Psychological Review
1982, Vol. 89, No. 5, 599-607

0033-295X/82/8905-0599$00.75

Deriving Exact Predictions From the Cascade Model

F. Gregory Ashby
Ohio State University

McClelland's (1979) cascade model is investigated, and it is shown that the model does not have a well-defined reaction time (RT) distribution function because it always predicts a nonzero probability that a response never occurs. By conditioning on the event that a response does occur, RT density and distribution functions are derived, thus allowing most RT statistics to be computed directly and eliminating the need for computer simulations. Using these results, an investigation of the model revealed that (a) it predicts mean RT additivity in most cases of pure insertion or selective influence; (b) it predicts only a very small increase in standard deviations as mean RT increases; and (c) it does not mimic the distribution of discrete-stage models that have a serial stage with an exponentially distributed duration.

Recently, McClelland (1979) proposed a continuous-time linear systems model of simple cognitive processes based on sequential banks of parallel integrators. This model, referred to by McClelland as the cascade model, exhibits some potentially very interesting properties. For example, McClelland argues that under certain conditions it mimics some of the reaction time (RT) additivities characteristic of serial discrete-stage models. Unfortunately, however, rigorous empirical testing of the model is precluded because McClelland (1979) offers no method for computing any of the RT statistics it predicts.

The format of this note is as follows: I will show that the model always predicts a nonzero probability that a response never occurs, which means, for example, that it always predicts infinite mean RTs. One way to circumvent this problem is to look only at trials on which a reponse does occur. By doing this it is possible to derive an RT probability density function predicted by the cascade model. From it, virtually any desired RT statistic can be accurately computed. Some of these (e.g., means and variances) will be examined, with particular regard to how well they correspond to known empirical results. For example, it turns out that the cascade model has little difficulty predicting standard mean RT additivities of the kind often found, say, in a Sternberg (1966) memory-scanning task but that it has great difficulty predicting the RT variability results. Finally, using these new tools it will be better possible to answer the question of how closely the cascade model comes to mimicking discrete-stage models.

Basically, the cascade model proposes that upon stimulus presentation, an overlapping, deterministic flow of information "cascades" through the processing system, thereby causing activation to build continuously at each of several response units. When the activation at any one unit exceeds a criterion value, the response associated with that unit is emitted. Thus, the cascade model proposes a race between the various response units; the first unit to accrue the criterion level of activation determines the response. Randomness in the observed RTs results because both the starting point and the asymptotic activation level are assumed to be normally distributed random variables.

Because the activation at the different parallel units accrues independently, the behavior of this model is completely determined by the activity at each individual response unit. Once the predictions of each of these are obtained, the predictions of the model in more general stimulus-response situations can be derived in a fairly straightforward fashion. Thus, for now, assume as in cases of perfect accuracy that a given stimulus always results in the same response.

In this case, RT is determined solely by the activity on one response unit. According to McClelland (1979), when there are n stages in the processing chain, the activation at each unit at time t is given by

This article was written while I was visiting Harvard University as a National Science Foundation Postdoctoral Fellow. I am grateful to W. K. Estes for the opportunity this year provided. I also wish to thank James T. Townsend and Joel Angiolillo-Bent for their helpful comments on the manuscript.

Requests for reprints should be sent to F. Gregory Ashby, Human Performance Center, Ohio State University, 404B West 17th Avenue, Columbus, Ohio 43210.

$$X_n(t) = A[1 - \sum_{i=1}^{n} K_i \exp(-k_i t)] + D$$

$$= AE_n(t) + D, \quad (1)$$

where D is a standard normal random variable (i.e., with mean 0 and unit variance), A is the normally distributed asymptotic activation level with mean a and variance σ_A^2 and is indepenent of D, and k_i is the processing rate of the ith of the n stages. Finally when $k_i \neq k_j$, for all i and $j \leq n$, K_i is defined by[1]

$$K_i = \prod_{j \neq i}^{n} k_j/(k_j - k_i).$$

This is the extent of McClelland's (1979) development. Even so, the process of deriving an RT probability distribution for this model is fairly straightforward. First, note that a response occurs when the activation level $X_n(t)$ exceeds a criterion value of R and therefore[2]

$$P(\mathrm{RT}_n \leq t) = P[X_n(t) \geq R].$$

Now $X_n(t)$ is itself normally distributed with mean $aE_n(t)$ and variance $1 + \sigma_A^2[\mathrm{E}_n(t)]^2$, and so after standardizing via a z transformation

$$P(\mathrm{RT}_n \leq t) = P[X_n(t) \geq R]$$

$$= P\{Z \geq [R - aE_n(t)]/\sqrt{1 + \sigma_A^2[E_n(t)]^2}\}$$

$$= P[Z \geq W_n(t)]$$

$$= 1 - (1/\sqrt{2\pi}) \int_{-\infty}^{W_n(t)} \exp(-x^2/2)dx. \quad (2)$$

The first thing to note about this function is that it is not a true probability distribution. Notice that $E_n(t)$ approaches one as t becomes infinite, and therefore, from Equations 1 and 2

$$P(\mathrm{RT} < \infty) = P(A + D \geq R)$$

$$= P[Z \geq (R - a)/\sqrt{1 + \sigma_A^2}] < 1,$$

which will be true for all normally distributed random variables A and D. This means there is a nonzero probability that a response never occurs, and so, as the model stands, no RT density function exists. This problem is minimized but still exists if more than one response is possible because even in these cases there will be trials when none of the activations exceed their criterion level. Note also that the problem becomes worse as the variance σ_A^2 increases. For example, suppose, as McClelland (1979) often does, that the mean asymptotic activation $a = 5$, the variance $\sigma_A^2 = 1$, and the criterion $R = 2.5$. Then

$$P(\mathrm{RT}_n < \infty) = 1 - Z(-1.77) = .9614,$$

where $Z(x)$ is the standard normal integral evaluated at x. Thus, on about 4% of the trials, a response never occurs. If $\sigma_A^2 = 2$ instead, then $P(\mathrm{RT}_n < \infty) = .9251$ and no response is given on 7.5% of the trials. This is clearly an incorrect prediction, especially for simple RT tasks in which no incorrect responses are possible.

In a similar fashion, the problem worsens as the difference between a and R decreases. For instance, lowering the mean asymptotic activation level (i.e., a) relative to the criterion R increases the probability that the activation never exceeds the criterion, and so it also increases the likelihood that a response never occurs. For this reason I will not investigate the effects of lowering mean asymptotic activation in this article.

RT density and distribution functions can be defined if we condition on the event that a response does eventually occur. Thus, define

$$F_n(t) = P(\mathrm{RT}_n \leq t | \mathrm{RT}_n < \infty)$$

$$= P(\mathrm{RT}_n \leq t)/\{1 - Z[(R - a)/\sqrt{1 - \sigma_A^2}]\}, \quad (3)$$

where $P(\mathrm{RT}_n \leq t)$ is given in Equation 2. $F_n(t)$ describes the distribution of RTs predicted by the cascade model in any experiment in which accuracy is high. An empirical distribution function can be estimated in an experimental setting by computing the proportion of RTs less than or equal to t for different values of t. The cascade model could then be tested by comparing the predicted $F_n(t)$ with the empirically obtained estimate.

The RT density function is found by differentiating $F_n(t)$ with respect to t. Thus,

---

[1] If some of the k_i are equal, $E_n(t)$ must be written differently (e.g., McGill & Gibbon, 1965), but fortunately, the function behaves the same in either case, so we need not be concerned with this possibility.

[2] Actually, McClelland proposes that RT is equal to this time plus the time it takes to excute a response (about 100 msec). Thus, he sees the response-execution stage as being discrete and not overlapping with the cascade process. This stage will not be included in any of the investigations reported in this article. McClelland assumes it to be invariant over the types of experimental manipulations I consider below, and so the effect of including it would be just to add a constant to all the reported means and variances. None of the qualitative conclusions that are drawn would be affected in any way by its inclusion.

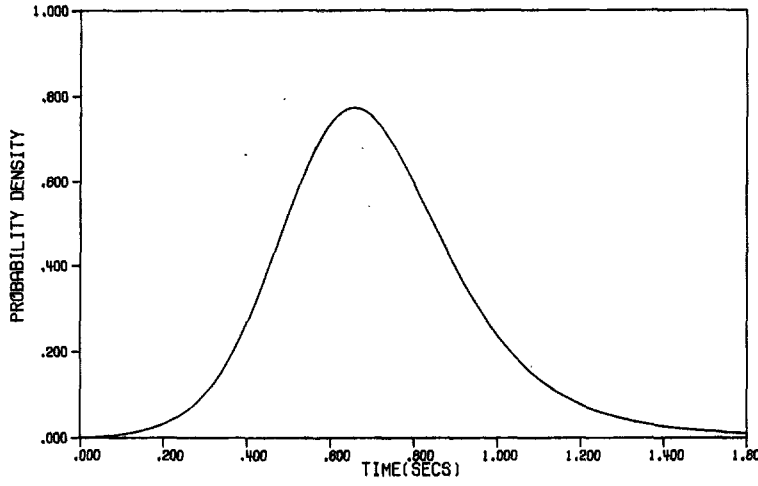

Figure 1. An example of an RT density function from McClelland's (1979) cascade model. (In this particular example there are three stages with rates 3, 4, and 5, respectively. The asymptotic activation has mean 5 and variance 0 and the criterion level of activation is 2.5.)

$$f_n(t) = (d/dt)F_n(t)$$

$$= \frac{[\sum_{i=1}^{n} k_i K_i \exp(-k_i t)][a + R\sigma_A^2 E_n(t)]\exp\{-[W_n(t)]^2/2\}}{\sqrt{2\pi}\{1 - Z[(R-a)/\sqrt{1+\sigma_A^2}]\}\{1 + \sigma_A^2[E_n(t)]^2\}^{3/2}}. \quad (4)$$

An example of this rather complex looking RT density function is illustrated in Figure 1 for the case in which the processing chain has three stages with rates 3, 4, and 5, respectively, the asymptotic activation has mean 5 and variance zero,[3] and the criterion activation level, R, is 2.5. These and all parameter values selected in this article are representative of the values chosen by McClelland (1979). The specific effect of the individual parameters will be discussed below (see also McClelland, 1979). The curve, although not as skewed as one might expect an RT density to be, is not unreasonable as a model of observed RTs.

Given the RT density and distribution functions, virtually any desired RT statistic can now be computed. For instance, in the case of the mean,[4]

$$E(\mathrm{RT}_n) = \int_0^\infty t f_n(t)dt.$$

Although the rather complicated density function in Equation 4 makes the problem of analytically evaluating this integral intractable, it is a fairly easy matter to evaluate it numerically. Of course, no numerical integration can be carried out to infinity, and so this approach only yields approximate results. Even so, any desired degree of accuracy is possible because the upper limit of integration can be chosen to be arbitrarily large. In the applications reported below, this limit was chosen so that more than 99% of $f_n(t)$ was always included in the integrations.

It is especially important that we have a method to compute mean RT accurately because it cannot be estimated from the mean activation versus processing-time curves utilized by McClelland (1979, e.g., Figure 10). For instance, the obvious estimate is the time at which the mean activation level first exceeds the criterion. However, because mean activation is negatively accelerated as it approaches asymptote, this procedure will tend to underestimate significantly the true mean.

[3] Although the RT variance is greater when $\sigma_A^2 = 1$, the shape of the density is qualitatively very similar to the case considered here.

[4] A somewhat more useful formula may be

$$E(\mathrm{RT}_n) = \int_0^\infty [1 - F_n(t)]dt$$

$$= \int_0^\infty Z[W_n(t)]dt,$$

where $Z[W_n(t)]$ is the area under the standard normal density to the left of $W_n(t)$.

The functions $F_n(t)$ and $f_n(t)$ of Equations 4 and 5 can also be used to find the RT density in more general experimental circumstances. For instance, consider the case of two stimuli, or two types of stimuli, S_a and S_b, and two responses, R_a and R_b. Let $g_{ij}(t)$ for i and j equal to a or b be the RT density function for the response unit associated with Response R_j when it is the only unit active and when Stimulus S_i is presented. Thus $g_{ij}(t)$ will be in the same form as $f_n(t)$ of Equation 4, except that the parameters may depend on the stimulus S_i as well as on the response unit R_j. For instance, the mean asymptotic activation level for unit R_j should usually be higher when Stimulus S_j is presented than when Stimulus S_i is. Finally define $q_{ij}(t)$ as the observable RT density in this experiment on trials when Stimulus S_i is presented and Response R_j is made.

Under these circumstances, the observable RT is determined by whichever of the two units is the first to achieve the criterion level of activation, and so when more than one response is possible, the cascade model is just a special kind of parallel process. Parallel-processing models have been extensively studied (e.g., see Townsend & Ashby, in press), and if the two units associated with Responses R_i and R_j operate independently, then it can be shown (e.g., Townsend & Ashby, in press) that if $i \neq j$, then

$$q_{ij}(t) = \frac{g_{ij}(t)[1 - G_{ii}(t)]}{\int_0^\infty g_{ij}(t)[1 - G_{ii}(t)]dt} \tag{5}$$

and

$$q_{ii}(t) = \frac{g_{ii}(t)[1 - G_{ij}(t)]}{\int_0^\infty g_{ii}(t)[1 - G_{ij}(t)]dt}. \tag{6}$$

Thus, given the distribution and density functions in Equations 3 and 4, it is possible to derive (numerically) the RT density function predicted by the cascade model in every condition of a great many experimental designs. In addition, response probabilities are also easily obtained because these are the denominators of Equations 5 and 6. Thus,

$$P(R_j|S_i) = \int_0^\infty g_{ij}(t)[1 - G_{ii}(t)]dt$$

and

$$P(R_i|S_i) = \int_0^\infty g_{ii}(t)[1 - G_{ij}(t)]dt. \tag{7}$$

Pure Insertion and the Cascade Model

This section briefly considers the effects on RT of inserting some new stage into the cascade process. To begin, one might ask whether the increase in RT caused by inserting an extra stage is independent of the rates of the unaffected stages. This type of independence, known as additivity, is characteristic of serial discrete-stage models. Table 1 illustrates the effects on the mean and standard deviation of RT of adding a fourth stage to the cascade process when the rate of the second stage is changed from 6.0 in Condition A to a value of 5.0 in Condition B. Mean RT additivity results if

$$E(RT_{A3}) + E(RT_{B4}) = E(RT_{B3}) + E(RT_{A4}),$$

where the second subscript refers to the number of stages in the processing chain and, as before, E denotes expectation.

It should be emphasized that all cascade-model predictions reported here, including those in Table 1, are based on the analytic results of Equations 2, 3, and 4 and are not the result of computer simulations. Table 1 indicates that in all cases,

Table 1
RT Means (msec) and Standard Deviations Predicted by the Cascade Model When a Stage With Rate k_4 is Inserted Into the Processing Chain

Stage rate		$E(RT_{A3})$		$E(RT_{A4})$		$E(RT_{B3})$		$E(RT_{B4})$		
k_3	k_4	M	SD	M	SD	M	SD	M	SD	$[E(RT_{A3}) + E(RT_{B4})] - [E(RT_{B3}) + E(RT_{A4})]$
19	20	502	236	553	238	535	247	586	249	0
14	15	522	238	590	242	554	249	622	252	0
9	10	563	244	665	252	595	255	697	262	0
4.5	5.5	673	275	857	299	706	285	890	307	0
1	2	1330	622	1823	683	1365	625	1856	686	−2
1	4	1330	622	1582	635	1365	625	1616	638	−1
20	4	499	236	749	285	532	247	782	293	0
20	10	499	236	602	244	532	247	634	254	−1
20	15	499	236	568	240	532	247	600	250	−1
20	25	499	236	540	237	532	247	573	248	0

Note. In all cases $a = 5.0$, $\sigma^2_A = 1$, $R = 2.5$, and the first stage rate $k_1 = 3.0$. The second stage rate k_2 equals 6.0 in Condition A, and 5.0 in Condition B.

mean RT additivity is closely approximated. McClelland suggested as a (conservative) rule of thumb that additivity should result so long as the rate of the inserted stage is at least four times as great as the rate of the slowest stage in the system. In the Table 1 examples there are four instances when the inserted-stage rate is less than four times greater than the rate of the slowest stage, yet in each of these additivity is virtually perfect. Furthermore, the additivity apparently does not depend on the magnitude of the variance parameter, σ_A^2 (1.0 in Table 1). All calculations in Table 1 were repeated with $\sigma_A^2 = 0$, and although this tended to decrease the means slightly, additivity was again virtually exact in all cases.

Although there may be some combination of parameters that leads to violations of additivity large enough to be empirically observable, Table 1 indicates that the conditions under which the cascade model predicts mean RT additivity do not depend simply on the relative magnitude of the inserted-stage rate and the rate of the slowest stage in the system.

Mean RT additivities are commonly reported in many experimental paradigms. For instance, in standard memory-scanning tasks, mean RT is usually found to increase with the size of the memory set at the same rate on target-absent (i.e., "no") trials as it does on target-present ("yes") trials (e.g., see Sternberg, 1975). Thus, it is likely that the cascade model can closely approximate these results. On the other hand, Table 1 makes it equally clear that the model will have a difficult time accounting for standard RT variability results, at least when it postulates that RT increases because an extra stage was inserted into the cascade process.

RT variability is almost always found to increase sharply with the mean (e.g., Chocholle, 1940; Green & Luce, 1973; Schneider & Shiffrin, 1977; Sternberg, Note 1), and although the very large standard deviations reported in Table 1 do tend to increase with mean RT, they increase much too slowly to account for the bulk of the empirical literature. For instance, in a memory-scanning task, Sternberg (Note 1) found that adding an item to the memory set increased the RT standard deviation by about 10 msec, and the mean by 35 msec. Similarly, Green and Luce (1973) found the standard deviation to increase 1 msec for every 3-msec increase in mean RT in a signal-detection task. In contrast, all standard deviations in Table 1 increase less than 1 msec for every 5-msec mean RT increase, and the average is only about 1 msec to every 18-msec increase in the means.

With σ_A^2 set to zero, the standard deviations tend to be about 15% smaller (although they are still very large compared to the mean), but they increase with mean RT increases at about the same rate (1 msec per 16-msec mean increase). Thus, manipulating this parameter does not appear to solve the model's variability problems. Another way the standard deviations can be reduced is to increase the parameters a and R. This will tend to minimize the effect of the starting-point variability introduced by the random variable D in Equation 1. For instance, increasing a to 10 and R to 5 reduces the standard deviations in Table 1 by more than 50% (mean RT additivity is unaffected), but at the same time it causes the standard deviations to increase with the mean at an even slower rate, and so this strategy also fails. Reducing mean asymptotic activation without changing the criterion will also affect predicted RT variability, but as mentioned earlier, this strategy significantly increases the probability that a response never occurs and for this reason is an unappealing alternative.

These results seem to indicate that empirical RT variability results will prove troublesome to the cascade model. The problem seems to be in the post hoc way that RT varibility is introduced into the model. The activation, $E_n(t)$ of Equation 1, grows in an essentially deterministic fashion. It has no inherent variability. Thus, inserting a stage into the processing chain has no direct effect on the variability of the cascade process.

One way around this problem might be to let the cascade rates be random variables and so vary from trial to trial. In this case inserting a stage into the cascade process should cause an increase in RT variability because it adds another source of variation to the RT process. The price that is paid though is greatly increased analytic complexity. Certainly, the RT distribution derived in Equation 2 would no longer be appropriate. In fact, it seems a good guess that the problem of deriving the correct RT distribution for such a model would be intractable, and so there would be no way to derive any exact RT predictions. Some idea of how the model would perform could be obtained from intuition, and some from computer simulation, but both methods are inexact, and in addition, the former is often faulty whereas the latter is inefficient.

Now that we have some idea about the ability of the cascade model to mimic the mean RT results predicted by discrete-stage models when pure insertion is assumed, we might ask if it can also mimic discrete-stage models at the distributional level; that is, does the RT distribution predicted by the cascade model resemble the RT distribution predicted by discrete-stage models? On first encounter with the cascade model, one is struck by the exponential terms in Equation 1. In fact, the activation level $E_n(t)$ of Equation 1 has the same form as the cumulative distribution function

of a general gamma distribution (e.g., McGill & Gibbon, 1965). The general gamma is the RT distribution of a serial discrete-stage model in which the duration of each stage is exponentially distributed. Even though the general gamma function describes a deterministic activation level in the cascade model rather than a probability distribution, it is still of interest to ask whether the cascade model mimics the distribution of discrete-stage models in which the duration of one or more stages is exponentially distributed. Fortunately, this question is easy to answer.

Suppose, for $n > 1$ that

$$RT_n = RT_{n-1} + T_n, \tag{8}$$

where RT_{n-1} and T_n are independent random variables and T_n is exponentially distributed with rate k_n. Thus, a discrete stage with an exponentially distributed duration has been inserted into the processing chain. Under these conditions, I (Ashby, 1982) showed that the density functions of RT_n and RT_{n-1}, that is, $f_n(t)$ and $f_{n-1}(t)$, must intersect at the mode of $f_n(t)$. In addition, Ashby and Townsend (1980) demonstrated that if (and only if) Equation (8) is valid, the following identity must hold for all $t > 0$:

$$k_n = f_n(t)/[F_{n-1}(t) - F_n(t)]. \tag{9}$$

Because the left-hand side is a constant, this identity provides a test of models satisfying Equation 8, because if Equation 8 is true, a plot of the right-hand side of Equation 9 versus time will yield a constant (flat) function. Data from a standard memory-scanning experiment (reported by Townsend & Roos, 1973) supports both of these predictions (Ashby, 1982; Ashby & Townsend, 1980), and thus, it is especially relevant to check whether the cascade model mimics them.

A number of cascade models were examined with regard to these predictions. In all of these the mean asymptotic activation was set to $a = 5$, the criterion was set to $R = 2.5$, and the first three stages had rates 3, 4, and 5, respectively. The inserted-stage rate varied from $k_4 = 1$ to $k_4 = 20$, and the variance parameter was either $\sigma_A^2 = 1$ or $\sigma_A^2 = 0$. In all cases values of k_4 between 1 and 20 led to predictions intermediate to the $k_4 = 1$ and the $k_4 = 20$ predictions, and thus, only the latter of these are reported here. In addition, the same qualitative results were obtained with $\sigma_A^2 = 0$ and $\sigma_A^2 = 1$. Figures 2 and 3 arbitrarily assume $\sigma_A^2 = 0$.

Figure 2 contains several examples of the cascade model densities, $f_3(t)$ and $f_4(t)$, and it is clearly seen that unlike the discrete-stage models of Equation 8, the cascade model does not predict that the two densities will intersect at the mode of $f_4(t)$. In all cases the intersection occurs before this mode, although it is closer to the mode for the larger value of k_4. This may be an artifact, however, because the greater the rate of the inserted stage, the closer the means of $f_3(t)$ and $f_4(t)$.

Figure 2 indicates that the cascade model does not mimic the discrete-stage model of Equation 8, at least when a is 5.0 and R is 2.5. Thus, a plot of $f_4(t)/[F_3(t) - F_4(t)]$ versus t should not yield a flat function for these parameter values. Even so, the cascade model may come closer to mimicking the Equation 9 result when the inserted-stage rate is large. Figure 3 shows plots of $f_4(t)/[F_3(t) - F_4(t)]$ versus time for the same parameter values of Figure 2. As expected, neither of these plots are flat, but surprisingly perhaps, the deviations from a flat function actually become more severe as the value of the inserted-stage rate increases. It appears that even though the activation level in the cascade process is composed of the convolution of exponential functions, the RT distribution predicted by the model is fundamentally different from the distribution of discrete-stage models containing a serial stage with an exponentially distributed duration.

Selective Influence and the Cascade Model

Next, I shall briefly consider the ability of the cascade model to mimic the additivities predicted by discrete-stage models when the duration of two (or more) stages are selectively influenced by some experimental factors. McClelland (1979) maintains that the cascade model mimics the additivity displayed by discrete-stage models under the assumption of selective influence (at, presumably, the mean level) almost perfectly "If the rate constant of either of the two manipulated processes is at least four times the rate of the slowest process in the system at all levels of both factors. Even outside this range, additivity is surprisingly good" (p. 311).

To test this claim, the rates of two cascade-model stages were factorially manipulated over a fairly wide range of values. It was assumed throughout that Factor A affects only the rate of the third stage whereas Factor B affects only the fourth stage rate. The two "base" stages had rates $k_1 = 3$ and $k_2 = 4$, whereas, as before, the mean asymptotic activation level, a, was set to 5 and the criterion activation level, R, to 2.5. Because the variance parameter, σ_A^2, seemed to have no effect on any qualitative predictions (i.e., mean RT additivity, the intersection points in Figure 2, the shape of the curves in Figure 3, etc.), it was set to zero for all calculations reported in this section.

Under these circumstances mean RT additivity holds if

$$E(RT_{A1,B1}) + E(RT_{A2,B2}) = E(RT_{A1,B2}) + E(RT_{A2,B1}).$$

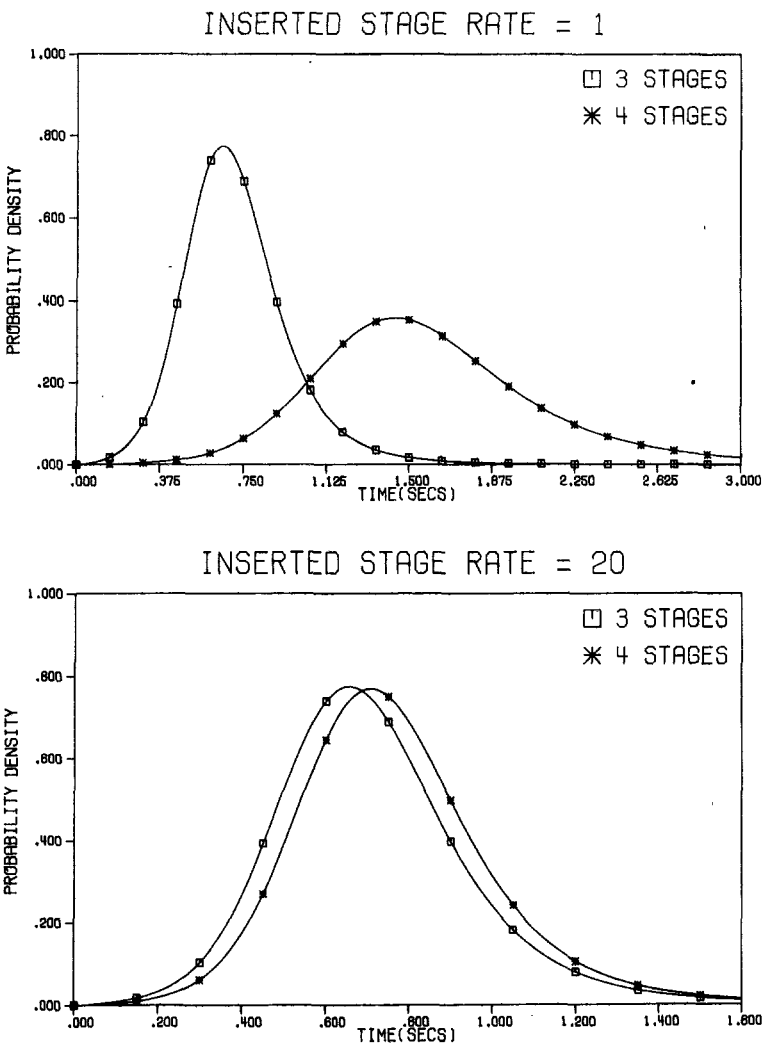

Figure 2. RT densities predicted by the cascade model. (In all cases the asymptotic activation has mean 5 and variance 0, the criterion level of activation is $R = 2.5$, and three stages have rates 3, 4, and 5, respectively. The rate of the fourth [inserted] stage varies from 1 to 20.)

Table 2 details the results of the factorial manipulations. It can be seen that a few substantial deviations are predicted, and these tend to correspond with McClelland's rule of thumb, but in most cases additivity is virtually exact. Even so, because of the large standard deviations predicted by the cascade model, only a remarkable empirical effort would succeed in detecting these nonadditivities. For instance, to detect a true deviation from additivity of 41 msec (as in Table 2) at a 95% confidence level requires 1,983 trials[5] per experimental condition if the cascade model is correct. Few experimenters have the time or resources for such an expenditure.

With regard to additivity at the distributional level, Ashby and Townsend (1980) showed that the assumption of selective influence implies that for all $t > 0$,

$$F_{A1,B1}(t) * f_{A2,B2}(t) = F_{A1,B2}(t) * f_{A2,B1}(t),$$

where $*$ stands for the convolution operation. The expression on each side of the equality is a cumulative distribution function, and thus, this condition can be checked by a statistical test such as the Kolmogorov-Smirnov. This nonparametric test, found to be more powerful than the better known chi-square test (e.g., Massey, 1951), uses the maximum deviation of the two distribution functions as its test statistic.

[5] This number will be even greater if $\sigma_A^2 > 0$.

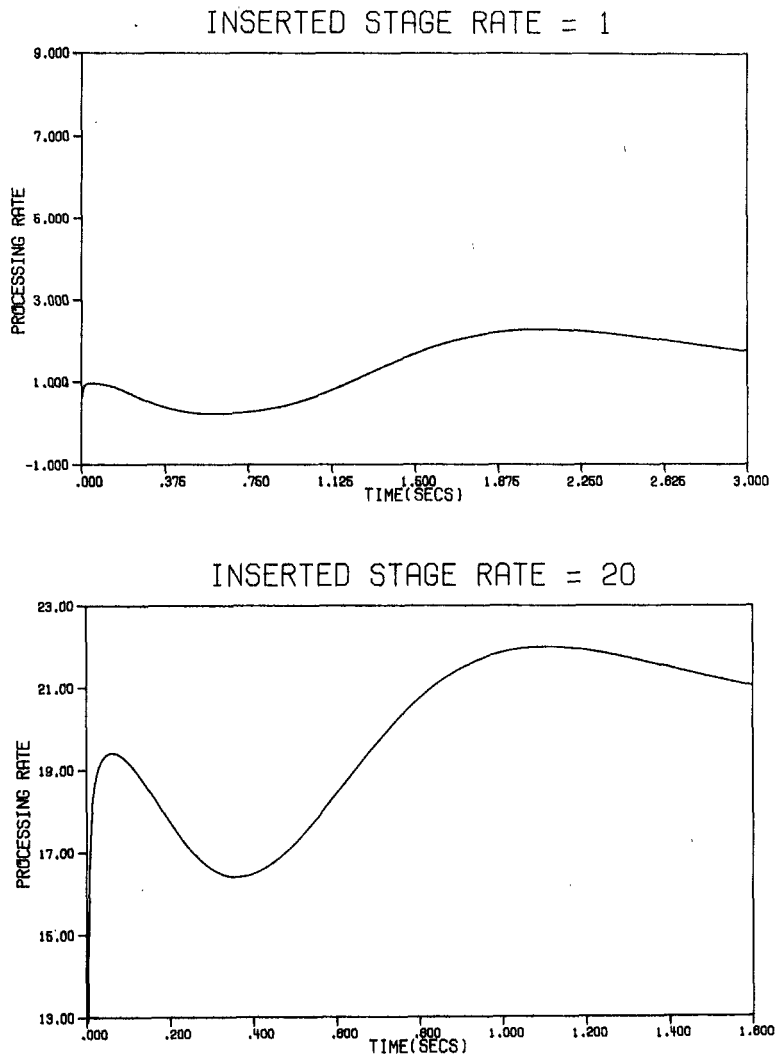

Figure 3. Plots of $f_4(t)/[F_3(t) - F_4(t)]$ versus t for cascade-model densities. (In all cases $a = 5$, $\sigma_A^2 = 0$, $R = 2.5$, $k_1 = 3$, $k_2 = 4$, and $k_3 = 5$. The rate of the inserted stage, k_4, varies from 1 to 20.)

Table 2 indicates that these maximum deviations agree ordinally with the mean additivity results, and in addition, it turns out that about the same number of trials are required to empirically detect the largest of these. Thus, the same conclusions can be drawn for the mean and the distributional results. If the rates of different processing stages in the cascade mode are selectively influenced, empirical additivity is to be expected.

Conclusions

The cascade model, as proposed by McClelland (1979), contains a fundamental flaw. It predicts that on a certain (sometimes quite significant) percentage of trials, subjects will fail to emit a response. Clearly, the model must be amended in some fashion to correct this problem. No solution is suggested here. Instead, predictions of the model are derived for trials on which a response does occur.

The derivation of RT density and distribution functions for these trials means that the cascade model can now be fit directly to virtually any data having RT and accuracy as dependent variables. Simulations are no longer necessary. Thus, quantitative fits of the model can now be compared with competitive theories so that the validity of the cascade idea can be rigorously tested.

This is very easily done when the RT statistics of interest are empirical density and distribution functions or some function of them—as, for ex-

Table 2
Results of Factorially Manipulating the Third and Fourth Rate Parameters in the Cascade Model

k_3		k_4		$[E(RT_{A1,B1}) + E(RT_{A2,B2})] - [E(RT_{A1,B2}) + E(RT_{A2,B1})]$	Maximum deviation of $[F_{A1,B1}(t) * f_{A2,B2}(t)] - [F_{A1,B2}(t) * f_{A2,B1}(t)]$
A_1	A_2	B_1	B_2		
5	6	1	2	0	.0008300
5	6	1	21	2	.0018586
5	6	7	8	0	.0000391
5	6	20	21	0	.0000047
1	20	2	19	41	.0285625
20	21	7	8	0	.0000041
20	21	19	22	0	.0000022
20	21	1	2	0	.0000061
20	21	1	22	0	.0000291
1	5	2	6	25	.0245825
1	5	2	20	38	.0267735
1	20	9	10	0	.0001900
1	5	9	10	0	.0002191
1	20	2	6	27	.0267118

Note. In all cases $k_1 = 3$, $k_2 = 4$, $a = 5$, $\sigma_A^2 = 0$, and $R = 2.5$.

ample, the hazard function, $h(t) = f(t)/[1 - F(t)]$). It is also easily done with order statistics, such as the median, and with the mode. More care must be taken when fitting means, variances, and other moments, however, because these depend so heavily on the tail of the density function. The numerical integrations necessary to calculate the moments must be carried out far enough into the tail to ensure that the approximations are accurate.

The rather limited set of predictions investigated here reveal that the cascade model (a) predicts mean RT additivity (or near additivity) in almost all cases in which a stage is inserted into the cascade process or in which a stage rate is selectively influenced, (b) predicts only a small increase in RT variability as the mean increases, (c) predicts very large RT standard deviations over much of its parameter space, and (d) does not behave at the distributional level like a discrete-stage model that has a serial stage whose duration is exponentially distributed. Of course, these conclusions must be tentative because a complete parametric study of the model was not made.

The primary aim of this note was to provide a method of obtaining exact quantitative predictions from McClelland's (1979) cascade model. Once this method is available, the model can be rigorously tested, and only then can its true validity be determined.

Reference Note

1. Sternberg, S. *Estimating the distribution of additive reaction-time components.* Paper presented at the meeting of the Psychometric Society, Niagara Falls, Ontario, Canada, 1964.

References

Ashby, F. G. Testing the assumptions of exponential, additive reaction time models. *Memory & Cognition,* 1982, *10,* 125–134.

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

Chocholle, R. Variation des temps de réaction auditifs en fonction de l'intensité a diverses fréquences. *L'Annee Psychologique,* 1940, *41,* 65–124.

Green, D. M., & Luce, R. D. Speed–accuracy trade-off in auditory detection. In S. Kornblum (Ed.), *Attention and performance IV.* N.Y.: Academic Press, 1973.

McClelland, J. L. On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review,* 1979, *86,* 287–330.

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

Massey, F. J., Jr. The Kolmogorov-Smirnov test for goodness of fit. *American Statistical Association Journal,* 1951, *46,* 68–78.

Schneider, W., & Shiffrin, R. M. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review,* 1977, *84,* 1–66.

Sternberg, S. High-speed scanning in human memory. *Science,* 1966, *153,* 652–654.

Sternberg, S. Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology,* 1975, *27,* 1–32.

Townsend, J. T., & Ashby, F. G. *The stochastic modeling of elementary psychological processes.* New York: Cambridge University Press, in press.

Townsend, J. T., & Roos, R. N. Search reaction time for single targets in multiletter stimuli with brief visual displays. *Memory & Cognition,* 1973, *1,* 319–332.

Received January 18, 1982
Revision received May 14, 1982 ■