}^2$, where the $\{M_{rjk}\}$ are the mean and *rth* central moments of RT_{jk} (see Kendall & Stuart, 1969, Volume 1, Ch. 3.). The quantity κ_4/κ_2^2 is a common measure of kurtosis, whose value is zero for the Gaussian distribution, and nonzero for other common distributions (Weisstein, 2014).

[2] To prove this, use Equation 5 with $r = 2$ and $r = 4$. For simplicity, let $C_{11} = a$, $C_{22} = b$, $C_{12} = c$, and $C_{21} = d$. Start with (A) $a^2 + b^2 = c^2 + d^2$ and (B) $a^4 + b^4 = c^4 + d^4$. Express the two sides of (A) and (B), respectively, in terms of $(a + b)^2$ and $(c + d)^2$, and of $(a + b)^4$ and $(c + d)^4$. Square the two sides of the equation derived from (A), and subtract from the equation derived from (B). This gives $ab = cd$, or $a/c = d/b = k$. Substituting in (A) gives $(k^2 - 1)(c^2 - b^2) = 0$. This implies either that $k = 1$, which means that $a = c$ and $b = d$; or that $b = c$, which requires $a = d$.

[3] Effects of factors on mean RT are almost always associated with nonzero effects on other aspects of the distribution, including var(RT). Indeed, Wagenmakers and Brown (2007) have argued for a lawful regularity in the relation between mean and variance: they claim that the standard deviations (SDs) of RT distributions increase linearly with their means. It will be seen that for the data to be presented below, effects on the mean are indeed accompanied by effects on the SDs, but that the relation between them is nonmonotonic.

Shape Differences of RT Distributions in Flash Detection

An example is provided by an experiment first reported by Backus and Sternberg (1988). It is called "Experiment 1" by Roberts and Sternberg (1993), who used the *summation test,* explained and applied in that article, to show that the data are consistent with a SIStage model with selective influence.[4] Subjects responded by pulling a lever if, after a variable foreperiod, they detected a flash in one of four locations. A central cue at the start of each trial indicated the most likely location. On 25% of the trials ("catch" trials) there was no flash. The factors foreperiod (six levels) and flash intensity (two levels) varied approximately randomly and independently from trial to trial.

For testing the SIStage model we used only the data from trials when the cue was valid and when the foreperiod was either 750 ms or 1,150 ms. Six subjects provided these data, each of whom served for six 1-hr test sessions after 3 hrs of practice. The data considered here thus reflect four conditions that are, from shortest to longest mean RT, (short, bright: Sb), (long, bright: Lb), (short, dim: Sd), and (long, dim: Ld). As shown in Table 26.2 of Roberts and Sternberg (1993), the overall mean RT is 222 ms; the main effects of foreperiod and intensity on mean RT are 15 ms and 36 ms, respectively, and their interaction is close to zero. That the summation test is well satisfied is shown in the Appendix of the present article as well as in Figures 26.2A and 26.3A in Roberts and Sternberg (1993); Figure 26.2A also shows the four distribution functions.

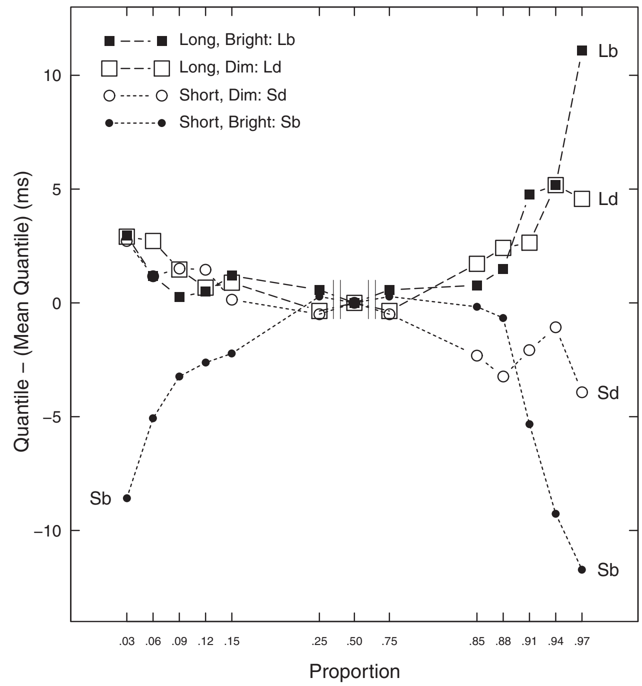

Figure 1. Data from four conditions in a detection experiment in which foreperiod and flash intensity were varied factorially: Means over five subjects of differences between the quantiles of normalized distributions in each condition and their means. To enable better visualization of the tails of the distributions, the *x*-axis is nonlinear, with breaks marked by vertical lines.

L-Skewness and L-Kurtosis

Skewness and kurtosis are frequently used to describe the shapes of distributions. We used measures based on L-moments for their evaluation.[5] To reduce heterogeneity for the present analysis, the first of the test sessions was omitted, as were the data for one subject whose variability was exceptionally high.[6] For each of the five subjects and each of the four conditions, the data were pooled over the remaining five test sessions. This resulted in 20 distributions of 80 observations each. For each distribution, estimates of the first four L-moments λ_1, λ_2, λ_3, and λ_4 and the derived measures of skewness, $\tau_3 = \lambda_3/\lambda_2$ ("L-skewness") and kurtosis, $\tau_4 = \lambda_4/\lambda_2$ ("L- kurtosis"), were calculated. Means and standard errors over the six subjects of estimates of λ_1, λ_2, τ_3, and τ_4 for the four conditions are provided in Table 1, which also includes the average standard deviation (*SD*) for each condition.

Table 1
Means and Standard Errors of Estimates of Five Parameters

Condition	Sb	Lb	Sd	Ld	
Foreperiod: Intensity:	Short Bright	Long Bright	Short Dim	Long Dim	
Measure					$\widehat{SE}$
$\hat{\lambda}_1$ (*ms*)	197.8	210.8	230.5	243.1	1.0
SD (*ms*)	15.94	25.30	21.01	26.19	.72
$\hat{\lambda}_2$ (*ms*)	8.78	12.57	11.37	14.00	.31
$\hat{\tau}_3$	.050	.259	.162	.215	.015
$\hat{\tau}_4$	.181	.242	.181	.185	.008

Note. Rows = parameters; columns = conditions. $\hat{\lambda}_1$ = estimate of the first L-moment = mean; *SD* = estimate of the standard deviation; $\hat{\lambda}_2$ = estimate of second L-moment, a measure of variability; $\hat{\tau}_3 = \hat{\lambda}_3/\hat{\lambda}_2$ = estimate of L-skewness; $\hat{\tau}_4 = \hat{\lambda}_4/\hat{\lambda}_2$ = estimate of L-kurtosis. Values from data for five subjects pooled over five sessions. Each $\widehat{SE}$ is based on the 12 *df Subjects* × *Conditions* mean square in an ANOVA.

Interpretations of $\hat{\tau}_3$ and $\hat{\tau}_4$ may be guided by the fact that they fall within the unit interval, $0 \leq \tau_3, \tau_4 \leq 1$, that for the exponential distribution, $\tau_3 = .333$ and $\tau_4 = .167$, and that for the Gaussian distribution, $\tau_3 = 0$ and $\tau_4 = .123$. The differences among the $\hat{\tau}_3$ values across conditions are striking: all five subjects show differences in the same direction for Sb versus Lb and Sb versus Ld, and four of the five show a difference in the same direction for Sb versus Sd. The mean difference between the means of the $\hat{\tau}_3$ values for Conditions Lb, Sd, and Ld, and of the values for Condition Sb, with $\pm$ *SE*, is 0.15 $\pm$ 0.03; a *t* test yields $p < .01$. (A similar comparison for the conventional measure of skewness, $\kappa_3/(\kappa_2^{3/2})$, yields a difference of 1.37–0.37 = 1.00 $\pm$ 0.36, and $p < .05$.) In an ANOVA in which effects were compared with their interaction with subjects, the effects on $\hat{\tau}_3$ of foreperiod, intensity, and their interaction yielded *p*

[4] See the Appendix for a description and an application of the summation test.

[5] L-Moments, λ_k, k = 1, 2, . . . , are linear combinations of order statistics that are influenced less than conventional moments by extreme observations, and have other desirable properties (Hosking, 1990, 1992, 2006; Hosking & Wallis, 1997; Jones, Rosco, & Pewsey, 2011; Royston, 1992). Calculations were performed using the R-package "lmom."

[6] For that subject, $\bar{\lambda}_2$, a measure of variability averaged over the four conditions, was 18.2; for the five other subjects, $11.1 < \bar{\lambda}_2 < 12.6$.

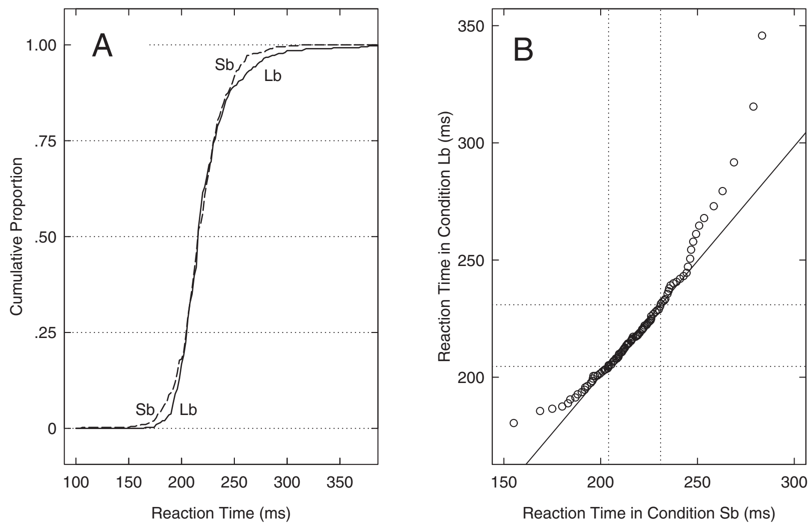

Figure 2. Effect of foreperiod on detection RT for bright flashes. A: Mean over five subjects of distribution functions of normalized RTs for Lb and Sb conditions. B: Mean quantile-quantile plot of the same pair of distributions, with 25% and 75% points indicated for each.

values of 0.03, 0.04, and 0.06, respectively. In a similar ANOVA for $\hat{\tau}_4$, the effect of foreperiod and its interaction with intensity yielded p values of 0.05 and 0.03, respectively. We can conclude that the different conditions produce RT distributions that differ in shape.

How great are the effects on τ_3? One basis for comparison is the increase in skewness with the size of the positive set, n_{pos}, in "memory scanning," a shape difference that has been emphasized by several investigators (e.g., Hockley, 1984; Hockley & Corballis, 1982; McElree & Dosher, 1989). In simulations of the ex-Gaussian distribution based on Hockley's parameter estimates (Hockley, 1984, Figure 4), the difference between the largest $\hat{\tau}_3$ (for $n_{pos} = 6$) and the smallest (for $n_{pos} = 3$) is $.296 - .211 = .085$ for positive responses and $.302 - .267 = .035$ for negative responses. In contrast, as shown in Table 1, the difference between the largest mean $\hat{\tau}_3$ (in Condition Lb of the present experiment) and the smallest (in Condition Sb), $.257 - .067 = .190$, is twice as large as the larger of Hockley's differences.

Quantiles of Normalized Distributions

To further explore the shape differences indicated by the effects on $\hat{\tau}_3$ and $\hat{\tau}_4$, we transformed the RTs linearly to normalize the 20 distributions so that they had equal medians (217 ms) and interquartile ranges (25 ms), equal to the means across the distributions of their medians and interquartile ranges, respectively. This enabled us to compare quantiles across subjects and conditions.[7] We did so because we believe that systematic differences are more likely to occur at points with equal proportions than at points with equal RTs.[8] For each normalized distribution a set of quantiles was estimated. Let q_{pcs} be the quantile for a given proportion, *p,* condition, *c,* and subject, *s.* From the $\{q_{pcs}\}$, their means over conditions, $\{q_{p\bullet s}\}$, and the differences $Q_{pcs} = q_{pcs} - q_{p\bullet s}$ could be determined. It is the $\{Q_{pc\bullet}\}$, the means over subjects of these differences, that are shown in Figure 1. If the distributions had the same shape, then, except for variations due to sampling error, the $\{Q_{pc\bullet}\}$ would all be zero. And, to the extent that quantile differences across conditions are large relative to quantile differences across subjects within conditions, we can conclude that the differences among conditions are real.

The interaction of the effects of foreperiod and intensity on τ_3 and τ_4 (striking, given that the effects of these factors on $\overline{RT}$ are additive) is also shown by their effects on the quantiles for both low and high tails: The effects of foreperiod on shape are substantially greater when the flash is bright than when it is dim. Separate ANOVAs for low and high tails show that proportion interacts significantly with condition (low tail: $p < .0001$; high tail: $p = .002$) and with the interaction of foreperiod with intensity (low tail: $p < .01$; high tail: $p < .01$). In an ANOVA in which tail (low or high) is a factor, and proportion is measured outward from 0.5, the interaction of proportion, condition, and tail is highly significant ($p < .0001$), confirming the impression that the effects of condition on the high tail are greater than on the low tail. That a separation between the Sd and Lb conditions shows up only for the high tail suggests a qualitative difference between the two tails: for the interaction of condition and tail in an ANOVA of just the data for Sd and Lb we found $p = .06$. We also noticed that, as shown by the mean squares in ANOVAs, variability across subjects is substantially greater for the high tail than the low tail: ratios of mean squares for *Proportion* × *Foreperiod* × *Subjects, Proportion* × *Interval* × *Subjects,* and *Proportion* × *Foreperiod* × *Interval* × *Subjects,* are 6.1, 4.3, and 8.1, with $p < .0001$ in each case.

To aid in understanding Figure 1, two additional ways of comparing the shapes of distributions are shown in Figure 2, in which the mean normalized distributions for the two conditions with the most contrasting shapes (Sb and Lb) are shown. In the quantile-

[7] All quantile estimates used the Hyndman and Fan (1996) Type 8 estimator.

[8] This goal is similar to that of Ratcliff's (1979) "Vincentizing" procedure.

quantile analysis[9] (Figure 2B), the plots for all five subjects are concave upward: the quadratic coefficient is significantly positive, with $p = .01$. Yet another way to compare these distributions is shown in Panel B of Figure A1 in the Appendix.

It seems likely that these effects on the shapes of RT distributions reflect interesting properties of the underlying process; it remains to be determined what these properties are.

Shape Invariance and the Ex-Gaussian Distribution

Because the ex-Gaussian distribution has been fitted to numerous sets of RT data, it is interesting to ask about the conditions under which two different ex-Gaussian distributions have the same shape. For the exponential distribution with scale parameter δ, the first three cumulants are δ, δ^2, $2\delta^3$, and in general, $\kappa_r = (r - 1)!\delta^r$. Those of the Gaussian distribution are μ, σ^2, 0 and for $r > 3$, $\kappa_r = 0$. The cumulants of the ex-Gaussian distribution are therefore the sums, $\delta + \mu$, $\delta^2 + \sigma^2$, $2\delta^3$, and, for $r > 3$, $(r - 1)!\delta^r$. It is easy to show[10] from Equation 3 that two different ex-Gaussian distributions have the same shape if and only if $\sigma_2/\sigma_1 = \delta_2/\delta_1$. Thus a minimum requirement for sameness of shape is that any factor that influences either δ or σ should also influence the other, and in the same direction. Yet in the Matzke and Wagenmakers (2009, Supplemental Materials) inventory of ex-Gaussian analyses, even this weak requirement is met in only 31 (about 21%) of the 147 cases where effects on δ and σ and their directions were observed.[11] Thus, to the extent that the ex-Gaussian distribution fit well, the shapes of most of the RT distributions that were analyzed were influenced by factor levels, violating shape invariance.

Conclusions

Given stages with variable and independent durations, and factors that influence more than the means of those durations selectively, the RT distributions in a factorial experiment are highly unlikely to have the same shape. Also, given variable and independent durations, and a factor that influences more than the mean of just one of those durations, the RT distributions for different levels of that factor are highly unlikely to have the same shape. Evidence from the fitting of the ex-Gaussian distribution to RT distributions has often revealed differences in shape. It follows that any theory that predicts shape invariance must be of limited generality. How well the distributions produced by an SIStage process can *approximate* shape invariance is a question for further research. (To the extent that the answer is "poorly," the observation of shape invariance in a particular case would constitute evidence against the SIStage model in that case.) In thinking about this issue it is important to consider the sensitivity of the standard tests for differences between distributions, and of the associated graphical displays. Acknowledgment of the existence of shape differences among RT distributions may lead to further understanding of the underlying processes.

---

[9] Quantiles associated with proportions .01, .02, . . . , .99 were determined for each distribution and each subject, their means over subjects determined, and the mean quantiles for Lb plotted against those for Sb.

[10] Let C be the scale factor that distinguishes the two distributions. From Equation 3, $\kappa_{32}/\kappa_{31} = \delta_2^3/\delta_1^3 = C^3$, and $\kappa_{22}/\kappa_{21} = (\delta_2^2 + \sigma_2^2)/(\delta_1^2 + \sigma_1^2) = C^2$. The first of these implies $\delta_2/\delta_1 = C$. Combining this with the second gives $\sigma_2/\sigma_1 = \delta_2/\delta_1$, which is thus a necessary condition for sameness of shape. And because $\kappa_{r2}/\kappa_{r1} = \delta_2^r/\delta_1^r = C^r$ for $r > 2$, it is also a sufficient condition. See also Thomas and Ross (1980, pp. 143–144), who show that this condition is required for two ex-Gaussian distributions to be members of the same "family."

[11] The requirement that both effects should be present and in the same direction is weak because it can be satisfied when the requirement that the effects be proportional is violated. We used this weaker requirement because the information in the Matzke and Wagenmakers inventory included the presence and direction of effects, but not their magnitudes.

References

Ashby, F. G., & Townsend, J. T. (1980). Decomposing the reaction-time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology, 21,* 93–123.

Backus, B. T., & Sternberg, S. (1988, November). *Attentional tradeoff across space early in visual processing*. Paper presented at the Psychonomic Society Annual Meeting, Chicago.

Bashore, T. R., Wylie, S. A., Ridderinkhof, K. R., & Martinerie, J. M. (2014). Response-specific slowing in older age revealed through differential stimulus and response effects on P300 latency and reaction time. *Aging, Neuropsychology, and Cognition, 21,* 633–673.

Borst, J. P., & Anderson, J. R. (2015). The discovery of processing stages: Analyzing EEG data with hidden semi-Markov models. *Neuroimage, 108,* 60–73.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin, 98,* 67–83.

Eckert, M. A. (2011). Slowing down: Age-related neurobiological predictors of processing speed. *Frontiers in Neuroscience, 5,* 25. http://dx.doi.org/10.3389/fnins.2011.00025

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked versus unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1731–1740.

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

Hockley, W. E., & Corballis, M. C. (1982). Tests of serial scanning in item recognition. *Canadian Journal of Psychology, 36,* 189–212.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society B, 52,* 105–124.

Hosking, J. R. M. (1992). Moments or L-moments? An example comparing two measures of distributional shape. *The American Statistician, 16,* 186–189.

Hosking, J. R. M. (2006). On the characterization of distributions by their L-moments. *Journal of Statistical Planning and Inference, 136,* 193–198.

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis: An approach based on L- moments*. Cambridge, UK: Cambridge University Press.

Hyndman, R. J., & Fan, Y. (1996). Sample quantiles in statistical packages. *The American Statistician, 50,* 361–365.

Jones, M. C., Rosco, J. F., & Pewsey, A. (2011). Skewness-invariant measures of kurtosis. *The American Statistician, 65,* 89–95.

Kendall, M. G., & Stuart, A. (1969). *The advanced theory of statistics, Vol. 1*. New York, NY: Hafner.

King, J-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences, 18,* 203–210.

Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16,* 798–817.

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General, 118,* 346–373.

Myerson, J., Adams, D. R., Hale, S., & Jenkins, L. (2003a). Analysis of group differences in processing speed: Brinley plots, Q-Q plots, and other conspiracies. *Psychonomic Bulletin & Review, 10,* 224–237.

Myerson, J., Hale, S., Zheng, Y., Jenkins, L., & Widaman, K. W. (2003b). The difference engine: A model of diversity in speeded cognition. *Psychonomic Bulletin and Review, 10,* 262–288.

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

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446–461.

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

Ratcliff, R., & Smith, P. L. (2010). Perceptual discrimination in static and dynamic noise: The temporal relation between perceptual encoding and decision making. *Journal of Experimental Psychology: General, 139,* 70–94.

Ratcliff, R., Spieler, D., & McKoon, G. (2000). Explicitly modeling the effects of aging on response time. *Psychonomic Bulletin & Review, 7,* 1–25.

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

Roberts, S., & Sternberg, S. (1993). The meaning of additive reaction-time effects: Tests of three alternatives. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive Neuroscience* (pp. 611–653). Cambridge, MA: MIT Press.

Rouder, J. N., Yue, Y., Speckman, P. L., Pratte, M. S., & Province, J. M. (2010). Gradual growth versus shape invariance in perceptual decision making. *Psychological Review, 117,* 1267–1274.

Royston, P. (1992). Which measures of skewness and kurtosis are best? *Statistics in Medicine, 11,* 333–343.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Sanders, A. F. (1998). *Elements of human performance: Reaction processes and attention in human skill*. Mahwah, NJ: Erlbaum.

Schall, J. D. (2003). Neural correlates of decision processes: Neural and mental chronometry. *Current Opinion in Neurobiology, 13,* 182–186.

Schall, J. D., Purcell, B. A., Heitz, R. P., Logan, G. D., & Palmeri, T. J. (2011). Neural mechanisms of saccade target selection: Gated accumulator model of the visual-motor cascade. *European Journal of Neuroscience, 33,* 1991–2002.

Sigman, M., & Dehaene, S. (2008). Brain mechanisms of serial and parallel processing during dual-task performance. *The Journal of Neuroscience, 28,* 7585–7598.

Sleimen-Malkoun, R., Temprado, J.-J., & Berton, E. (2013). Age-related dedifferentiation of cognitive and motor slowing: Insight from the comparison of Hick-Hyman and Fitts' laws. *Frontiers in Aging Neuroscience, 5,* 62. http://dx.doi.org/10.3389/fnagi.2013.00062

Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.) *An invitation to cognitive science, Vol. 4*: *Methods, models, and conceptual issues* (pp. 703–863). Cambridge, MA: MIT Press.

Sternberg, S. (2001). Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychologica, 106,* 147–246.

Thomas, E. A. C., & Ross, B. H. (1980). On appropriate procedures for combining probability distributions within the same family. *Journal of Mathematical Psychology, 21,* 136–152.

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.

Weisstein, E. W. (2014). *Kurtosis. MathWorld–A Wolfram Web Resource*. Retrieved from http://mathworld.wolfram.com

(*Appendix follows*)

Appendix

The Summation Test

We wish to test the hypothesis that the RT is generated by sequential processes (stages) **A** and **B,** whose durations are stochastically independent, and that factors *f* and *g* with two levels each influence stages **A** and **B** selectively. Equation 1 then follows, as shown by Ashby and Townsend (1980, p. 108).

Usually there are one or more additional factors that might influence both **A** and **B.** Examples are the level of practice, the particular stimulus, or whether the current stimulus repeats the previous one. If we ignore the levels of such factors, then their effects may cause the durations of **A** and **B** to covary, violating stochastic independence. Thus, tests of Equation 1 must be applied to subsets of data within which the levels of such "nuisance factors" are fixed. For many experiments, such subsets may be small. For the detection experiment described in the text, the nuisance factor was "session," and, for each subject, the subsets contained only 16 observations.

Ashby and Townsend (1980, p. 109) proposed testing Equation 1 by estimating the density functions $d_{ij}(t)$ for each of the four RT_{ij} sets, and determining whether the convolutions of $d_{12}(t)$ with $d_{21}(t)$ and $d_{11}(t)$ with $d_{22}(t)$ are equal. It isn't clear whether their method could work with small subsets of data. The summation test (Roberts & Sternberg, 1993) is simpler and more direct, and has been used successfully with small subsets. The basic idea is simply to add the observed RTs for Conditions 11 and 22, and for Conditions 12 and 21, within levels of the nuisance factors, to combine these sums across those levels, and to compare the resulting distributions. No estimation of density functions is required.

Procedure

The procedure, which is fully documented by Roberts and Sternberg (1993, Section 26.8) with examples, is as follows:

1. Partition each subject's data into what are hoped to be homogeneous subsets, that is, within levels of the nuisance factors.

2. For each of the subsets, indexed by *k,* sum the elements of the cartesian products of RT_{11k} with RT_{22k} and of RT_{12k} with RT_{21k}. (The cartesian product of sets S_1 and S_2 of sizes n_1 and n_2 is the set of all $n_1 \times n_2$ possible pairs of their members.) This produces two sets of sums for each data subset, one of which represents $RT_{11k} + RT_{22k}$ (the $S_{11.22.k}$ set) and the other of which represents $RT_{12k} + RT_{21k}$ (the $S_{12.21.k}$ set).

3. Before pooling these sets of sums across subsets, *k,* or comparing the results across subjects, adjust these sets by applying the same linear transformation to the members of each pair, $S_{11.22.k}$ and $S_{12.21.k}$, selecting the transformations for each pair so that the mean of their two medians and the mean of their two interquartile ranges are the same across all pairs, *k.* Application of the same transformation to the members of each pair perserves any differences between them. Call these normalized sets of sums $S^n_{11.22.k}$ and $S^n_{12.21.k}$.

4. Pool each of the normalized sets of sums over levels, *k,* to get $S^n_{11.22}$ and $S^n_{12.21}$ for each subject. Normalization before pooling is based on the belief that systematic failures of the test are more likely to occur at corresponding quantiles of the pair of distributions than at corresponding RTs.

5. The distributions of the sums $S^n_{11.22}$ and $S^n_{12.21}$ can now be compared. Because the prediction is that they should be identical, within sampling error, any measures of these distributions can be used, including ones, such as L-moments or other measures that depend on order statistics, which are less subject to the influence of extreme observations than are the variances or higher moments of the distributions. One possibility is to compute a pair of such distributions for each subject, determine the means over subjects of the measures of interest, and to estimate sampling error from their between-subjects variability.

Application to the Detection Data

This procedure was applied to the data discussed in the text. Let the distributions of $S^n_{12.21}$ and $S^n_{11.22}$ be denoted "Lb*Sd" and "Sb*Ld," respectively, indicating the pairs of conditions for which RTs were summed. Values of four L-statistics of these "summation distributions" are shown in Table A1. The differences are very small, especially when compared to the systematic differences shown in Table 1 among the component distributions. This result confirms the hypothesis of selective influence of foreperiod and intensity on sequential processes (stages) with stochastically independent durations.

Table A1
Summation Test L-Statistics: Means Over Five Subjects

Statistic	λ_1 (*ms*)	λ_2 (*ms*)	τ_3	τ_4
Mean for Lb*Sd	444.65	18.13	.150	.171
Mean for Sb*Ld	444.78	18.01	.160	.177
Difference	.13	−.12	.010	.006
SE of difference	1.23	.96	.039	.025

Note. Distributions Lb*Sd and Sb*Ld were determined for each of the five subjects. From these distributions, L-statistics and their differences were computed for each subject. "*SE*" is the standard error of the mean difference, based on between-subject variation.

(*Appendix continues*)

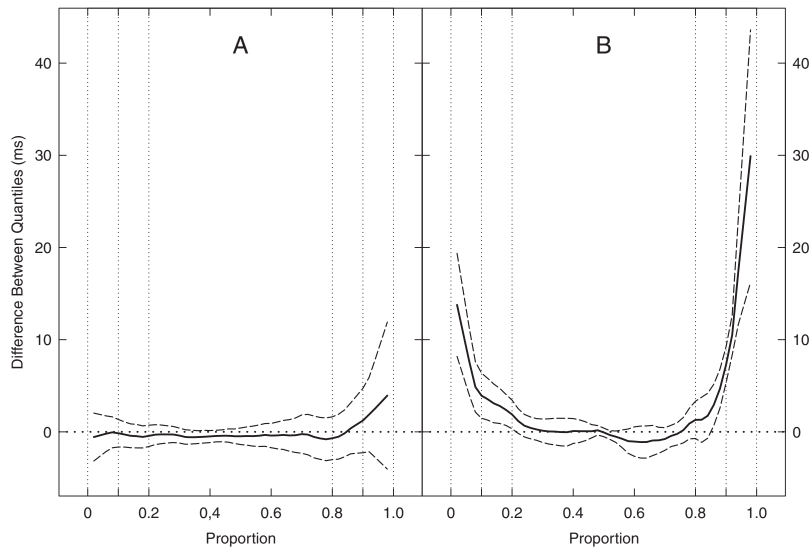

Figure A1. Distribution differences with ± *SE* curves. A. Mean differences between summation distributions Sb*Ld and Lb*Sd. B. Mean differences between component distributions Sb and Lb.

Further information about the relation between the Sb*Ld and Lb*Sd distributions is shown in Figure A1. Panel A shows how close to identical they are. For comparison, Panel B shows the difference between the two most different component distributions (also compared in Figure 2 of the text), plotted in the same way.

Received February 7, 2014
Revision received June 3, 2015
Accepted June 5, 2015 ■