

# Target-mask similarity in backward recognition masking of perceived tone duration

DOMINIC W. MASSARO

*University of Wisconsin at Madison, Madison, Wisconsin 53706*

and

WENDY L. IDSON

*University of Texas at Austin, Austin, Texas 78712*

Previous research using a backward recognition masking paradigm revealed that the perceived duration of a target increased with actual target duration, the time available for processing the target, and the duration of the mask. A model was developed which submits that the percept of target duration increases during the presentation of the target and the following silent interval. The mask terminates the growth of the perceived duration of the target. In addition, the perceived duration of the mask adds something to the final duration judgment. The model has provided a good description of duration perception for both tones and vowels, by several converging measures. Yet, in the previous studies, the target and mask were identical in all respects except for duration. The influence of the masking tone may possibly have been limited to this idiosyncratic situation. The present study explored this possibility by manipulating the frequency similarity between the target and masking tones. The target and masking tones could now differ in both duration and frequency. The earlier results were replicated, demonstrating that an identity of the target and masking tones is not a necessary condition for the previous findings.

A good deal of research has recently been directed towards the perception of duration (e.g., Allan, 1976; Cantor & Thomas, 1976; Idson & Massaro, 1977; Kristofferson, 1977; Massaro & Idson, 1976, 1978; Thomas & Cantor, 1975, 1976). This work has indicated that duration differs in some important respects from other stimulus attributes (see Allan & Kristofferson, 1974; Massaro & Idson, 1976, for reviews). One of the most important of these differences concerns the close correspondence which is generally found between the physical and psychological dimensions of a stimulus. For example, the perceived pitch of a tone is primarily due to its frequency, subject to the influence of factors such as intensity. Yet, in both audition (Efron, 1970a, 1970b, 1970c; Gol'dburt, 1961; Idson & Massaro, 1977; Massaro & Idson, 1976, 1978) and vision (Cantor & Thomas, 1976; Thomas & Cantor, 1975, 1976), the

perceived duration of a stimulus has been shown to be a function not only of its physical characteristics, but also of the amount of time spent on its processing. A given stimulus, having constant physical characteristics, will have a variety of perceived durations, dependent upon the amount of available processing time.

Massaro and Idson (1976) attempted to explore the role of temporal parameters in the perception of duration. A backward recognition masking paradigm was employed, in which one of the two target tones of differing durations was presented on each trial, followed after a variable silent intertone interval by one of three duration masking tones. The subject was asked to identify the target tone as being long or short, ignoring the masking tone. Performance averaged over the two targets produced results quite similar to those found for other perceptual attributes. When a subject is asked to identify pitch (Hawkins, Thomas, Presson, Cozic, & Brookmire, 1974; Massaro, 1970), timbre (Massaro, 1972), loudness (Moore & Massaro, 1973), or sound lateralization (Massaro, Cohen, & Idson, 1976), performance usually improves as a negatively accelerated monotonic function of the intertone interval, asymptoting at a level comparable to that found when no masking tone is presented. The top panel of Figure 1 presents the average percentage of correct test tone identifications, as a function of the intertone interval and

This research was supported by U.S. Public Health Service Grant MH-19399 to D. W. Massaro, and was conducted while W. L. Idson was supported by a University of Wisconsin Graduate Fellowship. The work represents a collaborative effort; order of authorship is arbitrary. We would like to thank Sheldon M. Ebenholtz for suggesting the necessity for the current study, Howard J. Kallman for helpful comments, and Michael M. Cohen for help on the model fitting. Requests for reprints may be sent to either author: Dominic W. Massaro at the Department of Psychology, University of Wisconsin, Madison Wisconsin 53706, or Wendy L. Idson at the Department of Psychology, University of Texas, Austin, Texas 78712.

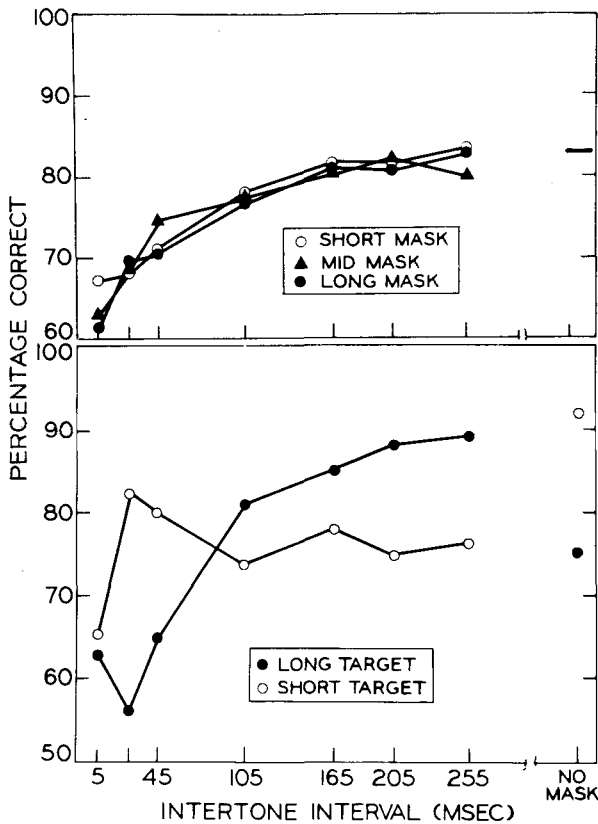


Figure 1. The top panel presents the percentage of correct identifications of the target, as a function of the duration of the mask and the intertone interval (ITI). The bottom panel presents the percentage of correct identifications of each of the target tones as a function of the ITI. (From Massaro & Idson, 1976.)

masking tone duration. Average performance improved as a negatively accelerated monotonic function of the intertone interval. This result was obtained for all three duration masking tones. Thus, the average masking function for duration reflects similar processes to those occurring for other stimulus attributes.

The difference between duration and other attributes becomes apparent when performance on the alternative target tones is considered. With other attributes, the alternative targets yield essentially equivalent performance; performance on both targets improves with the intertone interval. In contrast, the bottom panel of Figure 1 shows that markedly different results were found for the short and long target tones. While identification of the long target resembled the average function in the top panel, identification of the short target was better at short than at long intervals. Moreover, though identification of the short target was quite good on no-mask trials, identification of the long target was quite poor.

Massaro and Idson (1976) proposed a model to account for these results which constitutes an exten-

sion of a more general model of auditory recognition (Massaro, 1972, 1975). A sound is assumed to be initially stored in a very early preperceptual memory, having a capacity limit of a single sound. During the primary recognition process, information in this store is read out continuously over the course of approximately 250 msec. The preperceptual information is matched to information in long-term memory, in order to generate a synthesized percept of the sound. The resulting percept is stored in a synthesized auditory memory, which can hold multiple items for a period of several seconds. If a second masking sound is presented before processing of the first sound is completed, storage of the second sound will disrupt the representation of the first sound in preperceptual memory and interfere with its processing. Thus, the improved performance found in a backward recognition masking task with increases in the intertone interval between sounds represents the extraction of successively greater amounts of information from the target, given longer processing times prior to the onset of the mask. In the context of this model, the accuracy with which the target tone can be identified will increase with the intertone interval, but the average perceptual experience of the target will not change. For example, while the pitch of a tone will be resolved more accurately at longer intertone intervals, that tone will not come to sound consistently higher or lower as the interval is lengthened.

Perception of duration is assumed to occur in essentially the manner described by Massaro's model. However, duration differs from other attributes of the auditory stimulus in that the perceptual experience will also vary with increased processing time. Both the accuracy with which the duration of a target can be resolved and the perceived duration of that target will increase with the intertone interval in a backward masking task. If sufficient time is allowed for complete processing of the target, its perceived duration will be directly related to the temporal extent. If less than complete processing time is available, the target will be perceived as having a duration shorter than its asymptotic value. Consequently, perceived duration can be conceptualized as a negatively accelerated monotonic function of the intertone interval. If a masking tone is presented before the target has been completely processed, it will have the dual effect of decreasing the target's perceived duration and decreasing its discriminability from the other target. To the extent that the subject is able to process the target prior to the onset of the mask, both discriminability and perceived duration will increase. This insures that, at long intertone intervals, the perceived duration of the target will approach its asymptotic duration. At short intertone intervals, both the long and the short targets will have relatively short perceived durations, while at

long intertone intervals, both targets will have relatively long perceived durations. Therefore, at short intertone intervals, the long target will be inaccurately identified as short quite often, while identification of the short target will be relatively good. With increases in the intertone interval, a successively greater proportion of the targets will come to be classified as long, simultaneously increasing accuracy on the long target and decreasing accuracy on the short target. These were exactly the results obtained (see Figure 1).

The model assumes that the judged perceived duration, in the masking paradigm, is not solely a function of the information extracted from the preperceptual representation of the target. The perceived duration of the mask also influences the judged duration of the target. Given the relatively short intertone intervals between target and mask, it is unlikely that the subjects would be able to judge the perceived duration of the target prior to the onset of the mask (see Massaro & Idson, 1976, Experiment 5). Instead, the perceived duration of the mask will be incorporated into the judgment of the perceived duration of the target. The model assumes these two components to be independent; some proportion of the perceived duration of the mask is simply added to that of the target. Four possible mechanisms for the influence of the mask on the judged duration of the target will be evaluated in the Discussion section.

According to the model, the contribution of the mask will lengthen the judged duration of the target. Given sufficient processing time, a target will have a longer judged duration on a masking trial than on a no-mask trial. Thus, both targets will be classified as short disproportionately often on no-mask trials, yielding excellent performance on the short target and poor performance on the long target (see Figure 1). The lengthening effect of the mask was also supported by the strong Target by Mask interaction observed. For example, to the extent that the target and masking tones were both short, performance was better than if the target was short and the mask was long.

In order to provide converging evidence for this model, Idson and Massaro (1977) attempted to find a more direct index of the changes in perceived duration which occur in a backward masking task. Instead of categorizing the target as short or long, the subject's task was to estimate the duration of the target tone along a continuous scale of perceived duration, allowing the subjects to directly report the apparent duration of the tone. In accord with the model, the rated duration of the target tone increased monotonically with increases in the intertone interval. The top panel of Figure 2 presents the rated duration of each target tone, as a function of the intertone interval. As can be seen from the figure, the rated duration of the target increased as the intertone inter-

val was lengthened from 25 to 255 msec. The longer rated durations at the 5-msec interval reflect an integration of the target and masking tones into a single composite sound (Massaro, 1975), which yields a longer perceived duration at this than at any other interval (Idson & Massaro, 1977). Consonant with the assumption that the physical duration of the target contributes to its perceived duration, the long target was always rated as having a longer perceived duration than the short target at the same interval. For both targets, however, rated duration increased monotonically with the intertone interval, indicating that perceived duration increased with increasing processing time. Moreover, the targets were both rated as shorter on no-mask trials than on trials on which a mask was presented, supporting the assumption that the mask contributes a component to the perceived duration of the target. Additional support for this assumption came from the finding, shown in the bottom panel of Figure 2, that the rated duration of both targets increased with the duration of the mask. The targets were rated shortest when followed

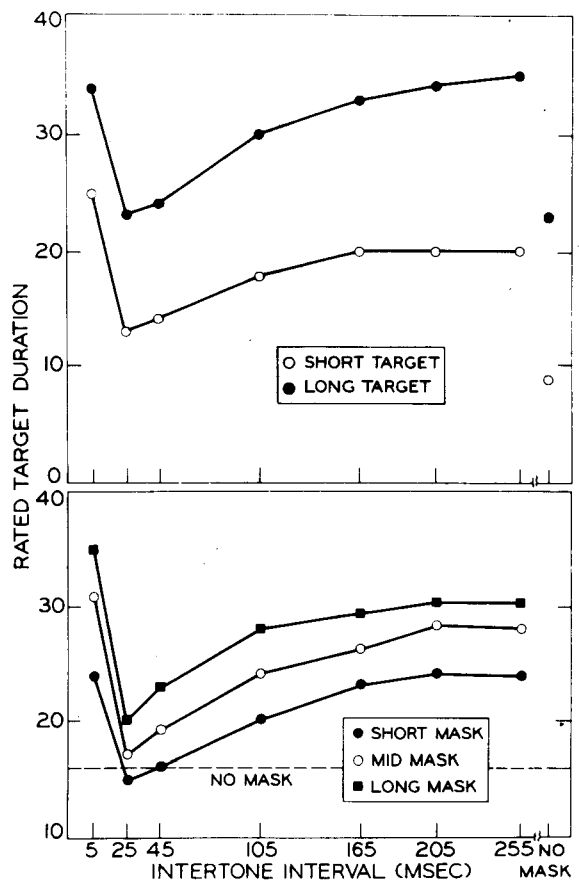


Figure 2. The top panel presents the rated durations of each of the target tones, as a function of the ITI. The bottom panel presents the rated durations of the target, as a function of the ITI and the duration of the mask (from Idson & Massaro, 1977).

by a short mask, and longest when followed by a long mask. Thus, the pattern of results for the rating data conform precisely to the predictions of the model.

The close correspondence between the rating data and the binary choice data provides qualitative support for the Massaro and Idson (1976) model. A more precise evaluation is possible by developing a quantitative description of changes in perceived duration. The perceived duration of the target is assumed to increase as a negatively accelerated function of processing time, with the rate of growth differing during the actual duration of the target and the silent interval prior to the onset of the mask. Consequently, the perceived duration (PD) of the target is given by

$$PD = x + y, \tag{1}$$

where  $x = \alpha[1 - e^{-(\theta_D t_D)}]$  and  $y = (\alpha - x)[1 - e^{-(\theta_I t_I)}]$ . The value of  $x$  is the perceived duration obtained during the actual duration of the target; the value of  $y$  is the perceived duration obtained during the silent interval prior to the onset of the mask. The value of  $\alpha$  is the asymptotic value of perceived duration,  $\theta_D$  represents the rate of growth of PD during the time of target presentation,  $t_D$ , and  $\theta_I$  represents the rate of growth of PD during the silent interval,  $t_I$ , following target offset.

When the mask is presented, it terminates further growth of the duration percept. Moreover, the mask serves to lengthen the judged duration (JD) of the target by a constant, proportional to the duration of the mask. Incorporating this assumption yields

$$JD = PD + K(t_m), \tag{2}$$

where  $t_m$  is mask duration and  $K$  is a constant of proportionality, which reflects the weight given the masking tone. If  $JD_S$  and  $JD_L$  represent the judged durations of the short and long targets, respectively, then the growth of the duration percept over time is given by:

$$JD_S = x_S + y_S + K_S(t_m), \tag{3}$$

where  $x_S = \alpha_S[1 - e^{-(\theta_D t_D)}]$ ,  $y_S = (\alpha_S - x_S)[1 - e^{-(\theta_S t_S)}]$ , and

$$JD_L = x_L + y_L + K_L(t_m), \tag{4}$$

where  $x_L = \alpha_L[1 - e^{-(\theta_D t_D)}]$ ,  $y_L = (\alpha_L - x_L)[1 - e^{-(\theta_L t_L)}]$ . The values of  $x_S$  and  $x_L$  are the perceived durations obtained during the actual presentation of the short and long targets, respectively. The values of  $y_S$  and  $y_L$  are the perceived durations obtained during processing after the offset of the short and long targets, respectively. The value of  $\theta_D$  represents the

growth rate during target presentation,  $t_D$ , and  $\theta_S$  and  $\theta_L$  represent the rates of growth of the short and long targets over the silent interval,  $t_I$ , to asymptotes  $\alpha_S$  and  $\alpha_L$ , respectively. The values of  $K_S$  and  $K_L$  are constants of proportionality for the short and long targets, respectively. The discriminability of the two targets, as given by  $d'$ , is then equal to the difference between their judged durations (see Massaro & Idson, 1976).

The quantitative version of the model just presented differs from that originally offered by Massaro and Idson (1976). The earlier version of the model assumed a single value of  $\theta$ , representing the rate of growth of PD, which differed for the long and short targets:

$$PD_S = \alpha_S(1 - e^{-\theta_S t}) \tag{5}$$

and

$$PD_L = \alpha_L(1 - e^{-\theta_L t}). \tag{6}$$

In order for the value of  $\theta$  to differ for the long and short targets, the duration percept would need to be processed differentially from target onset. Yet the stimulus events corresponding to the short and long targets do not differ until offset of the short target. Compartmentalizing the growth rate into a value for the time of target duration and the interval following target offset, as in Equations 3 and 4, resolves this discrepancy and several attendant problems which have been raised (Allan & Rousseau, 1977; Kristofferson, 1977) concerning the original model.<sup>1</sup>

The model presented in Equations 3 and 4 provides a good fit to both the binary choice data of Figure 1 and the rating data of Figure 2. In order to fit the model to the obtained data, the assumption is made that noise in the sensory system creates a distribution of judged durations for each experimental condition. The noise distribution is assumed to be normal with a variance equal to 1. The binary choice data were transformed into relative standard deviation units, which provide an index of judged duration. In the context of the model, the probability of responding long is a measure of the judged duration of the target. Thus, the probability of responding long given a long target,  $P(L | L)$ , and the probability of responding long given a short target,  $P(L | S)$ , were computed at each experimental condition. The values of  $P(L | L)$  or  $P(L | S)$  can be used to determine the location of the mean of the subjective durations at that condition with respect to the fixed criterion value. The  $z$  distance between the mean of the relevant distribution and the criterion value can be computed for both the long and the short targets. If  $P(L | L)$  or  $P(L | S)$  is greater than .5, this indicates that the mean of the distribution is to the right

of the criterion and the  $z$  distance will be positive. If either probability is less than .5, the mean of the distribution is to the left of the criterion and the  $z$  distance will be negative. In both cases, the  $z$  distance will correspond to the transform of the proportion of the long or short distribution lying between the mean and criterion (see Massaro, 1975, Chapter 7). The value of +3 was added to each  $z$  score to insure that all scores would be positive, allowing Equations 3 and 4 to be applied. The model was then fit to the values of  $z[P(L | L)]$  and  $z[P(L | S)]$ , using the iterative computer program STEPIT (Chandler, 1969). The 5-msec intertone interval was excluded from the fit, as the integration process which occurs at this interval cannot be described with the same parameter values (see Massaro & Idson, 1976).

The fit of the new model, given in Equations 3 and 4, to the binary choice data (Massaro & Idson, 1976, Experiment 1) is given in Figure 3, which presents the predicted and observed values of  $z[P(L | L)]$  and  $z[P(L | S)]$ . The predicted values were computed by estimating the parameter values  $\alpha_L$ ,  $\alpha_S$ ,  $\theta_L$ ,  $\theta_S$ ,  $\theta_D$ ,  $k_L$ , and  $k_S$  by minimizing the squared deviations between the predicted and observed values. The scale values of  $t_D$ ,  $t_I$ , and  $t_m$  are specified in seconds in the fit of the model. The model is providing a reasonably good fit to the data. Seven parameters were used to fit 38 data points (2 targets  $\times$  3 masks  $\times$  6 intertone intervals + 2 no-mask targets), with an average squared deviation between the predicted and observed values of .009. The parameter estimates were in accord with the model. The values of  $\alpha_L$  and  $\alpha_S$  were 3.84 and 1.81, respectively, the long target having a greater potential perceived duration than

the short target. The values of  $\theta_L$  and  $\theta_S$  were 14.47 and 102.36, the short target being processed at a faster rate than the long target. The estimated value of  $\theta_D$  was 9.93. The duration of the mask had a greater effect on the short than on the long target, the parameter estimates of  $k_L$  and  $k_S$  being 5.58 and 6.32, respectively.

The model can provide an equally good fit to the rating data. A data transformation analogous to that performed upon the binary choice data was performed upon the rating data. The ratings were assumed to provide a direct index of the judged duration of a target tone under a particular experimental condition. Given the additional assumption that the subjects were using an interval response scale, a linear transformation can be performed which will preserve the relations among the scale values under the various experimental conditions (Marks, 1974, Chapter 7), allowing a direct test of the predictions of Equations 3 and 4. A linear transformation of the form

$$\text{Transformed data} = a(\text{observed data}) + b \quad (8)$$

was employed in the following way. Values of  $a$  and  $b$  were estimated simultaneously with the estimates of the free parameters in Equations 3 and 4, in order to minimize the squared deviations between the predicted and observed values, using the iterative minimization routine STEPIT (Chandler, 1969). To insure that the transformed scores were in the same range for the rating data as for the binary choice data, the value of  $\alpha_L$  was set to that obtained in the fit to the binary choice data.

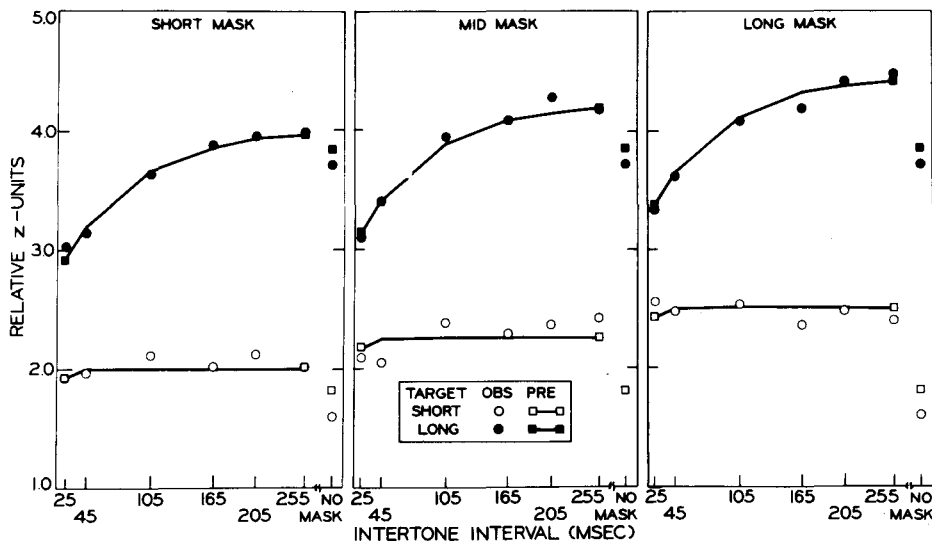


Figure 3. Predicted and observed values of  $z[P(L | L)]$  and  $z[P(L | S)]$  as a function of both the duration of the masking tone and the intertone interval. The no-mask values do not differ as a function of masking tone duration, since the latter is a dummy variable under the no-mask condition (observed data from Massaro & Idson, 1976, Experiment 1).

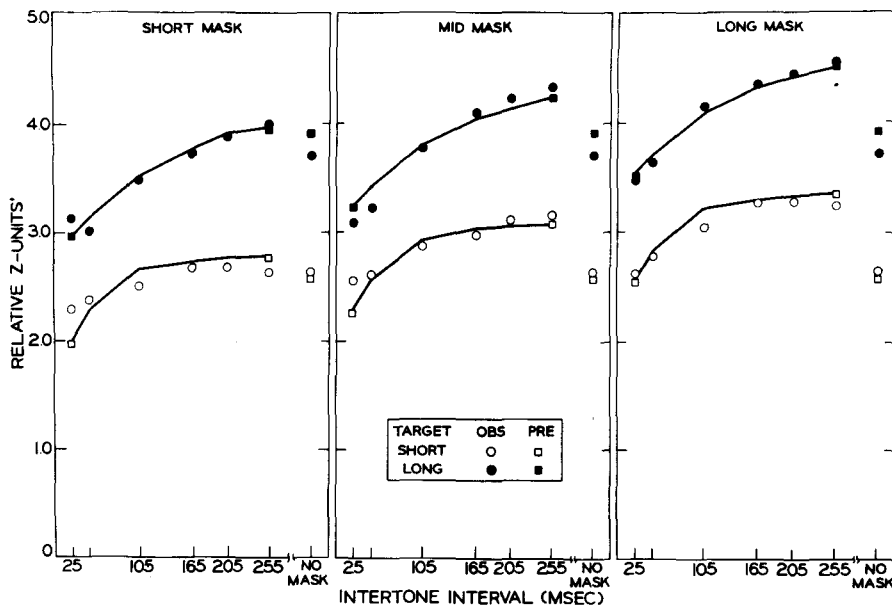


Figure 4. Predicted and observed value of  $z[P(L|L)]$  and  $z[P(L|S)]$  as a function of both the duration of the masking tone and the intertone interval. The no-mask values do not differ as a function of masking tone duration, since the latter is a dummy variable under the no-mask condition (predicted and observed data from Idson & Massaro, 1977, Experiment 2).

Figure 4 presents the predicted and observed values for the short and long target tones. The predicted values were computed by estimating the parameter values  $\alpha_S$ ,  $\theta_S$ ,  $\theta_L$ ,  $\theta_D$ ,  $K_S$ ,  $K_L$ ,  $a$ , and  $b$  by minimizing the squared deviations between the predicted and observed values. The model again provides a reasonable fit to the data. Thirty-eight points (2 targets  $\times$  3 masks  $\times$  6 ITIs + 2 no-mask targets) were fit by eight parameters, with an average squared deviation between the predicted and the observed values of .014. The parameter estimates were in accord with the model. The value of  $\alpha_L$  was set to 3.90, and the value of  $\alpha_S$  was estimated at 2.58. The long target again had a greater potential perceived duration than the short target, though the discriminability between the two targets was less than that observed for the binary choice data (see Idson & Massaro, 1977, for a discussion of this point). The parameter estimates for  $\alpha_L$  and  $\alpha_S$  were 8.81 and 25.00, respectively, indicating that the short target was processed at a faster rate than the long target. The parameter estimate for  $\theta_D$  was 11.15. The duration of the mask had a slightly larger effect upon the short target, the parameter estimates for  $K_S$  and  $K_L$  being 6.96 and 6.85, respectively. The parameter estimates for  $a$  and  $b$  were .075 and 1.75, respectively.

The ability of a single model to describe both the binary choice and the rating data suggests that the processes postulated by the model play an important role in the processing of duration. However, in all of the previous studies, the mask was identical to the target along all dimensions except duration. This

stimulus situation contrasts markedly with the more natural conditions in which temporally successive items differ along multiple dimensions. It is entirely possible that different processes are involved in duration perception when successive stimuli vary in terms of several attributes. If this were the case, then the applicability of the Massaro and Idson model would be limited, with little potential for describing the dynamics of duration processing in areas such as speech perception (see Massaro & Idson, 1978).

The present research was designed to evaluate the ability of the model to describe duration processing when the target and mask were qualitatively different. In the current study, the masking tone could differ from the target not only in duration, but also in frequency. The rating, rather than the binary choice task, was employed, since it provides a more direct measure of changes in judged duration. If the Massaro and Idson model is applicable to more natural conditions of perception, then the earlier results should be replicated, even when the target and mask differ in frequency.

## METHOD

### Subjects

The subjects were 13 University of Wisconsin undergraduates, who received credit in an introductory psychology course for their participation.

### Apparatus and Stimuli

Four subjects could be tested simultaneously in separate sound-insulated booths. All experimental events were controlled by a PDP-8/L computer. The tones were generated as sine waves by

a digitally controlled oscillator (Wavetek Model 155) and were presented binaurally over matched headphones (Grason-Stadler Model TDH-49). The tones were turned on and off at the zero crossing and had essentially instantaneous rise times. A continuous response scale was employed, consisting of a pointer connected to a potentiometer and a pushbutton. The subject set the pointer to the desired location along a 5.5-cm scale, and then pushed a pushbutton to indicate that a response had been made. The placement of the pointer was indicated by the voltage passed from the potentiometer to an analog-to-digital converter. The voltage given by a linear potentiometer allows the 5.5-cm scale to be mapped into 50 categories from left to right, which were assigned integer values from 0 to 49.

Two target tones were employed in the study, having durations of 50 msec (short) and 90 msec (long). Three masking tone durations, symmetrical around the durations of the target tones, were used—30 msec (short mask), 70 msec (mid mask), and 110 msec (long mask). The target tones were presented at a frequency of 700 Hz. The frequencies of the masking tones varied symmetrically about the frequency of the target—500 Hz (low-frequency mask), 700 Hz (same-frequency mask), and 900 Hz (high-frequency mask). The intensities of the target and masking tones were chosen randomly and independently on each trial from a set of seven possible intensities: 75 to 81 dB SPL in 1-dB steps. Intensity was randomized in this range to insure that observers would utilize duration and not loudness as the relevant dimension to be judged.

#### Procedure

The experiment was conducted on 5 consecutive days. Each day was divided into two 20-min sessions, separated by a 10-min rest break. Day 1 was considered to be practice. In the first session of that day, the subjects heard one of the two target tones, which they were to rate on the continuous scale from short to long. Presentation of the two targets was random, and each tone was programmed to occur equally often. The subjects were given 3 sec in which to make their responses, by locating the pointer along the continuous scale. Following the response interval, a 500-msec display of an asterisk (\*) was presented over a visual display of light-emitting diodes (Monsanto Model MDA-III) to indicate the end of the trial. The intertrial interval was 1 sec.

Session 2 of Day 1 and both sessions of the four subsequent experimental days consisted of 300 test trials. All of Session 2 of Day 1 and the first 12 trials of the eight additional sessions were treated as unscored practice trials, though the subjects were not aware of this fact. On each trial of the test sessions, the subject heard one of the two target tones. On 7/8 of the trials, the target was followed by one of the masking tones after a variable silent intertone interval (ITI) of 5, 25, 45, 105, 165, 205, or 255 msec. On the remaining 1/8 of the trials, no masking tone was presented (no-mask condition). The responses, feedback, and intertrial intervals were identical to those employed in the practice session. All 144 experimental conditions (2 targets  $\times$  3 mask durations  $\times$  3 mask frequencies  $\times$  7 ITIs + the no-mask) were completely random and were programmed to occur equally often within a session. Both the duration and the frequency of the mask were dummy variables under the no-mask condition.

## RESULTS

The ratings of target duration, indicated by placement of the pointer along the continuous response scale, were recorded as integers between 0 and 49, with 0 representing the shortest tones and 49 the longest tones. The responses were pooled over the 4 experimental days. The mean ratings of duration

were compiled for each subject, at each target duration by mask duration by mask frequency by ITI condition. Two separate analyses of variance were performed on these ratings. In one analysis, the no-mask was treated as an ITI of  $\infty$  and entered the analysis as a level of the ITI factor. In the second analysis, the no-mask was eliminated, and the ITI factor had only seven levels. This second analysis was conducted so as to allow a direct evaluation of the effects of masking tone duration and frequency, by eliminating those trials on which no masking tone was presented. All of the results given below which involve the masking tone as a factor were drawn from this second analysis, excluding the no-mask trials. All other results were taken from the analysis in which the no-mask trials were included.

Figure 5 presents the rated durations of the target tone, as a function of the ITI, for the three frequencies of the masking tones. The left-hand panel presents the results for the low-frequency mask, the middle panel the results for the same-frequency mask, and the right-hand panel the results for the high-frequency mask. The same results were observed for all three masking tone frequencies. The long target was consistently rated as being longer in duration than the short target. The rated durations of both targets increased consistently with the ITI. Moreover, for both the long and the short targets, rated duration on no-mask trials was shorter than the average rated duration of that target on masking trials. Both the main effect of target duration [ $F(1,7) = 38.77, p < .001$ ] and the interaction of this factor with the ITI [ $F(7,84) = 4.07, p < .005$ ] were significant in the analysis, though neither the interaction of target duration and masking tone frequency nor the three-way interaction of ITI, masking tone frequency, and target tone duration were significant.

Figure 6 presents the average rated durations of the target tone, as a function of both the ITI and the duration of the masking tone, for each of the three masking tone frequencies. The left-hand panel presents the results for the low-frequency mask, the middle panel the results for the same-frequency mask, and the right-hand panel the results for the high-frequency mask. For all three masking tone frequencies, the same pattern of results was observed. Rated target duration increased with the duration of the mask; at all ITIs, the target was rated as shortest when followed by the short mask and longest when followed by the long mask. For all three masking tone durations, the rated duration of both targets increased consistently with the ITI. The overall effect of increasing the duration of the masking tone was to increase the rated duration of the target, raising the overall level of performance on both targets. Both the Target Duration by Mask Duration interaction [ $F(2,24) = 7.71, p < .005$ ] and the three-way

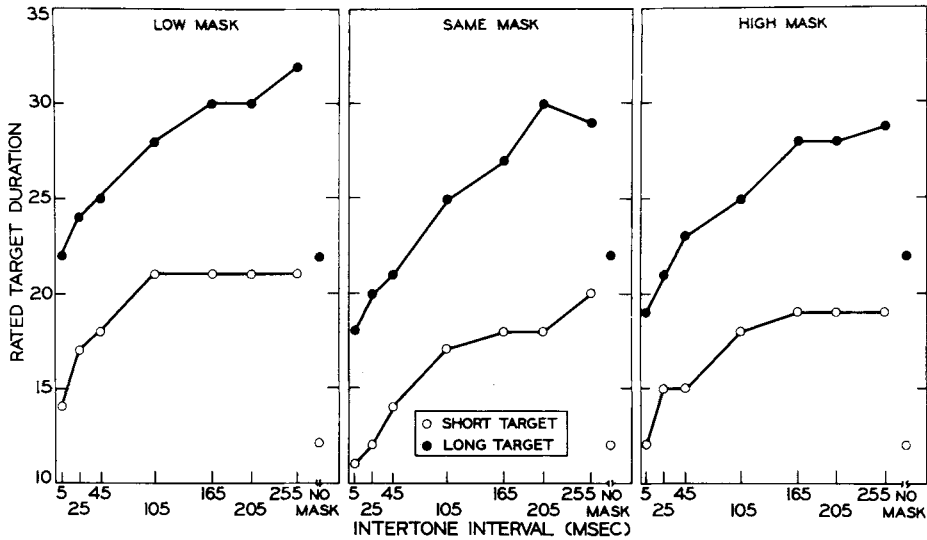


Figure 5. Rated duration of each of the target tones as a function of the ITI, for each of the masking tone frequencies. The left-hand panel presents the results for the low-frequency mask, the middle panel for the same-frequency mask, and the right-hand panel for the high-frequency mask.

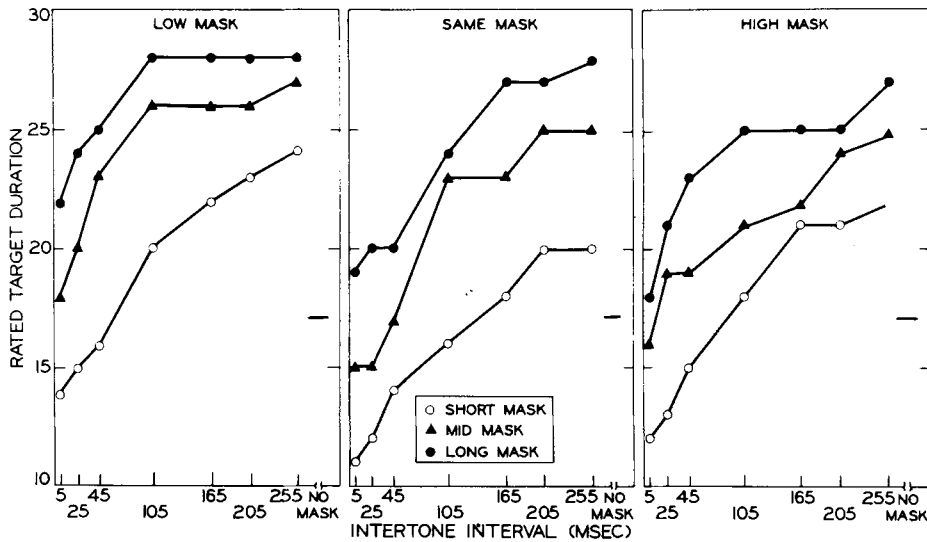


Figure 6. Rated duration of the target as a function of the ITI and the duration of the mask, for each of the three masking tone frequencies. The left-hand panel presents the results for the low-frequency mask, the middle panel for the same-frequency mask, and the right-hand panel for the high-frequency mask.

interaction of ITI, target duration, and mask duration [ $F(14,165) = 1.78, p < .05$ ] were significant in the analysis.

Rated target duration also increased with the ITI: for all three masking tone durations, the ratings of target tone duration increased consistently with the ITI. The target was always rated as shorter on no-mask trials than on masking trials with ITIs greater than 45 msec. The main effects of both the ITI

[ $F(7,84) = 16.47, p < .001$ ] and the duration of the mask [ $F(2,24) = 41.40, p < .001$ ] were significant in the analysis, as was the interaction of these two factors [ $F(14,168) = 11.66, p < .001$ ]. Despite the similarity in the overall pattern of results, the rated duration of the target tone was significantly longer with the low-frequency masking tone. The main effect of masking tone frequency [ $F(2,24) = 8.53, p < .001$ ] was significant in the analysis. Neither



the interaction of masking tone frequency with masking tone duration nor the three-way interaction of frequency, duration, and the ITI were significant.

## DISCUSSION

The current results appear to replicate those of the earlier experiments (Idson & Massaro, 1977). Under all experimental conditions, the long target was rated as having a longer duration than the short target. For both targets, with masks of all durations and frequencies, rated duration increased monotonically with increases in the ITI. Finally, for all three frequency masks, the rated duration of both targets increased consistently with increases in the duration of the mask. The targets were rated as longest when followed by a long mask and shortest when followed by a short mask. These results indicate that judged duration is a function of the actual duration of the target, the duration of the mask, and the amount of available processing time, regardless of the frequency relationship between target and mask. The model would appear to provide a good qualitative description of duration processing when the target and mask differ along multiple dimensions.

A more precise evaluation of the ability of the model to handle the present results can be given by testing Equations 3 and 4 against the obtained data. In order to fit the equations, a linear transformation of the ratings was taken, in the manner described in the introduction. The model was then fit to the resulting values of  $z[P(L | L)]$  and  $z[P(L | S)]$ , using the iterative computer program STEPIT (Chandler, 1969). As in the earlier studies, the 5-msec interval was excluded from the fit. To insure that the rating scores would be transformed in the same way as those obtained in the earlier studies, the values of  $a$  and  $b$  were set to the values obtained in the description of rating study (Idson & Massaro, 1977, Experiment 2).

Two versions of the model were fit to the data in order to assess the contribution of the duration of the masking tone as a function of its frequency. In the simple version, the contribution of the masking tone was assumed to be the same regardless of its frequency similarity to the target tone. In this model, only one value of  $K_S$  and one value of  $K_L$  were estimated across the three frequencies of the masking tone. In the complex version, the contribution of the duration of the masking tone was assumed to be a function of its frequency. Therefore, different values of  $K_S$  and  $K_L$  were estimated for each of the three frequencies of the masking tone. The complex version requires six parameter values to account for the contribution of the mask, whereas the first version requires just two. Both models also require estimates of  $\theta_D$ ,  $\theta_S$ ,  $\theta_L$ ,  $\alpha_S$ , and  $\alpha_L$ , giving a total of

7 parameters for the first version and a total of 11 for the second.

The predicted values were computed by estimating the parameter values to minimize the squared deviations between predicted and observed values. Table 1 presents the parameter values and the goodness-of-fit for each of the two versions of the model. The model provided a very good description of the data in both cases considering the fact that 110 independent data points are being described by only 7 and 11 parameters, respectively. The predictions show, however, that the contribution of the duration of the masking tone cannot be considered to be completely independent of its frequency similarity to the target tone. The parameter values corresponding to the contribution of the mask indicate that the low-frequency masks add more to the judged duration of the target than do the same-frequency or high-frequency masks. The observed and predicted results given by the complex model are given in Table 2.

Idson and Massaro (1977) offered four alternative possibilities for the underlying process responsible for the additive contribution of the mask to the judged duration of the target. The most obvious alternative is a mechanism based on the integration theory of masking. Hawkins and Presson (1977) have articulated such a mechanism to explain target-mask similarity effects in backward recognition masking of pitch judgments. Their formulation of integration theory predicts that the contribution of the mask to judgments about the target would decrease with increases in the interval separating the two tones. This prediction is contradicted by the finding that the observed data are adequately described with the

Table 1  
The Parameter Estimates and Averaged Squared Deviations ( $d^2$ )  
Between the Predicted and Observed Values of the Simple and  
Complex Versions of the Model Given by Equations 3 and 4

Parameter	Simple	Complex
$\theta_D$	15.62	15.61
$\theta_S$	40.47	40.51
$\theta_L$	10.18	10.19
$K_S$	7.45	
Low		9.09
Same		6.25
High		6.98
$K_L$	5.85	
Low		7.38
Same		5.08
High		5.10
$\alpha_S$	2.68	2.68
$\alpha_L$	3.62	3.62
$a$	.075	.075
$b$	1.75	1.75
$d^2$	.020	.013

Table 2  
The Observed and Predicted z Scores for the Long and Short Targets as a Function of ITI,  
Masking Tone Duration, and the Frequency of the Masking Tone

	ITI (msec)	Mask Frequency					
		Low		Same		High	
		Long Target					
Short Mask	25	3.19	3.16	2.95	3.09	2.95	3.09
	45	3.37	3.28	3.04	3.21	3.16	3.21
	105	3.69	3.54	3.33	3.47	3.43	3.47
	165	3.72	3.68	3.58	3.61	3.70	3.61
	205	3.79	3.74	3.70	3.67	3.72	3.67
	255	4.00	3.78	3.46	3.71	3.83	3.71
Mid Mask	25	3.51	3.45	3.17	3.29	3.43	3.29
	45	3.71	3.58	3.30	3.42	3.44	3.42
	105	3.96	3.84	3.75	3.67	3.65	3.68
	165	4.03	3.97	3.76	3.81	3.79	3.82
	205	4.09	4.03	4.05	3.87	3.92	3.87
	255	4.13	4.07	4.01	3.91	3.93	3.91
Long Mask	25	3.86	3.75	3.55	3.49	3.56	3.50
	45	3.76	3.87	3.55	3.62	3.72	3.62
	105	3.99	4.13	3.83	3.88	3.76	3.88
	165	4.17	4.27	4.05	4.02	3.96	4.02
	205	4.13	4.33	4.16	4.07	3.92	4.07
	255	4.33	4.37	4.21	4.12	4.17	4.12
No Mask = 3.39, 3.62							
		Short Target					
Short Mask	25	2.58	2.51	2.32	2.42	2.50	2.44
	45	2.59	2.76	2.56	2.67	2.53	2.69
	105	2.85	2.94	2.62	2.85	2.79	2.87
	165	3.09	2.95	2.69	2.87	2.96	2.89
	205	3.11	2.95	2.76	2.87	2.87	2.89
	255	3.12	2.95	3.04	2.87	2.97	2.89
Mid Mask	25	3.05	2.87	2.65	2.67	2.95	2.72
	45	3.25	3.12	2.72	2.92	2.98	2.97
	105	3.42	3.30	3.18	3.10	3.07	3.15
	165	3.37	3.32	3.23	3.12	3.18	3.17
	205	3.39	3.32	3.20	3.12	3.22	3.17
	255	3.37	3.32	3.19	3.12	3.36	3.17
Long Mask	25	3.27	3.24	2.96	2.92	3.16	3.00
	45	3.48	3.48	3.01	3.17	3.19	3.25
	105	3.68	3.66	3.29	3.35	3.43	3.43
	165	3.60	3.68	3.45	3.37	3.34	3.45
	205	3.56	3.68	3.42	3.37	3.37	3.45
	255	3.37	3.68	3.47	3.37	3.36	3.45
No Mask = 2.58, 2.68							

Note—The first entry is the observed score and the second is the predicted score.

assumption that the contribution of the mask is independent of the intertone interval between the target and masking tones (Idson & Massaro, 1977).

The second possibility is that the mask influences the decision criterion used by the observer in the task. Upon hearing the masking tone, subjects may be biased to evaluate the duration of the target tone in the same direction as that of the perceived duration of the mask. This alternative seems somewhat unlikely, given that identical results are observed in both binary choice and continuous rating experiments. The rating experiment allows an observer to provide a direct index of judged duration and there is no

reason to attribute the differences in ratings to purely decision processes.

The third explanation is that the observer confuses the temporal order of the target and mask and inadvertently rates the mask duration on some of the trials. This explanation would also predict a decreasing contribution of the mask with an increasing intertone interval, since temporal order judgments improve dramatically with small increases in relative onsets between two events. Analogous to an integration explanation, a temporal confusion hypothesis also has difficulty with the observation of a fixed mask contribution with large changes in the intertone

interval separating the target and the mask.

The fourth explanation is based on the idea that the mask has two independent effects on processing target duration. The mask not only terminates primary recognition of the target tone, it can also modify the representation of the target tone in synthesized auditory memory. The primary evidence for this interpretation is the strictly additive effects of intertone interval and mask duration in the perceived duration experiments. The idea is that the representation of the target in synthesized auditory memory can be modified by the perceived duration of the mask even though the duration of the mask is not critical for how effective it is in terminating primary recognition of the target tone. The fact that the low mask produced more of an increase in judged duration than did the high and same mask might be explained in terms of the low mask producing less interference in synthesized auditory memory. This explanation is consistent with other findings that masks lower in frequency than the target produce less interference in backward recognition masking (Watson, Wroton, Kelly, & Benbassat, 1975).

The strictly additive effects of target processing time and mask duration are consistent with recent experiments on similarity effects in backward recognition masking (Kallman & Massaro, Note 1). Having found significant test-mask similarity effects in backward recognition masking of pitch judgments, the goal was to isolate the stage of processing responsible for the similarity effects. Their results indicated that the decrease in backward recognition masking with decreasing test-mask similarity was not due to the mask's effectiveness in terminating the readout of preperceptual auditory store. Instead, the similarity effects appear to result from differences in interference with the synthesized auditory memory of the test tone. Therefore, the contribution of mask duration to judged target duration and test-mask similarity effects are *not* inconsistent with the idea that the mask terminates primary recognition of the target tone. These effects can be attributed to processes occurring at the later stage of synthesized auditory memory.

## CONCLUSION

The model given by Idson and Massaro (1977) provided a good description of the temporal course of the perceived duration of the target tone. Moreover, masking tones of both the same and of different frequencies than the target tones appear to have similar effects upon the perceived duration of the target. These results suggest that the model is not limited to the case in which the target and mask differ only in duration. Rather, the model is capable of describing the dynamics of duration perception

under the more natural conditions in which temporally successive items differ along multiple dimensions.

The utility of the model is further supported by a recent study in the visual modality (Bernstein, Fisicaro, & Fox, 1976). Bernstein et al. independently varied the target and mask energies of a metacontrast display. In addition to the typical finding that the accuracy of performance was a U-shaped function of the SOA, they found that judgments of brightness were influenced by the energy of the mask. At SOAs longer than the period of integration, brightness judgments were positively correlated with the energy of the mask. These results parallel the present findings that duration judgments are positively correlated with the duration of the mask. Bernstein et al.'s analysis also suggested that the target and the mask were first processed separately, with the information from each being combined for the classification judgment. This is in accord with the present model, which assumes that the target and mask are processed separately and then combined additively, as shown in Equation 2. The strength of the model is further increased by its ability to describe analogous results in auditory and visual information processing.

## REFERENCE NOTE

1. Kallman, H. J., & Massaro, D. W. *Similarity effects in backward recognition masking*. Wisconsin Human Information Processing Program, WHIPP, Report No. 4, 1978.

## REFERENCES

- ALLAN, L. G. Is there a constant minimum perceptual duration? *Quarterly Journal of Experimental Psychology*, 1976, **28**, 71-76.
- ALLAN, L. G., & KRISTOFFERSON, A. B. Psychophysical theories of duration discrimination. *Perception & Psychophysics*, 1974, **16**, 26-34.
- ALLAN, L. G., & ROUSSEAU, R. Backward masking in judgments of duration. *Perception & Psychophysics*, 1977, **21**, 482-486.
- BERNSTEIN, I. H., FISICARO, S. A., & FOX, J. A. Metacontrast suppression and criterion content: A discriminant function analysis. *Perception & Psychophysics*, 1976, **20**, 198-204.
- CANTOR, N. E., & THOMAS, E. A. C. Visual masking effects on duration, size, and form discrimination. *Perception & Psychophysics*, 1976, **19**, 321-327.
- CHANDLER, J. P. Subroutine STEPIT - finds local minima of a smooth function of several parameters. *Behavioral Science*, 1969, **14**, 81-82.
- EFRON, R. Effects of stimulus duration on perceptual onset and offset latencies. *Perception & Psychophysics*, 1970, **8**, 231-234. (a)
- EFRON, R. The measurement of perceptual duration. *Stadium Generale*, 1970, **23**, 550-561. (b)
- EFRON, R. The minimum duration of a perception. *Neuropsychologia*, 1970, **8**, 57-63. (c)
- GOL'DBURT, S. N. Investigation of the stability of auditory processes in micro-intervals of time (new findings in back masking). *Biophysics*, 1961, **6**, 809-817.
- HAWKINS, H. L., THOMAS, G., PRESSON, J., COZIC, A., & BROOKMIRE, D. Tonal specificity and masking in auditory recognition. *Journal of Experimental Psychology*, 1974, **103**, 530-538.

- HAWKINS, H. L., & PRESSON, J. C. Masking and preperceptual selectivity in auditory recognition. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, N.J.: Erlbaum, 1977.
- IDSON, W. L., & MASSARO, D. W. Perceptual processing and experience of auditory duration. *Sensory Processes*, 1977, **1**, 316-337.
- KRISTOFFERSON, A. B. A real-time criterion theory of duration discrimination. *Perception & Psychophysics*, 1977, **21**, 105-117.
- MARKS, L. E. *Sensory process: The new psychophysics*. New York: Academic Press, 1974.
- MASSARO, D. W. Preperceptual auditory images. *Journal of Experimental Psychology*, 1970, **85**, 411-417.
- MASSARO, D. W. Stimulus information versus processing time in auditory pattern recognition. *Perception & Psychophysics*, 1972, **12**, 55-56.
- MASSARO, D. W. *Experimental psychology and information processing*. Chicago: Rand McNally, 1975.
- MASSARO, D. W., COHEN, M. C., & IDSON, W. L. Recognition masking of auditory lateralization and pitch judgments. *Journal of the Acoustical Society of America*, 1976, **59**, 434-441.
- MASSARO, D. W., & IDSON, W. L. Temporal course of perceived auditory duration. *Perception & Psychophysics*, 1976, **20**, 331-352.
- MASSARO, D. W., & IDSON, W. L. The temporal course of perceived vowel duration. *Journal of Speech and Hearing Research*, 1978, **21**, 37-55.
- MOORE, J. J., & MASSARO, D. W. Attention and processing capacity in auditory recognition. *Journal of Experimental Psychology*, 1973, **99**, 49-54.
- THOMAS, E. A. C., & CANTOR, N. E. On the duality of simultaneous time and size perception. *Perception & Psychophysics*, 1975, **18**, 44-48.
- THOMAS, E. A. C., & CANTOR, N. E. Simultaneous time and size perception. *Perception & Psychophysics*, 1976, **19**, 353-360.
- WATSON, C. S., WROTON, H. W., KELLY, W. J., & BENBASSAT, C. A. Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals. *Journal of the Acoustical Society of America*, 1975, **57**, 1175-1185.

## NOTE

1. In two more recent papers (Idson & Massaro, 1977; Massaro & Idson, 1978), the predictions of the revised model were expressed as

$$JD_S = \alpha_S [1 - e^{-(\theta_D D + \theta_S t)}] + K_S(t_m) \quad (7)$$

and

$$JD_L = \alpha_L [1 - e^{-(\theta_D D + \theta_L t)}] + K_L(t_m). \quad (8)$$

Although Equations 7 and 8 are mathematically identical to Equations 3 and 4, respectively, Equations 3 and 4 were chosen in the present formulation because they have the advantage of making the psychological assumptions of the model more transparent.

(Received for publication March 10, 1978;  
revision accepted May 11, 1978.)