

Bias produced by fast guessing in distribution-based tests of race models

JEFF MILLER and ALBANO LOPES
University of California, San Diego, California

A comparison involving cumulative probability distributions of reaction time (RT) has been used to test race models of the redundancy gain observed in certain divided-attention paradigms. It has been pointed out, however, that the presence of fast guesses would interfere with this test, biasing it to accept race models. The present paper reports simulations carried out to determine the size of the bias introduced by fast guesses. In absolute terms, this bias can be quite large—exceeding 175 msec in some conditions. Simulations indicate that the bias increases with the percentage of fast guesses and with the latency difference between the lower tails of guess and nonguess RT distributions. Discarding and rerunning errors reduces bias somewhat, but a more elaborate “kill-the-twin” procedure reduces it much more.

A variety of reaction time (RT) paradigms include some trials on which a response can be initiated as soon as any one of several redundant pieces of stimulus information (often called *targets*) is detected. In a visual search task, for example, a “target-present” response can be initiated as soon as a target stimulus is found. If the display contains several targets, the multiple targets are redundant because the response can be initiated as soon as any one is found. Bimodal divided-attention tasks are closely analogous; targets can be presented to any single modality, or they can be presented redundantly to more than one modality at the same time. Similarly, in the *same-different* discrimination paradigm, stimuli may differ along one attribute, along another attribute, or redundantly along both attributes. In this paradigm, differing attributes are analogous to targets, because the *different* response can be initiated as soon as any differing attribute is found.

In all divided-attention paradigms of this type, responses are typically faster, on average, when several redundant targets are present than when only one target is present (e.g., Egesh, 1966; Holmgren, Juola, & Atkinson, 1974). In such situations, a *race model* is an obvious candidate to explain the faster responses obtained with redundant targets (Raab, 1962). According to race models, responses to redundant targets are generated by the faster of two separate processes, each of which is responsible for detecting one of the targets. That is, if RT_1 and RT_2 are

random variables corresponding to the finishing times for processes that detect Targets 1 and 2, respectively, then the finishing time for detecting redundant targets, RT_R , is

$$RT_R = \min(RT_1, RT_2). \quad (1)$$

Miller (1978) noted that race models predict the following relationship among the cumulative probability density functions (CDFs) of the finishing times for detecting targets under three conditions (Target 1, Target 2, and redundant targets). For every value of t ,

$$\Pr(RT_R \leq t) \leq \Pr(RT_1 \leq t) + \Pr(RT_2 \leq t) \quad (2)$$

(for discussion and elaboration, see Ashby & Townsend, 1986; Colonius, 1988; Ulrich & Giray, 1986). The left side of the inequality is simply the CDF of RT in the redundant condition, $F_R(t) \equiv \Pr(RT_R \leq t)$. Assuming that the racing processes operate at the same speed in the single- and redundant-target conditions, which would be true for a large class of race models, the two terms on the right side of the inequality are simply the CDFs of RT in the two single-target conditions, $F_1(t)$ and $F_2(t)$. Given this assumption, the inequality predicted by race models can be written in a compact and testable form as

$$F_R(t) \leq F_1(t) + F_2(t), \quad (3)$$

where $F_1(t)$ and $F_2(t)$ are the CDFs of RT in Conditions 1 and 2, respectively.

Inequality 3 is necessarily satisfied for large values of t , because the left and right sides of the inequality reach asymptotic values of 1 and 2, respectively. For small values of t , however, violations are predicted by a variety of models in which two redundant targets mutually activate a single response or otherwise interact during response activation (Grice, Canham, & Boroughs, 1984; Miller, 1982, 1986, 1991; Mordkoff & Yantis, 1991; Schwarz, 1989). For example, if each redundant target produces response activation and the two separate activations sum together before being compared to a response

This research was supported by National Institute of Mental Health Grant PHS-MH40733 to the first author. We would like to thank Patricia Haden, Craig Kunimoto, Rolf Ulrich, and an anonymous reviewer for helpful comments on earlier drafts of the manuscript, and Richard Freedman of the Ford Scientific Research Lab, Larry Riddle of Emory University, and Juliet Schaffer of U. C. Berkeley, who independently provided information useful in implementing the simulations. Correspondence concerning this article should be addressed to Jeff Miller, Department of Psychology, C-009, UCSD, La Jolla, CA 92093-0109 (e-mail: jomiller@ucsd.edu).

criterion, then the minimum RT in the redundant condition may well be less than the minimum in either single-target condition. This relation among minima violates the inequality, because for t equal to the redundant minimum $F_R(t) > 0$, although $F_1(t) + F_2(t) = 0$.

A number of researchers have tested race models using Inequality 3 (e.g., Diederich & Colonius, 1987; Dijkstra, 1990; Grice et al., 1984; Miller, 1978, 1981, 1982, 1986, 1991; Mordkoff & Yantis, 1991; Sanocki et al., 1985). In many situations, the inequality is clearly violated (e.g., Grice et al., 1984; Miller, 1981, 1982, 1986), ruling out race models for those situations. This conclusion is theoretically useful, although the type of model responsible for the violations is not yet clear (e.g., Diederich & Colonius, 1991; Miller, 1986, 1991; Mordkoff & Yantis, 1991; Schwarz, 1989). In some situations, though, the data satisfy Inequality 3, indicating that race models may be a plausible account of the speedup associated with redundant targets (e.g., Grice et al., 1984, Experiments 4 and 5; Miller, 1982, Experiments 4 and 5).

Although Inequality 3 seems to provide a useful distribution-level test of race models, its diagnostic power is threatened by a potential bias that can arise in experiments where subjects make occasional fast guesses. Fast guesses are responses given without processing the stimulus, naturally assumed to be much faster than regular responses. There is good reason to believe that they are often present, at least in small proportions, in RT experiments (e.g., Ollman, 1966; Yellott, 1971). As mentioned by Miller (1982) and discussed in detail by Eriksen (1988), the presence of fast guesses in an experiment biases tests of Inequality 3 in favor of race models.

Basically, bias arises because fast guesses will inflate the right side of Inequality 3 twice as fast as they inflate the left side, assuming that guesses are equally likely in all conditions, as indeed they must be if conditions are randomly intermixed and if subjects make fast guesses without processing the stimulus. More formally, suppose fast guesses are made with probability g . Then, the CDF in condition k is

$$F_{k,g}(t) = (1-g) \cdot F_k(t) + g \cdot G(t), \quad (4)$$

where $G(t)$ is the CDF of fast-guess latencies. If the subject makes fast guesses with probability g in all conditions, then Inequality 3 becomes

$$(1-g)F_R(t) + gG(t) \leq [(1-g)F_1(t) + gG(t)] + [(1-g)F_2(t) + gG(t)], \quad (5)$$

which simplifies to

$$F_R(t) \leq F_1(t) + F_2(t) + \frac{g}{1-g} \cdot G(t). \quad (6)$$

Since $g/(1-g) \cdot G(t)$ is necessarily nonnegative, Inequality 6 can be satisfied even if Inequality 3 is not. Thus, the presence of fast guesses creates a bias that tends to favor race models. The bias is a nuisance particularly at the low end of the observed RT distributions, because

values of $G(t)$ tend to be relatively large even for t values near the bottom of the nonguess RT distributions. Thus, fast guesses work against finding violations of race models in exactly the range of t values where such violations are to be expected.

It might seem that the solution to this potential source of bias is simply to exclude fast guesses from the analysis, but this is easier said than done. There is large random variation in RT, so the probability distributions of guess and nonguess RTs may overlap. If there is overlap, no RT-based criterion will separate the guesses from the nonguesses perfectly. In other words, excluding trials with responses faster than some cutoff will miss some fast guesses, exclude some nonguesses, or both. In any of these cases, Inequality 3 may no longer provide a valid test of race models.

Given that fast guesses may distort the picture obtained using Inequality 3 in almost any experiment, the influence of fast guesses on this inequality needs quantitative study. Although the direction of the effect of fast guesses is quite clear, the size of the effect is not. It might appear that the effect size would be relatively easy to determine using Inequality 6. It turns out, however, that this inequality cannot be used for one of the experimental procedures we are particularly interested in evaluating (i.e., the kill-the-twin procedure described below); in this article, therefore, we used simulation methods.

The present paper reports computer simulation studies examining the size of the bias produced by fast guesses on tests of Inequality 3. Our results show that the bias can be quite large—so, experimenters cannot afford to ignore this problem when testing race models, as they have sometimes done in the past (e.g., Miller, 1982). Fortunately, the results also show that there is an available experimental procedure that can greatly reduce this type of bias.

The current simulations assessed the biasing effect of fast guesses within three experimental procedures—any of which might plausibly be used by researchers testing race models—differing slightly in their attacks on the problem of identifying and excluding fast guesses. The *include-errors* procedure adopts the head-in-the-sand approach: RTs from correct and error trials are simply pooled together to form a single RT distribution for each condition, just as if fast guesses were not present. Obviously, this procedure was expected to suffer most from the bias introduced by fast guesses. The *rerun-errors* procedure is somewhat better, because errors are discarded from the data set. When this procedure is used, erroneous trials are typically rerun later in the block, to ensure that the estimated RT distributions contain the desired number of observations and are based only on correct responses. Obviously, however, this procedure is not perfect either, because half of the fast guesses (i.e., the correct ones) are included in the data set.

A third procedure, discussed by Eriksen (1988) and aptly named the *kill-the-twin* procedure by A. van der Heijden (see Eriksen, 1988), has special intuitive appeal for the

current problem. This procedure is based on the assumption that each error reflects a fast guess. Naturally, then, error trials are discarded and rerun later in a block. Furthermore, however, it is explicitly acknowledged that correct fast guesses are just as likely to occur as erroneous ones (assuming the usual two-choice task), and that correct fast guesses should also be excluded from the analysis to obtain an accurate picture of the nonguess RT distribution. Thus, each time an error occurs, the correct trial with the most similar RT (i.e., the "twin" trial) is also excluded and rerun. This procedure has recently been used by Mordkoff, Yantis, and Egeth (1990) and by Mordkoff and Yantis (1991), and it was evaluated here because it seems a plausible approach to the problem of eliminating fast guesses from the distribution of correct RTs.¹

SIMULATION ASSUMPTIONS

In principle, it is simple to assess the effect of fast guesses on tests of Inequality 3. One first sees how well the inequality fits data from simulated experiments without any fast guesses, and then one sees how the fits change in simulated experiments with nonzero proportions of fast guesses. In essence, the effect of guessing is evaluated by comparing "experimental" simulations including guesses against "control" simulations without them. The present simulations were conducted with five different fast-guessing probabilities (g): 0.00 (control), 0.02, 0.05, 0.10, and 0.20.

A number of explicit assumptions had to be made in order to generate simulated RTs from the appropriate sort of divided-attention experiment. First, it was necessary to have a model to generate RTs for each of the single-target conditions. The Ex-Gaussian model (Burbeck & Luce, 1982) was chosen, because it is computationally simple and provides a reasonable fit to observed RT distributions (e.g., Blough, 1988; Ghohson & Hohle, 1968a, 1968b; Hockley, 1984; Hohle, 1965; Ratcliff & Murdock, 1976). In accordance with this model, each single-target RT was generated by summing independently generated *normal* and *exponential* random variables. The means of the normal and exponential distributions varied across simulations; the standard deviation of the normal was fixed at 20 msec.

Second, it was necessary to have a model to generate RTs for the trials with redundant targets. Since the nature of the processes generating redundant RTs is at the center of the ongoing debate to which this research is relevant, the choice of a model for this purpose was quite difficult. The candidates include race models in which the processes detecting different targets have positively correlated, negatively correlated, or uncorrelated finishing times. There are also several alternatives to race models, and these have general forms quite different from one another (compare, e.g., the diverse approaches considered by Grice et al., 1984, Miller, 1986, 1991, and Mordkoff & Yantis, 1991).

The approach taken here was to generate redundant RTs in accordance with a boundary case of the race model. Specifically, redundant RTs were generated such that there would be *equality* between the left and right sides of Inequality 3. Although it may seem peculiar to base the simulations on the race model—given that the cases of interest are precisely those in which this model is false (and yet fails to be rejected because of fast guesses)—this seemed to be the most reasonable approach. We are mainly interested in finding out how likely it is that fast guesses could allow the race model to be accepted even though the data were actually generated by some other model. For this to be possible, even given the biasing effects of fast guesses, the actual model cannot produce results too different from those consistent with the extreme race model we simulated. Thus, simulations of the race model reveal the effects of fast guesses with a distribution of redundant RTs fairly close to the true one in the cases of interest. Another advantage of using this race model in the simulations is that it is a reasonably well understood benchmark against which other models can be compared. Finally, it should be noted that the choice of a model for redundant RTs probably does not make too much difference. Our major question concerns the *effect* of fast guesses on Inequality 3—that is, the differential bias on the left and right sides of the inequality. It seems likely that this differential will be approximately the same regardless of the exact values of these quantities, especially when guesses are considerably faster than nonguesses in all conditions. Thus, the effect of fast guesses may be roughly the same, regardless of the true model generating redundant RTs.

Third, it was necessary to have a model to generate RTs for fast guesses. Little information is available about the distribution of fast-guess times, but it seems plausible that these should be approximately the same as simple RTs—fast and with relatively low variance. Because RT distributions tend to become less skewed as their means become smaller (e.g., Hockley, 1984), it also seemed reasonable to use a symmetric, rather than a skewed, distribution for fast-guess RTs. Thus, fast-guess RTs were generated as normally distributed random variables, with $\mu = 200$ msec and $\sigma = 20$ msec.

SIMULATION PROCEDURE

A total of 1,215 sets of simulations were run, with each set defined by a combination of levels of the following six factors:

- Experimental procedure: include errors, rerun errors, or kill-the-twin.
- Probability of a fast guess, g : 0.00 (control), 0.02, 0.05, 0.10, or 0.20.
- Probability of a correct response on nonguess trials: 1.00, 0.95, or 0.90. This parameter was expected to be of special importance for the rerun-errors and kill-the-twin procedures.

- Number of trials per subject per condition, n : 10, 20, or 50.
- Mean of the normal distribution used in generating nonguess RTs: 200, 300, or 500 msec.
- Mean of the exponential distribution used in generating nonguess RTs: 50, 100, or 200 msec.

Simulating each subject involved generating RTs for three experimental conditions: the two single-target conditions and the redundant-target condition. To generate an individual RT, the first step was to randomly determine whether the trial was a guess or a nonguess according to the probability of a guess (g). If the trial was a guess, the RT was randomly sampled from a normal distribution, with $\mu = 200$ msec and $\sigma = 20$ msec. If not, the RT was randomly selected from the distribution for the appropriate experimental condition. On a single-target trial, the RT was simply the sum of independently generated normal and exponential random variables, with parameters varying across simulations as indicated above. On a redundant-target trial, the RT was randomly sampled from the distribution representing the boundary case of the race model, for reasons outlined in the previous section. At the beginning of each simulation, this distribution was approximated by constructing a table of 1,000 equally spaced percentile points (i.e., RT values having percentile ranks of 0.05, 0.15, 0.25, . . . 99.95 within this distribution). More formally, the table contained a list of 1,000 values of t_i such that $F_R(t_i) = F_1(t_i) + F_2(t_i) = (i - 0.5)/1,000$, for $i = 1, \dots, 1,000$. On any trial, then, a redundant-trial RT was generated simply by randomly selecting 1 of the 1,000 values in the table, with equal probabilities. To give a feeling for what these redundant-trial distributions look like, Figure 1 shows the CDFs of RT for guesses, single-target nonguess trials, and redundant-target nonguess trials for the parameter values producing the fastest single- and redundant-target responses.

All RTs (i.e., guesses and nonguesses) generated for each simulated subject were combined together and analyzed in terms of Inequality 3 using a data-analysis procedure typical in actual research (e.g., Miller, 1978). The n RTs obtained in the redundant-target condition, $T_{R,i}$, $i = 1, \dots, n$, were used to estimate corresponding true percentile points of the redundant distribution (cf. Ratcliff, 1979). Specifically, the i 'th fastest RT value, $T_{R,(i)}$, is an estimate of the true percentile point, $\xi_{R,i}$, defined as the value for which $F_R(\xi_{R,i}) = (i - 0.5)/n$.² For example, the fastest and slowest of 10 RTs are taken as the estimates of the 5th and 95th percentile points, respectively. The $2n$ RTs obtained in the two single-target conditions, $T_{1,i}$, $i = 1, \dots, n$ and $T_{2,i}$, $i = 1, \dots, n$, were pooled to obtain estimates for the corresponding true percentile points of $F_1(t) + F_2(t)$. Specifically, the i 'th fastest of the pooled RT values, $T_{1+2,(i)}$, estimates the true percentile point $\xi_{1+2,i}$, defined as the value for which $F_1(\xi_{1+2,i}) + F_2(\xi_{1+2,i}) = (i - 0.5)/n$. For example, the fastest and slowest of 20 pooled data values are estimates of the points at which the sums of percentiles (summing

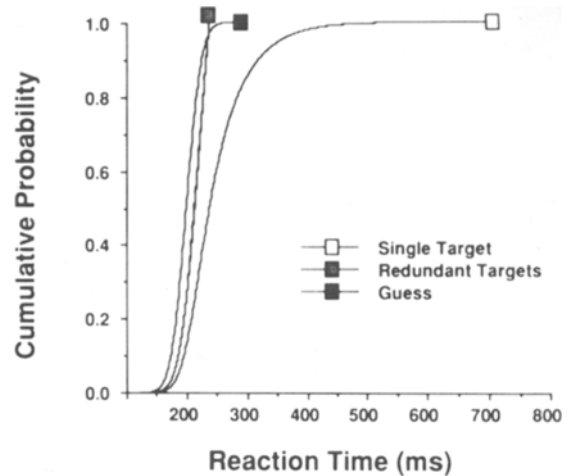


Figure 1. Cumulative probability distributions of RTs for fast guesses, single-target trials, and redundant-target trials. The fast-guess distribution is a normal with $\mu = 200$ and $\sigma = 20$. The single-target distribution is an Ex-Gaussian with normal parameters $\mu = 200$ and $\sigma = 20$ and exponential parameter $\mu = 50$. At each value of T , the redundant-target distribution has a cumulative probability exactly twice as large as the single-target distribution, up to the bound of 1.0.

across F_1 and F_2) are 5 and 195, respectively. In terms of percentile points, the prediction of race models corresponding to Inequality 3 is that $\xi_{R,i} \geq \xi_{1+2,i}$, for $i = 1, \dots, n$; in essence, the prediction is that a larger RT is needed to exceed a given percentage of the redundant-target distribution than is needed to exceed the same summed percentage across the two single-target distributions. This prediction can be evaluated statistically by comparing the estimated percentile points for the redundant-target distribution against estimates of the points with the same summed percentage across the single-target distributions. That is, we compared $\xi_{R,i}$ against $\xi_{1+2,i}$, where $\xi_{k,i}$ is the average across (simulated) subjects of the corresponding $T_{k,(i)}$'s ($k = 1 + 2$ or R).

Within each set of simulations, at least 200 subjects were simulated. Since it was not clear in advance how many subjects would be required to get stable estimates, the stopping rule was defined in terms of the desired accuracy for differences in estimated percentile points being compared. After each 100 simulated subjects in a given set of simulations, the standard error of the difference $\hat{\xi}_{R,i} - \hat{\xi}_{1+2,i}$ was computed for $i = 1, \dots, n$. This standard error was simply the standard deviation of the differences $T_{R,(i)} - T_{1+2,(i)}$, computed across simulated subjects, divided by the square root of the number of subjects simulated. Each set of simulations was terminated when the standard error of the difference was less than or equal to 1 msec for all n percentile points.³ The main values recorded for subsequent analysis were the averages, across simulated subjects, of $\xi_{R,i}$ and $\xi_{1+2,i}$ for $i = 1, \dots, n$.

Each simulation set was summarized in terms of n difference scores,

$$D_i = \hat{\xi}_{R,i} - \hat{\xi}_{1+2,i}, \quad (7)$$

for $i = 1, \dots, n$. Positive difference scores indicate that the percentile point of the redundant-target distribution is larger than the point having the same sum of percentiles across the two single-target conditions, as predicted by race models. Of particular interest were the biasing effects of fast guesses on these differences; therefore, it is convenient to explicitly indicate the dependence of these differences on guessing probability by writing them as $D_i(g)$, although of course the D_i values may depend on other parameters of the simulation as well. In any case, the biasing effect observed in a simulation with guess probability g may then be defined as

$$B_i(g) = D_i(g) - D_i(0). \quad (8)$$

That is, the bias measure $B_i(g)$ is the change in D_i values produced by a given proportion of fast guesses, with all other simulation parameters remaining fixed.

RESULTS AND DISCUSSION

Figure 2 displays the biasing effect of fast guesses as a function of the four simulation parameters having the largest effects:⁴ experimental procedure, probability of fast guess, mean of normal component of nonguess RT, and percentile.⁵ On the basis of the stopping rule for the simulations, these estimated bias values are very precise, each having a 95% confidence interval width of less than ± 2 msec, so there is little need to worry about random error in interpreting the large differences shown in the graphs.

Two very striking results of the simulations are immediately apparent from Figure 2. First, fast guesses can produce a huge bias in favor of race models, especially when errors are either included in the analysis or simply discarded and rerun. When errors are included, the presence of even 5% fast guesses produces bias approaching 90 msec at the 5th percentile with the normal mean of 500 msec. Since 5% fast guesses would translate into an error rate of only 2.5%—quite acceptable in most RT experiments—this must be seen as a strong signal that the fast-guessing problem needs to be addressed in order to test race models with the CDF approach. Furthermore, with 10% or 20% fast guesses, bias is both larger in absolute terms and more extensive over percentiles of the RT distribution, and even the corresponding 5% and 10% error rates would not be considered outrageous in many experiments.

The second striking result is that the three different experimental procedures have quite different sensitivities to bias (note the greatly magnified vertical scale for kill-the-twin). Across all simulations and percentiles, the average biases were 22.0, 12.2, and 1.6 msec for the include-errors, rerun-errors, and kill-the-twin procedures, respectively. Comparing the three graphs shown in Figure 2, it is clear that this difference in averages reflects a consistent ordering of procedures across simulation parameters and percentiles. It is not surprising that bias is much larger for the include-errors procedure than for the rerun-

errors procedure, because the former procedure includes twice as many fast guesses (i.e., correct and incorrect ones) in the analysis as the latter (i.e., only correct ones). What is perhaps surprising is that the kill-the-twin procedure is so very effective, relative to the other two, in excluding

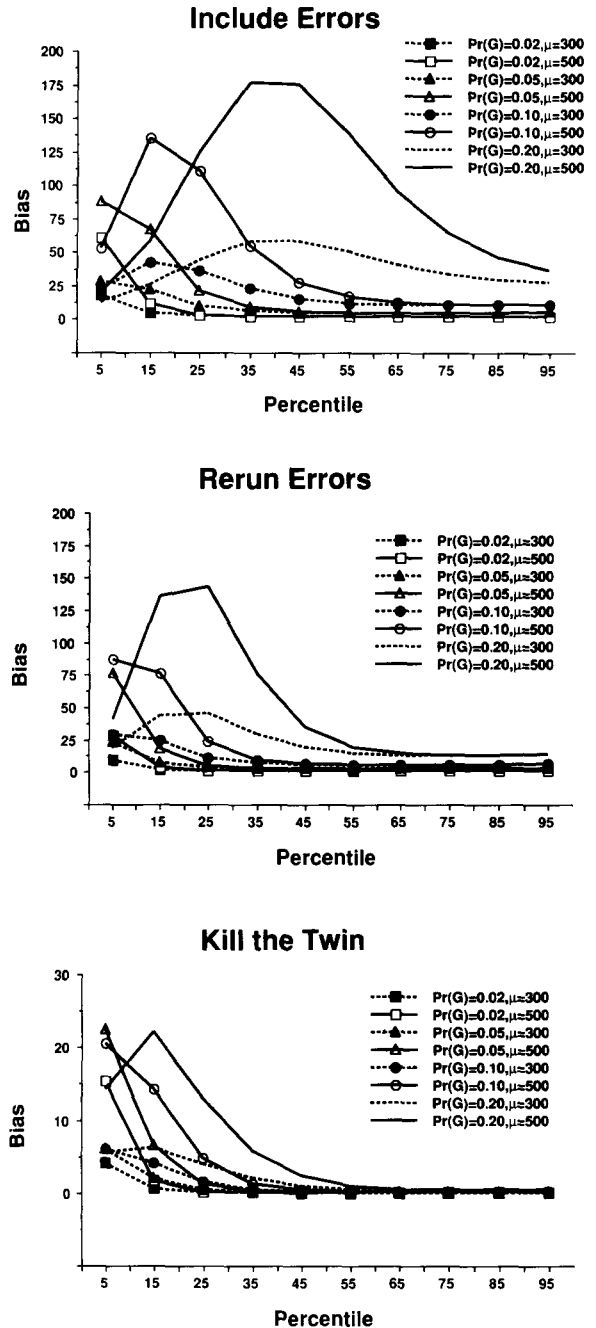


Figure 2. Bias introduced by fast guesses, plotted as a function of the probability of a fast guess, the mean of the normal component of the nonguess RT distribution, and percentile. Separate plots are shown for experiments including errors, excluding errors, and using the kill-the-twin procedure. Note the different scale on the vertical axis for the latter procedure.

fast guesses from the analysis altogether. Clearly, this procedure must have been very successful in identifying and excluding a high percentage of the *correct* fast guesses. Obviously, it cannot be perfect, of course, because fast guesses do not always come in offsetting pairs (i.e., one correct, one error) with closely matched latencies.

Two other factors having large effects were, predictably enough, the proportion of fast guesses and the RT percentile. As expected, the biasing effects of fast guesses were larger when there were more guesses, with average bias effects of 2.6, 6.5, 12.7, and 25.8 msec for guess probabilities of 2%, 5%, 10%, and 20%, respectively. Biasing effects also tended to be larger at the lower percentiles, as is apparent in the general decrease from left to right within each panel of Figure 2. This dependence is natural, because fast guesses are most likely to be used as estimates of lower percentile values.

There was also an interesting interaction between these two factors, most clearly apparent with $\mu = 500$. When the probability of a fast guess was low, bias decreased monotonically as percentile increased. When the probability of a fast guess was relatively high, however, bias was an inverted-U-shape function of percentile, peaking near the 40th percentile for the include-errors procedure and the 25th percentile for the rerun-errors procedure. Why did bias actually drop off at the lowest percentiles when guessing probability was high? The reason is that, for the low percentiles, the simulated values $T_{R,(i)}$ were very likely to come from the guessing distribution, just as the simulated values $T_{1+2,(i)}$ usually did. Since the two T values came from the same distribution, there was little difference between them, and therefore little bias introduced by guessing at this point. For higher percentiles, on the other hand, there was a much better chance that $T_{R,(i)}$ would come from the nonguess distribution while $T_{1+2,(i)}$ came from the guessing distribution, because there were likely to be twice as many guesses in the pool of $2n$ single-target trials as in the n redundant-target trials. In this case, there would be a large difference between the two T values from different distributions and, hence, a large bias introduced by guessing.

Another parameter having a very large effect on bias is the mean of the normal component of the normal-plus-exponential RT distribution for nonguesses. When this mean was 200 msec, there was virtually no bias—regardless of the mean of the exponential component—so the simulations using this value have been omitted from Figure 2. Biases of as much as 50 msec were observed, however, when this mean was 300 msec; biases exceeding 175 msec were observed when this mean was 500 msec.

Interestingly, the mean of the exponential component had virtually no effect on bias, so this factor was not included in the figure. In fact, average bias changed only 4 msec as the exponential mean increased from 50 to 200 msec. Thus, bias is not directly sensitive to the overall mean RT of nonguesses, because changing this mean by adjusting the parameter of the exponential had little effect.

From the very different effects of the normal and exponential components on bias, it appears that a crucial

determinant of bias is the extent to which the responses in the lower tail of the guess distribution are faster than those in the lower tail of the nonguess distribution. Increasing the mean of the normal component essentially shifts the whole RT distribution, including the lower tail, up to larger values. Increasing the mean of the exponential component, on the other hand, mainly increases the positive skew, extending the upper tail of the RT distribution but having little effect on the lower tail. Since the former manipulation increases bias dramatically and the latter has virtually no effect, it seems reasonable to conclude that bias is mostly sensitive to the relative locations of the lower tails of the guess and nonguess distributions, at least for reasonably small guessing probabilities. This conclusion is also supported by the fact that there was virtually no bias when the normal component of the nonguess distribution had the same mean (i.e., 200 msec) as the guessing distribution, even though the nonguess distribution had an additional exponential component.

The remaining simulation parameters—the number of observations per condition, n , and the probability of a correct response on nonguess trials—had negligible main effects and interactions with other factors.

CONCLUSIONS

Researchers interested in conducting powerful tests of race models must deal with the possibility that fast guessing will prevent rejection of race models even when they are false. This is especially likely to happen to the extent that fast guesses are common and the lower tail of the distribution of guessing times is below the lower tail of the distribution of nonguessing times. In extreme cases, fast guessing could conceal true violations of race models larger than 175 msec.

The best solutions to this problem are procedural rather than statistical in nature. Increasing the number of observations per subject, for example, clearly has no effect on the bias. What would be effective, first and foremost, is to make sure that subjects do not adopt a strategy of making fast guesses. Thus, the experimenter should demand high accuracy in general and penalize unusually fast, inaccurate responses in particular (for a reasonable payoff schedule, see, e.g., Meyer, Yantis, Osman, & Smith, 1985). Unfortunately, one cannot demand perfect performance, because it is very difficult to interpret average RT in a condition in which no errors are observed, as discussed by Pachella (1974).

The kill-the-twin procedure should almost certainly be adopted in any circumstance where even a small number of fast guesses are likely, as indicated by data patterns such as (1) errors significantly faster than correct responses, (2) near-chance accuracy for the fastest responses, and/or (3) equal minimum RTs across conditions differing in mean RTs. Because of the potentially large effects of fast guesses, failure to correct for fast guesses would seriously weaken the conclusions of any experiment with sizable error rates that failed to find violations of Inequality 3.

REFERENCES

- ASHBY, F. G., & TOWNSEND, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154-179.
- BLOUGH, D. S. (1988). Quantitative relations between visual search speed and target-distractor similarity. *Perception & Psychophysics*, *43*, 57-71.
- BURBECK, S. L., & LUCE, R. D. (1982). Evidence from auditory simple reaction times for both change and level detectors. *Perception & Psychophysics*, *32*, 117-133.
- COLONIUS, H. (1988). Modeling the redundant signals effect by specifying the hazard function. *Perception & Psychophysics*, *43*, 604-606.
- DIEDERICH, A., & COLONIUS, H. (1987). Intersensory facilitation in the motor component? A reaction time analysis. *Psychological Research*, *49*, 23-29.
- DIEDERICH, A., & COLONIUS, H. (1991). A further test of the superposition model for the redundant-signals effect in bimodal detection. *Perception & Psychophysics*, *50*, 83-86.
- DIJKSTRA, A. F. J. (1990). *Cross-modal contacts between graphemes and phonemes: Explorations in bimodal processing*. Doctoral dissertation, University of Nijmegen, Nijmegen, The Netherlands.
- EGETH, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, *1*, 245-252.
- ERIKSEN, C. W. (1988). A source of error in attempts to distinguish coactivation from separate activation in the perception of redundant targets. *Perception & Psychophysics*, *44*, 191-193.
- GHOLSON, B., & HOHLE, R. H. (1968a). Choice reaction times to hues printed in conflicting hue names and nonsense words. *Journal of Experimental Psychology*, *76*, 413-418.
- GHOLSON, B., & HOHLE, R. H. (1968b). Verbal reaction times to hues vs. hue names and forms vs. form names. *Perception & Psychophysics*, *3*, 191-196.
- GRICE, G. R., CANHAM, L., & BOROUGHS, J. M. (1984). Combination rule for redundant information in reaction time tasks with divided attention. *Perception & Psychophysics*, *35*, 451-463.
- GRICE, G. R., NULLMEYER, R., & SPIKER, V. (1977). Application of variable criterion theory to choice reaction time. *Perception & Psychophysics*, *22*, 431-449.
- HOCKLEY, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 598-615.
- HOHLE, R. H. (1965). Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology*, *69*, 382-386.
- HOLMGREN, J. E., JUOLA, J. F., & ATKINSON, R. C. (1974). Response latency in visual search with redundancy in the visual display. *Perception & Psychophysics*, *16*, 123-128.
- MEYER, D. E., YANTIS, S., OSMAN, A. M., & SMITH, J. E. K. (1985). Temporal properties of human information processing: Tests of discrete versus continuous models. *Cognitive Psychology*, *17*, 445-518.
- MILLER, J. O. (1978). Multidimensional same-different judgments: Evidence against independent comparisons of dimensions. *Journal of Experimental Psychology: Human Perception & Performance*, *4*, 411-422.
- MILLER, J. O. (1981). Global precedence in attention and decision. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 1161-1174.
- MILLER, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247-279.
- MILLER, J. O. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics*, *40*, 331-343.
- MILLER, J. O. (1991). Channel interaction and the redundant targets effect in bimodal divided attention. *Journal of Experimental Psychology: Human Perception & Performance*, *17*, 160-169.
- MORDKOFF, J. T., & YANTIS, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception & Performance*, *17*, 520-538.
- MORDKOFF, J. T., YANTIS, S., & EGETH, H. E. (1990). Detecting conjunctions of color and form in parallel. *Perception & Psychophysics*, *48*, 157-168.
- OLLMAN, R. T. (1966). Fast guesses in choice reaction time. *Psychonomic Science*, *6*, 155-156.
- PACHELLA, R. G. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41-82). Hillsdale, NJ: Erlbaum.
- RAAB, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*, 574-590.
- RATCLIFF, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*, 446-461.
- RATCLIFF, R., & MURDOCK, B. B., JR. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*, 190-214.
- SANOCCI, T., GOLDMAN, K., WALTZ, J., COOK, C., EPSTEIN, W., & ODEN, G. C. (1985). Interaction of stimulus and contextual information during reading: Identifying words within sentences. *Memory & Cognition*, *13*, 145-157.
- SCHWARZ, W. (1989). A new model to explain the redundant-signals effect. *Perception & Psychophysics*, *46*, 498-500.
- ULRICH, R., & GIRAY, M. (1986). Separate-activation models with variable base times: Testability and checking of cross-channel dependency. *Perception & Psychophysics*, *39*, 248-254.
- YELLOTT, J. I., JR. (1971). Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology*, *8*, 159-199.

NOTES

1. Both articles by Mordkoff and his colleagues cite Grice, Nullmeyer, and Spiker (1977) as originators of the kill-the-twin procedure. Although Grice et al. did adjust their correct RT distributions as a function of the distribution of error RTs, their adjustment procedure was intended to serve a different purpose than the kill-the-twin procedure and is algebraically quite different from it (Grice, personal communication, July 1991), as is now acknowledged by Mordkoff (personal communication, July 1991). Thus, it appears that the first published discussion of the procedure was that of Eriksen (1988).
2. The subtraction of 0.5 assigns each ordered RT value to the middle of the $1/n$ th section of the distribution that it represents.
3. Since each of the simulations with $g > 0$ was to be compared against a corresponding control simulation with $g = 0$, it was desirable to have greater accuracy for the simulations with $g = 0$. Thus, a criterion of 0.1 msec rather than 1.0 msec was used for simulations with $g = 0$.
4. Using bias as the dependent variable, mean squares were computed for a seven-factor design, including percentile as a seventh factor in addition to the six factors varying across simulations. These mean squares were used to order the main effects and interactions of the different simulation parameters, and it is on this basis that factor effects and interactions are referred to as being relatively large or small.
5. For comparability across the different values of n , 10 percentile values were used: 0.05, 0.15, . . . , 0.95. The ξ 's corresponding to these values were provided directly by the simulations for $n = 10$ and $n = 50$, and they were obtained by linear interpolation for $n = 20$.

(Manuscript received April 22, 1991;
revision accepted for publication July 29, 1991.)