

The Weibull Distribution, the Power Law, and the Instance Theory of Automaticity

Gordon D. Logan
University of Illinois

H. Colonius (1995) agreed with the fundamental tenets of the instance theory of automaticity. His article addressed the mathematical development of the theory, pointing out an error in one of two arguments that G. D. Logan (1988, 1992) used to justify the choice of the Weibull as the distribution of retrieval times and suggesting an alternative argument that places different emphasis on the power function speedup and the Weibull distribution. This article attempts to clarify the problematic argument, point out some practical limitations on H. Colonius's (1995) alternative argument, and suggest important future directions for the mathematical development of the theory.

For several years, I have been developing the instance theory of automaticity to address a number of phenomena in the literature on skill acquisition and automaticity (Logan, 1988, 1990, 1992). The theory accounts for differences between automatic and nonautomatic processing, the conditions required to produce automaticity, the role of attention in the acquisition and expression of automaticity, and the power function speedup that characterizes skill acquisition (see Compton & Logan, 1991; Lassaline & Logan, 1993; Logan, 1988, 1990, 1992; Logan & Etherton, 1994; Logan & Klapp, 1991). The most important contribution of the theory was to shift the focus of research from resource-based conceptions of automaticity (e.g., LaBerge & Samuels, 1974; Posner & Snyder, 1975; Shiffrin & Schneider, 1977) to memory-based conceptions, arguing that automaticity was a memory phenomenon governed by the theoretical and empirical principles that govern memory (see also Anderson, 1982, 1992; Schneider, 1985).

Colonius (1995) agreed with the substance of the theory, arguing that the data and the statistics of extreme values support the idea that the power function speedup results from a race between traces retrieved from memory. His article focused on a part of the logic I used to derive the power function prediction, addressing my assumptions about the Weibull distribution. He pointed out a subtle but important error in one of the arguments I used to justify the assumption that reaction times in well-practiced tasks conform to the Weibull distribution, and he suggested a different role for the Weibull distribution in testing the theory. In this article, I wish to clarify and expand on Colonius's argument, point out limitations on the test he proposed, and suggest new directions for the mathematical development of the instance theory.

This research was supported by Grant SBR 9410406 from the National Science Foundation. I would like to thank Brian Ross and Ehtibar Dzhafarov for their help in preparing this article and Hans Colonius and John Wixted for helpful comments on the manuscript.

Correspondence concerning this article may be addressed to Gordon D. Logan, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign Illinois 61820. Electronic mail may be sent via Internet to GLOGAN@S.PSYCH.UIUC.EDU.

The Instance Theory and the Power Function Speedup

The instance theory consists of three main assumptions: *obligatory encoding*, which says that attention to a stimulus or event is sufficient to cause it to be stored in memory; *obligatory retrieval*, which says that attention to a stimulus or event is sufficient to cause things associated with it to be retrieved from memory; and *instance representation*, which says that stimuli or events are stored and retrieved separately (Logan, 1988, p. 493). These assumptions imply a learning mechanism: When people perform the same task in the same environment repeatedly, obligatory encoding will result in the acquisition of a task-relevant knowledge base, which obligatory retrieval will make available to them when stimuli or events are repeated. This is the essence of the theory: Performance is automatic when it can be supported by traces retrieved from memory instead of costly algorithmic computation (Logan, 1988, 1990). Differences between algorithmic computation and memory retrieval account for the differences between nonautomatic and automatic performance. Consistency of the task and the environment (i.e., consistent mapping; Shiffrin & Schneider, 1977) is important because repetition allows the knowledge base to grow and it allows it to be useful in supporting current performance. Colonius (1995) had no quarrel with this part of the theory.

A major goal in developing the theory was to account for the power function speedup. Newell and Rosenbloom (1981) argued that a power function speedup was an empirical law that governed all situations in which skill was characterized as speed. Since their seminal chapter, the power function speedup has become a benchmark prediction that theories have to make to be taken seriously as accounts of skill acquisition (e.g., Anderson, 1982, 1992; Cohen, Dunbar, & McClelland, 1990; MacKay, 1982; Schneider, 1985). The obligatory retrieval assumption and the instance representation assumption suggested a race between traces as a possible mechanism to account for the power law. Intuitively, the more traces there are in memory, the more likely it is that one will be retrieved exceptionally quickly, which accounts for the speedup with practice. However, the faster a trace is retrieved, the less likely it is that another trace will be retrieved even faster; so increasing the number of traces will have diminishing returns. This explanation accounts for the

negative acceleration that is characteristic of the power function. The remaining step was to formalize these intuitions, and I looked to the statistics of extreme values—maxima and minima—to characterize the race mathematically.

Formalizing the theory required three further assumptions: Each trace was retrieved independently, each trace had the same distribution of retrieval times, and the first trace to be retrieved governed performance. The problem then amounted to figuring out how the time required for the first trace to be retrieved changed as a function of the number of traces in memory—the number of runners in the race. This is a well-studied problem in the statistics of extreme values, and I found two kinds of answers in the statistics of extreme values that suggested that not only mean reaction time, but the whole distribution of reaction times, should decrease as a power function of practice.

One answer involved starting with an initial distribution and working forward by increasing n , the number of traces. The expression for the cumulative distribution function for minima sampled from the same parent distribution, $F(x)$, is

$$F_{1:n}(x) = 1 - [1 - F(x)]^n. \quad (1)$$

I tried the exponential distribution and the Weibull distribution and found that both predicted a power function decrease in reaction time over practice (see Logan, 1988, Appendix A). The cumulative distribution function for the Weibull is

$$F(x) = 1 - \exp(-x^\alpha). \quad (2)$$

The cumulative distribution function of minima sampled from the Weibull distribution can be obtained by substituting Equation 2 into Equation 1:

$$F_{1:n}(x) = 1 - \exp[-(n^{1/\alpha}x)^\alpha]. \quad (3)$$

The other answer involved working backward from the asymptotic distribution of minima. There are three distributions of minima that are asymptotic in the sense that minima sampled from any parent distribution come to be distributed as one of the three as sample size approaches infinity if the minima converge on any distribution at all (e.g., see Castillo, 1988). The asymptotic distribution that was relevant to a particular parent distribution depended on the lower tail of the parent. Distributions that are bounded at the lower tail and have a lower tail with a particular shape (Gnedenko, 1943) converge on the Weibull distribution. I argued that the reaction time distributions contributing to the race should be of this type, and I claimed consequently that reaction time distributions at high levels of practice must conform to the Weibull.

The Asymptotic Argument

Distributions of Minima Degenerate as n Approaches Infinity

Colonius (1995) pointed out, correctly, that my claim that distributions of minima sampled from reaction time distributions converged on the Weibull was erroneous. Working forward from an initial distribution does not result in a Weibull distribution of minima, as I claimed. Colonius, citing Castillo (1988), noted that the distribution of minima sampled from

any parent distribution, including the Weibull expressed in Equation 3, become degenerate as sample size approaches infinity. The value of $F_{1:n}(x)$ must equal 0 in regions where the parent distribution, $F(x)$, equals 0 and it must equal 1.0 in regions where the parent distribution is greater than 0 and less than or equal to 1.0. Thus, minima sampled from a Weibull distribution do not converge on a Weibull distribution as n approaches infinity, contrary to my claim.

The tendency for distributions of minima to become degenerate as sample size approaches infinity means that one cannot work forward from Equations 1 and 3 to asymptotic distributions of minima. However, it does not mean that one cannot predict a power function speedup with Equations 1 and 3 using nonasymptotic arguments. Essentially, this is what Colonius (1995) and I both do (see below). The degeneracy of distributions of minima as sample size approaches infinity is important in principle but it may not be very important in practice. Consider, for example, a parent retrieval time distribution with a mean of 800 ms and a standard deviation of 300 ms, numbers that are typical of real data in experiments to which instance theory has been applied (e.g., see Logan, 1992). If the distribution were Weibull with an exponent of 2.0 (which is also typical), the 1st percentile would be 300 ms and the 99th percentile would be 1,630 ms, a difference of 1,330 ms. With 1,000,000 samples, the nondegenerate range of the distribution of minima (i.e., $0 < F_{1:n}(x) < 1.0$) would have reduced considerably. The difference between the 1st and 99th percentiles would be reduced by a factor of $n^{-1/\alpha} = 1/1,000$ to 1.3 ms. Let us assume, for the sake of argument, that when n equals 1,000,000, the distribution is approximately degenerate. The important question is, How likely is this to occur in practice?

The instance theory makes an important distinction between practice on the task and practice on particular instances. The power function predictions address practice on particular instances rather than practice on the task itself. Most tasks studied in the automaticity literature involve many different stimuli, and instance theory says that it is repetitions of the individual stimuli that count, not repetitions of the task itself (e.g., see Logan & Klapp, 1991). In real-world skills (i.e., the domains to which theories of automaticity should generalize), there are thousands of instances to be learned. Thus, it will take several orders of magnitude more than 1,000,000 trials to present a single instance 1,000,000 times. Moreover, this analysis assumes that people learn every instance and forget none of them, which is an unrealistic assumption (Anderson & Milson, 1989; Grant & Logan, 1993). In practice, the number of retrieved instances may be much smaller than the number of presented ones, which further mitigates the degeneracy of the distribution of minima.

Asymptotic Distributions Are Distributions of Transformed Scores

In 1988 and 1992, I justified my choice of the Weibull distribution with an asymptotic argument as well as a nonasymptotic argument. I claimed that the distribution of minima sampled from any parent distribution whose form was a reasonable description of empirical distributions of reaction times would converge on the Weibull because the Weibull was the relevant

one of the three asymptotic distributions of minima. As Colonius (1995) pointed out, this claim was mistaken because it missed a subtle but extremely important point in the mathematical development of asymptotic distributions: The reason why asymptotic distributions do not become degenerate as sample size approaches infinity is that they are distributions of a linear transformation of the original variable and not distributions of the original variable itself. This point is extremely important. There is nothing in the instance theory to justify transforming the retrieval times, and that makes arguments about asymptotic distributions irrelevant. Moreover, it makes arguments about the domains of attraction for the different asymptotic distributions irrelevant as well. Theorists must look elsewhere for explanations of the power function speedup.

This issue is important enough to be explained in some detail: The variable in all three of the asymptotic distributions is a linear transformation of the variable, x , in the parent distribution:

$$x \rightarrow a_n + b_n x.$$

The additive and multiplicative terms (a_n and b_n , respectively) are functions of n . For each asymptotic distribution, there is a series of additive and multiplicative terms whose function is to “undo” the effect of raising $[1 - F(x)]$ in Equation 1 to the n th power and thereby prevent the asymptotic distribution from degenerating. For the Weibull distribution, the additive term, a_n , drops out because the Weibull is bounded at the lower tail and the variable, x , can be transformed so that the lower bound is 0.0. The multiplicative term, b_n is a power function, $n^{1/\alpha}$. Transforming the scale of the Weibull distribution in Equation 3 by this power function prevents the asymptotic distribution from degenerating in the following manner:

$$\begin{aligned} F_{1:n}(n^{-1/\alpha}x) &= 1 - \exp[-(n^{1/\alpha}(n^{-1/\alpha}x))^\alpha] \\ &= 1 - \exp[-x^\alpha]. \end{aligned} \quad (4)$$

The result of the transformation is a Weibull distribution, not a degenerate function that is either 1 or 0.

It is important to note that the Weibull distribution that results from the transformation is not necessarily the same as the parent distribution (though it is in Equation 4). The Weibull distribution that results from the transformation depends on the exact value of the multiplicative term, b_n , and that term can take on an indefinitely large number of values. The power function, $n^{1/\alpha}$, could be multiplied or divided by a constant, have something added to it or subtracted from it, and so forth. In those cases, the asymptotic Weibull distribution would not be the same as the parent Weibull. It would have the same form (i.e., Weibull) but not the same parameters. According to Castillo (1988, p. 109), the multiplicative term, b_n , could be replaced by another, b_n^* . The only restriction on the substitution is that a candidate term, b_n^* , must satisfy

$$\lim_{n \rightarrow \infty} \frac{b_n^*}{b_n} = 1.$$

The important point is that the asymptotic distribution of minima does not result directly from the parent distribution. In-

stead, it requires a transformation of the variable in the parent distribution. There is no justification for transforming finishing times in the initial assumptions of the instance theory or in the additional assumptions that are required to formalize the race model. Thus, arguments based on asymptotic distributions are simply inappropriate. Justification for assuming the Weibull distribution and the proof of the power function speedup must be based on nonasymptotic arguments or arguments that work forward from an initial distribution (see Logan, 1988, 1992).

The Nonasymptotic Argument

Colonius (1995) suggested a nonasymptotic argument, based on the *min-stable* property of the Weibull distribution. The *min-stable* property refers to the fact that the distribution of minima sampled from a parent Weibull distribution is itself a Weibull distribution. The form of the distribution remains the same—remains stable—as sample size increases. The distribution simply changes in scale, following a power function, as is shown in Equation 3. This is the same nonasymptotic argument I made in 1988 and 1992, assuming that the parent distribution is the Weibull and inferring the power function speedup.

Colonius (1995) drew a different inference, arguing that the (observed) power function speedup and the (assumed) race between traces imply a Weibull distribution for reaction times. Apparently, minima sampled from a parent distribution will decrease as a power function of sample size (practice) only if the parent distribution is Weibull. This finding is important because it provides an a priori justification for assuming the Weibull distribution. Previously, the a priori justification for the Weibull was based on the asymptotic argument. Colonius showed that the asymptotic argument was flawed and that removed the a priori basis for assuming the Weibull distribution. The nonasymptotic argument, based on the min-stable property, provides a new justification for assuming the Weibull distribution, so the choice of distributions remains principled.

Colonius (1995) cited a theorem in Huang (1989) that proves that a sequence of means of minima uniquely determines the distribution function of the minima. This is an important theorem that would be well worth explaining to the psychological community because means generally do not constrain distributions uniquely (Townsend, 1990). The theorem is important to instance theory because it implies that the power function speedup in the means together with the race assumption (i.e., the assumption that the means are means of minima) implies a power function speedup in the distribution of reaction times. Moreover, Colonius argued that the power function speedup in the means and the race assumption imply a Weibull distribution of reaction times. This finding is important as well because it provides a further justification for choosing the Weibull distribution. This is an important contribution that adds to the strength of the instance theory.

Colonius (1995) proposed a shift in emphasis in testing instance theory as a result of Huang's (1989) theorem. Whereas I assumed a race and the Weibull distribution and predicted the power function speedup, Colonius recommended assuming the power function speedup and a race and predicting the Weibull distribution. He proposed a specific test of the Weibull distribution prediction:

$$n^{1/\alpha} X_{1:n} = X \quad \text{for all } n \geq 1,$$

where X is a nonnegative random variable with a nondegenerate continuous distribution function F . He claimed that this relation can be tested by comparing only two levels of practice, n_1 and n_2 , where $\log n_1 / \log n_2$ is irrational. He claimed that this test is better than simply testing for the Weibull shape at some given level of practice.

The test that Colonius proposed may not be better than simply testing for a Weibull shape. Colonius's test amounts to comparing one distribution against another, and that is essentially the same as comparing an empirical distribution against a theoretical one (i.e., the Weibull). In fact, it may be worse because it involves comparing two empirical distributions, each of which is noisy, whereas comparing an empirical distribution with a theoretical one involves only one source of noise (there is no noise in "estimating" the theoretical distribution).

Moreover, there is a danger in comparing only two levels of practice. Colonius ignored the consequences of a central assumption of the instance theory, that memory retrieval races against a general algorithm for performing the task. Colonius's arguments (and mine) require the assumption that the minima are sampled independently from identical distributions (the *iid* assumption). The *iid* assumption may be reasonable if only memory retrieval is involved because memory traces of different instances may well have the same distribution of retrieval times. But it is unlikely that an arbitrary algorithm will have the same distribution of finishing times as memory retrieval of a single instance. This is an important consideration, especially at low levels of practice, in which the algorithm dominates. In studies that collected strategy reports, it is evident that people persist in using the algorithm even at advanced levels of skill (Compton & Logan, 1991; Logan & Klapp, 1991). The danger in testing only two levels of practice is that one may be tempted to test too early, when performance is still dominated by the algorithm. That could lead to an inappropriate rejection of the Weibull distribution and thus, instance theory. Indeed, Logan (1992) found better agreement with the distributional predictions of the instance theory when the first few sessions were deleted from analysis (see also Logan & Etherton, 1994).

It seems to me that the best test of the Weibull assumption involves fitting a set of constrained Weibull distributions to empirical distributions over a wide range of practice. The Weibulls are constrained to have the same shape parameter, α , and the (power function) reduction in the scale of the distribution over practice is constrained to have an exponent equal to the reciprocal of the Weibull exponent (i.e., $1/\alpha$). I performed this test on 12 data sets in Logan (1992) and found reasonable fits, especially at higher levels of practice for which the algorithm was assumed to have dropped out.

Why the Weibull?

Colonius (1995) suggested a shift of focus from the power function speedup to the Weibull distribution in testing the instance theory. He argued that if one assumes a race and observes a power function speedup, then the distribution must be Weibull. If the empirical distributions are not Weibull, then the race assumption is invalid. This is an important contribution be-

cause it makes the theory falsifiable. So far, the few data that have been subjected to the test have found a Weibull shape (Logan, 1992), so the race assumption appears to be valid. Nevertheless, I am uneasy about the shift in emphasis for three reasons.

First, the power function speedup that Newell and Rosenbloom (1981) addressed is a speedup in mean reaction time only, not a speedup in the entire distribution of reaction times. The power function speedup in distributions is a relatively new discovery, and most of the evidence has come from my laboratory (but see Strayer & Kramer, 1990). I would like to see more evidence for its robustness and generality before making it a cornerstone assumption of instance theory instead of a prediction to be tested. In general, ordering of means does not imply ordering of distributions (though ordering of distributions implies ordering of means; Townsend, 1990), so a power function speedup in mean reaction time does not in itself imply a power function speedup in the distribution of reaction times.¹ Indeed, most theorists focused only on mean reaction times, and it was the instance theory that drew attention to the speedup in the entire distribution of reaction times (Logan, 1988, 1992). Other construals of skill acquisition do not imply a power function speedup in the distribution. For example, a probability mixture of slow nonautomatic processes and fast automatic processes could produce a power function speedup in the means if the mixture probability changed appropriately, but a mixture model would not predict a power function reduction in the distribution. Instead, variability would increase and then decrease over practice (Compton & Logan, 1991). Thus, the power function speedup in the distribution of reaction times is a major discovery that was inspired by instance theory, and the instance theory was the first to account for it (see Anderson, 1992; Cohen et al., 1990). Moreover the instance theory predicts a constraint between the shape of the reaction time distribution and the shape of the power function that no other theory predicts (i.e., the exponent, α , of the Weibull, which determines its shape, is the reciprocal of the exponent of the power function, $1/\alpha$, which determines its shape; Logan, 1992).

Second, the Weibull is a very flexible distribution. Its shape can vary dramatically, depending on the exponent, α , ranging from exponential to normal. Logan (1992, Appendix B) showed that the Weibull closely mimicked the ex-Gaussian distribution (a convolution of a normal distribution and an exponential distribution) that provides an excellent approximation to observed reaction time distributions (Ratcliff & Murdock, 1976). Van Zandt and Ratcliff (1995) showed that the Weibull closely mimicked a mixture of gamma distributions. Moreover, a mixture of gammas chosen in the right way could approximate the power function reduction in the scale of the distribution with practice that is characteristic of the Weibull. Thus, it may be difficult to discriminate the Weibull from similar-look-

¹ Huang's (1989) theorem, which states that a power function speedup in the means implies a power function speedup in the distribution, assumes that the means are means of minima (i.e., it assumes a race between memory traces). Other theories account for the power function speedup without assuming a race, and consequently, the speedup in the means does not imply a speedup in the distributions for those theories.

ing alternatives, such as the log-normal. Given the present state of knowledge, it may be reasonable to assume the Weibull distribution to develop the power-function predictions or to use it as an approximation to empirical distributions.

Finally, I am uneasy about using the Weibull distribution to diagnose the race model because the test may be unrealistically stringent. Other processes besides memory retrieval contribute to observed reaction times. There must be some encoding processes that transform physical stimuli to neural codes and some motor processes that transform neural codes to overt responses. These other processes must take time, and it is unlikely that the time they take will be constant. Thus, the observed reaction time distribution should be a convolution of the Weibull distribution and one or two other distributions that represent the time for encoding and responding. It may be that the durations of these processes are relatively short and, more important, relatively constant, because the Weibull distribution by itself provides a reasonable description of empirical reaction time distributions (Logan, 1992). However, there is no guarantee that encoding and response processes will be relatively invariant in all applications. When more powerful experiments are conducted, variability in encoding and response may become more problematic. Indeed, Logan (1992) found some perturbations in fitting Weibulls to empirical reaction time distributions that were accounted for by assuming variability in encoding and response processes.

One solution to this "intercept" problem would be to model the encoding and response processes in sufficient detail to specify the distributions of their processing times. That would be an important step but a difficult one given the current state of knowledge and the (normal) amount of controversy in the literatures on perception and motor control. In the mean time, it may be better to compare the instance theory with specific alternatives than to attempt to perfect its fit to a given data set. As with Compton and Logan's (1991) comparison between the instance theory and a probability mixture of fast (automatic) and slow (algorithmic) processes, it may be possible to find qualitatively different predictions that distinguish the alternatives without requiring either theory to fit the data perfectly. No other theory models the entire reaction process, and it is unreasonable to expect the instance theory to do so.

Future Mathematical Developments

It is encouraging to have someone with Colonius's mathematical knowledge and ability working on the instance theory. The statistics of extreme values are not very familiar to psychologists, and I am no exception. For me and many others I have spoken to, much of the attraction of the instance theory lies in its mathematical development. It is interesting and surprising that simple assumptions can lead to such powerful predictions. The assumptions may be overly simple, however, and it would be worthwhile to generalize them and make them more realistic.

The main problem with the mathematical development of the instance theory is that it relies heavily on the iid assumption. Many people express skepticism when I tell them I assume identical distributions. I already explained why the algorithm is unlikely to have the same distribution of finishing times as the

memory retrieval process. Many people think that it is likely that different traces may have different retrieval time distributions. There may be primacy and recency effects, such that the first- and last-learned items are retrieved faster than the others. Robert Nosofsky and Tom Palmeri (personal communication, November, 1994) are working on a generalization of the instance theory that assumes that retrieval time is a function of similarity. Stimulus encoding may be context dependent (Logan & Etherton, 1994) such that the same stimuli encoded in different contexts may be retrieved at different speeds.

It is easy to show with simulations that the power function predictions are robust with respect to violations of the iid assumption. Logan (1992, Appendix A) reported simulations from Weibull distributions that vary in location, scale, and shape parameters that continue to show the power function speedup for distributions. Logan (1992, Appendix A) provided proof that the power function speedup for distributions held for Weibull distributions that varied in location and scale parameters, with the shape parameter held constant. More work along these lines would advance theoretical development considerably.

Another important direction for mathematical development is to separate "intercept" processes such as perceptual encoding and response generation from the "central" processes that the instance theory addresses. The clearest separation may require well-developed theories of perception and motor control to interface with the instance theory, but it may be possible to use more general techniques that deconvolve the distributions without specifying the processes that generate them. The intercept problem is pervasive in models of reaction time because theorists focus primarily on one or two processes and hope that the other processes do not contaminate their predictions too much. A general solution to the intercept problem would be an important step forward for the instance theory in particular and for models of reaction time in general.

In summary, Colonius (1995) agreed with the substance of the instance theory, accepting the idea that automaticity is well characterized as a race between traces of past instances retrieved from memory. His article primarily addressed one of two arguments I used to account for the ubiquitous power function speedup in reaction time. Although there was an error in the asymptotic argument, the nonasymptotic argument remains sound, and the instance theory remains a viable account of automatization.

References

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 192-210.
- Anderson, J. R. (1992). Automaticity and the ACT* theory. *American Journal of Psychology*, 105, 165-180.
- Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96, 703-719.
- Castillo, E. (1988). *Extreme value theory in engineering*. San Diego, CA: Academic Press.
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97, 332-361.
- Colonius, H. (1995). The instance theory of automaticity: Why the Weibull? *Psychological Review*, 2, 744-750.
- Compton, B. J., & Logan, G. D. (1991). The transition from algorithm

- to memory retrieval in memory-based theories of automaticity. *Memory & Cognition*, *19*, 151–158.
- Gnedenko, B. V. (1943). Sur la distribution limitée du terme maximum d'une série aléatoire [On the limit distribution of maximum terms from a random series]. *Annals of Mathematics*, *44*, 423–453.
- Grant, S. C., & Logan, G. D. (1993). The loss of repetition priming and automaticity as a function of degree of initial learning. *Memory & Cognition*, *21*, 611–618.
- Huang, J. S. (1989). Moment problem of order statistics: A review. *International Statistical Review*, *57*, 59–66.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.
- Lassaline, M. E., & Logan, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 561–581.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.
- Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, *22*, 1–35.
- 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.
- Logan, G. D., & Etherton, J. L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1022–1050.
- Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 179–195.
- MacKay, D. G. (1982). The problem of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, *89*, 483–506.
- Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.
- Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*, 190–214.
- Schneider, W. (1985). Toward a model of attention and the development of automatic processing. In M. I. Posner & O. S. Marin (Eds.), *Attention and performance XI* (pp. 475–492). Hillsdale, NJ: Erlbaum.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.
- Strayer, D. L., & Kramer, A. F. (1990). An analysis of memory-based theories of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 291–304.
- Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, *108*, 551–567.
- van Zandt, T., & Ratcliff, R. (1995). Statistical mimicking of reaction time data: Single-process models, parameter variability, and mixtures. *Psychonomic Bulletin and Review*, *2*, 20–54.

Received February 23, 1995

Revision received April 6, 1995

Accepted April 11, 1995 ■