PSYCHOMETRIKA— VOL. 70, NO. 3, 427–430
SEPTEMBER 2005
DOI: 10.1007/s11336-004-1215-4

THE APPLICABILITY OF DEADLINE MODELS: COMMENT ON GLICKMAN, GRAY, AND MORALES (2005)

JEFFREY N. ROUDER

UNIVERSITY OF MISSOURI-COLUMBIA

Glickman, Gray, and Morales (this issue) propose a statistical model for measuring the unobserved latency of stimulus-controlled processes. The model accounts for both speed and accuracy and does so by assuming that participants set an internal deadline. If a stimulus-controlled response is not produced by the deadline, the participant then guesses. The applicability of the model is discussed in this Comment. The deadline model yields specific predictions for the case in which stimulus difficulty is manipulated in a within-block manner. In this case, it is reasonable to assume that stimulus difficulty does not affect the deadline. It is shown that in common perceptual and cognitive domains, extant data do not fully meet these predictions. Hence, practitioners need be aware of the possibility and consequences of model misspecification.

Key words: Race model, Deadline Model Speed-Accuracy Trade-off, RT modeling

Models for response time (RT) in perception and cognition range from simple linear models (such as those used in ANOVA) to complex and detailed nonlinear models of specific psychological processing. Examples of the latter include neural network models (e.g., Anderson, 1991) and models based on nonlinear stochastic processes (e.g., Usher and McClelland, 2001). Whereas linear models are often too simple to capture first-order psychological properties, complex nonlinear models have unknown statistical properties and are often intractable for detailed analysis. The model of Glickman, Gray, and Morales (2005) is a rarity in this landscape: it has substantive psychological content with well-developed statistical properties. The goal of this Comment is to provide some context as to the applicability of the model in perception and cognition.

1. The Deadline Model Revisited

Glickman, et al.'s model is nearly identical to the *deadline model* (Ollman and Billington, 1972). In this model a stimulus-controlled process races a deadline. If the stimulus-controlled process terminates before the deadline, then a correct response is produced. If the deadline occurs before the stimulus-controlled process terminates, then a guessing response is produced. This model was initially viewed as attractive because of its parsimony and because of its ability to explain basic speed accuracy trade-off effects. To respond quickly, participants set earlier deadlines. The consequence is a loss of accuracy. To respond accurately, participants set a later deadline. The consequence is slower response times.

More recently, Ruthruff (1996) concluded that deadline models do not hold in basic tasks. His evidence is based on a *selective influence* test. He manipulated two factors in his experiments: the difficulty of the stimulus (easy versus hard) and the instructions regarding speed versus accuracy stress (speed stress versus accuracy stress). The model posits that the speed versus accuracy stress should affect the deadline exclusively. In particular, the deadline for speed stress should stochastically dominate that for accuracy stress. This constraint can be expressed in terms of the

---

This research is supported by NSF grant SES - 0095919 to J. Rouder, D. Sun, and P. Speckman.

Requests for reprints should be sent to Jeffrey N. Rouder, Department of Psychological Sciences, 210 McAlester Hall, University of Missouri, Columbia, MO 65211, USA. E-mail: rouderj@missouri.edu.

cumulative distribution functions (CDFs). Let F_{R_S} and F_{R_A} be the CDFs of the deadline under speed and accuracy stress, respectively,

$$F_{R_S}(t) \geq F_{R_A}(t), \quad \forall t. \tag{1}$$

Ruthruff intermixed hard and easy stimuli randomly within a block of trials, such that the participant did not know the degree of stimulus difficulty before stimulus onset. Under these conditions, stimulus difficulty should affect the stimulus-controlled process exclusively. In particular, the stimulus-controlled latency for easy stimuli should stochastically dominate that for hard stimuli. Let F_{T_E} and F_{T_H} be the CDFs of the stimulus-controlled process for easy and hard stimuli, respectively,

$$F_{T_E}(t) \geq F_{T_H}(t), \quad \forall t. \tag{2}$$

Under these conditions, Ruthruff derived the following prediction for CDFs which he terms the *deadline inequality*,

$$F_{ES} \geq F_{HS} + F_{EA}. \tag{3}$$

The subscripts *ES*, *HS*, and *EA* refer to the conditions of easy stimuli with speed stress, hard stimuli with speed stress, and easy stimuli with accuracy stress, respectively. A proof of the inequality can be found in Ruthruff (1996, pp. 58–59). To gain insight into the inequality, consider the quickest responses in the hard-stimuli/speed-stressed condition. These quick response times reflect the quickest deadlines for speed stress. Likewise, the quickest responses in the easy-stimuli/accuracy-stressed condition reflect the quickest stimulus-controlled times for easy stimuli. The quickest stimuli in the easy-stimuli/speed-stressed condition is the minimum of these two times, which cannot be arbitrarily fast.

Given the reasonableness of the assumptions, the deadline inequality is a good benchmark for the deadline model. Ruthruff tested the inequality in tasks in which participants assessed the relative brightness of stimuli or whether strings of characters are valid words (lexical decision task). Estimated CDFs for the easy stimuli with speed stress were too small; they were less than the sum of the other two CDFs in the relevant conditions. These results are problematic for Glickman, et al. deadline model.

2. Relationship Between Error and Correct Response Times

A second challenge for the deadline model comes from the relative speeds of error and correct responses in perceptual tasks. The basic finding is that when tasks are difficult, error responses are slower than correct responses, but when tasks are easy the reverse holds (Luce, 1986). This generalization not only holds across various experiments, but holds within individuals as well. When participants are presented a mixed block of easy and difficult stimuli, they often have faster error responses than correct responses for the easy ones, but slower error responses than correct responses for the more difficult ones. I term this the *hard/slow-easy/fast error pattern*. Ratcliff and I found this pattern holds for brightness discrimination, color discrimination, and same–different tasks (Ratcliff and Rouder, 1998). Figure 1 shows data from a typical participant in Ratcliff and Rouder's Experiment 1. Stimulus difficulty was collapsed into seven levels. For each level two points are plotted, one for correct responses and one for error responses; these two points are connected by a dotted line. The four easiest levels have accuracies above 90% and response times below 800 ms. For these stimuli, the correct responses are slower than the corresponding errors, i.e., the dotted line has positive slope. The three difficult conditions have accuracies below 90% and response times above 800 ms. For these stimuli, correct responses are faster than the corresponding errors, i.e., the dotted line has negative slope. Others have also found

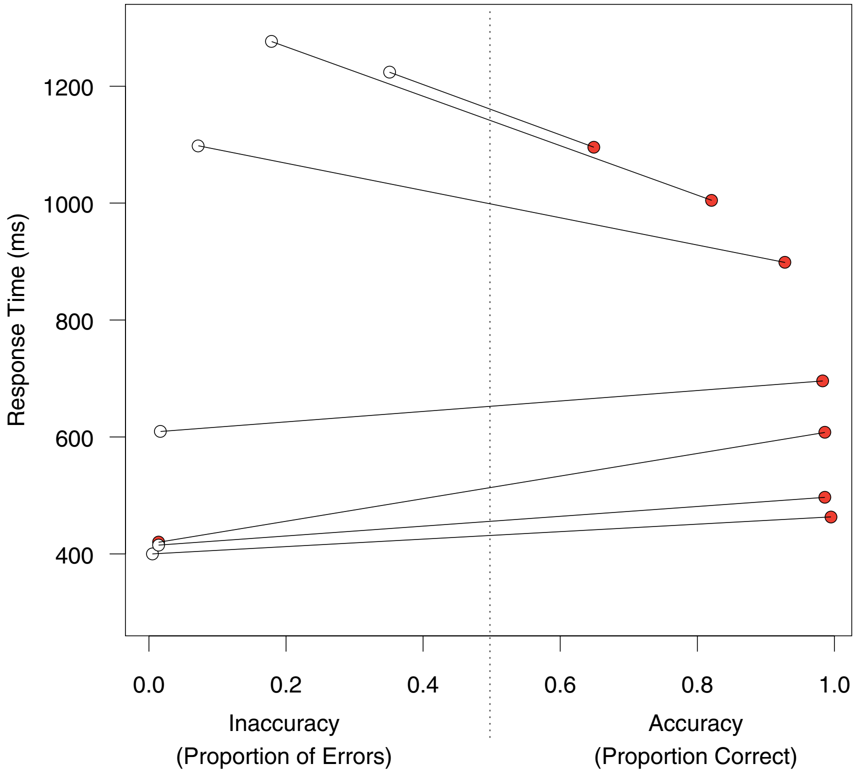

FIGURE 1.
Response time as a function of accuracy for Participant KR in Ratcliff and Rouder's (1998) Experiment 1. Filled circles denote correct responses; the *x* axis is the accuracy and the *y* axis is the response time. Open circles denote error responses; the *x* axis is the inaccuracy and the *y* axis is the response time. Solid lines connect points from the same condition. In the four conditions in which accuracy was above 90%, correct responses were slower than their corresponding errors (positively sloped lines). In the other three more difficult conditions, correct responses were quicker than their corresponding errors (negatively sloped lines).

the hard/slow-easy/fast error pattern (Espinoza-Varas and Watson, 1994; Smith and Vickers, 1988) and, as a result, this pattern has become an informal benchmark for models of RT and accuracy (Brown and Heathcote, 2005; Usher and McClelland, 2001).

Because stimulus difficulty is manipulated in a within-block manner, the deadline cannot depend on stimulus difficulty. Further, it is reasonable to assume that the stimulus-controlled latencies for easier conditions dominate those for more difficult conditions, i.e., equation (2) holds. Glickman, et al. provide an example in which the deadline model does indeed provide for the pattern (their Figure 1). Unfortunately, the deadline depends on stimulus difficulty in this figure; hence the selective influence condition is not met.

To gather some insight into whether the deadline model can account for the hard/slow-easy/fast pattern, I performed computer simulations of the model (see www.missouri.edu/˜pcl). The deadline model can qualitatively account for the hard/slow-easy/fast pattern, but only for tightly constrained parameter values. The shape of the deadline distribution must be highly symmetric or leftward skewed. In these cases, however, predicted RT distributions for hard stimuli are also symmetric—far too symmetric to account for empirical RT distributions. For example, predicted RT distributions had a skewness of 0.29 whereas empirical RT distributions for hard stimuli in Ratcliff and Rouder's (1998) data had skewnesses between 1.5 and 2.5.

3. Conclusion

Glickman, et al.'s deadline model is a strong contribution. It provides a natural means of modeling treatment group effects as well as accounting for nuisance covariates (e.g., heart rate), even when participants may have individualized criteria for trading speed for accuracy. These benefits should not be discounted. Cognitive and perceptual psychologists have not focused on developing models with statistically desirable properties. Glickman, et al.'s development serves as a model for what can be accomplished in a relatively simple framework. The above critiques serve as a cautionary note rather than a rejection of the model—researchers intending to use this model should be aware of the possibility of misspecifications and should explore the consequences on inference. I have shown the strong probability of misspecification in relatively simple tasks such as word and object identifications. It is an open question as to whether these misspecifications exist in more complex domains (e.g., learning algebra).

My colleagues and I have been concerned with developing substantive models with tractable statistical properties; our approach is based on Weibull models with hierarchical priors (Rouder, Sun, Speckman, Lu, and Zhou, 2003; Rouder, Lu, Speckman, Sun, and Jiang, 2005). We offer interpretations of Weibull shift, scale, and shape parameters in terms of psychological processes. We also allow for a great degree of individual variation—each individual is accorded his or her unique shift, scale, and shape parameters. By using hierarchical priors, efficient estimation through Bayesian shrinkage is obtained. The advantage of Glickman, et al.'s current model is the ability to account for error responses as well as factors that lead participants to trade speed for accuracy. The two approaches are complementary rather than oppositional: they each provide unique advantages and disadvantages and either one may be more appropriate for certain questions in certain domains.

References

Anderson, J.A. (1991). Why, having so many neurons, do we have so few thoughts? In W.E. Hockley & S. Lewandowsky (Eds.), *Relating Theory and Data: Essays on Human Memory in Honor of Bennet B. Murdock* (pp. 477–507). Hillsdale, NJ: Erlbaum.

Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review, 112*, 117–128.

Espinoza-Varas, B., & Watson, C. (1994). Effects of decision criterion on latencies of binary decisions. *Perception & Psychophysics, 55*, 190–203.

Luce, R.D. (1986). *Response Times.* New York: Oxford University Press.

Ollman, R.T., & Billington, M.J. (1972). The deadline model for simple reaction times. *Cognitive Psychology, 3*, 311–336.

Ratcliff, R., & Rouder, J.N. (1998). Modeling response times for decisions between two choices. *Psychological Science, 9*, 347–356.

Rouder, J., Lu, J., Speckman, P.L., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin and Review, 12*, 195–223.

Rouder, J.N., Sun, D., Speckman, P.L., Lu, J., & Zhou, D. (2003). A hierarchical Bayesian statistical framework for response time distributions. *Psychometrika, 68*, 589–606.

Ruthruff, E. (1996). A test of the deadline model for speed-accuracy tradeoffs. *Perception & Psychophysics, 55*, 56–64.

Smith, P.L., & Vickers, D. (1988). The accumulator model of two choice discrimination. *Journal of Mathematical Psychology, 32*, 135–168.

Usher, M., & McClelland, J.L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review, 108*, 550–592.

Manuscript received 18 MAY 2004
Final version accepted 16 JUNE 2004