SENSORY PROCESSES 1, 316–337 (1977)

Perceptual Processing and Experience of Auditory Duration[1]

WENDY L. IDSON AND DOMINIC W. MASSARO
University of Wisconsin

Received January 21, 1977

This study explored changes in the apparent duration of a pure tone over time. Previous research employing a backward-masking, recognition paradigm suggested that the perception of duration increases with actual target duration, the time available for processing the target, and the duration of the mask. The present study attempted to explore more directly these changes in apparent duration. A backward-masking paradigm was again used, in which one of two targets of different duration was presented on each trial, followed after a variable, silent interval by a masking tone, which could assume one of three possible durations. The subject's task was to estimate the duration of the target tone along a continuous scale of perceived duration. The estimates of perceived duration increased directly with the durations of the target, the intertone interval, and the mask. The results provide strong quantitative support for a model developed in previous research using accuracy measures. The model submits that the percept of target duration increases over the duration of the target and the following silent interval. The masking tone terminates the continuous growth of perceived duration of the target tone. The model also assumes that the mask adds a constant, proportional to its duration, to the perceived duration of the target.

One of the fundamental problems in auditory perception concerns the relationship between the physical and psychological attributes of sound. In general, a close correspondence has been found between these dimensions of an auditory stimulus. For example, the perceived pitch of a tone depends primarily on its frequency, subject to the influence of factors such as intensity. Yet recent work in both audition (Efron, 1970a,b,c; Gol'dburt, 1961; Massaro & Idson, 1976, in press) and vision (Cantor & Thomas, 1976; Thomas & Cantor, 1975, 1976) suggests that duration may differ in this respect from other psychological attributes. The perceived duration of a stimulus is a function not only of its physical characteristics, but also of the amount of time spent processing it. A given stimulus, having constant physical characteristics, will have a variety of perceived durations, dependent upon the amount of available processing time.

In audition, two sets of studies indicate the validity of the argument that perceived duration may vary over time. The first set of studies was designed to

[1] This research was supported by U. S. Public Health Service Grant No. MH-19339 to D. W. M. and was conducted while W. L. I. was supported by both a University of Wisconsin Graduate Traineeship and Fellowship. The work represents a collaborative effort; order of authorship is arbitrary. We would like to thank Lola L. Lopes and Gregg C. Oden for valuable discussions. Michael M. Cohen wrote the computer program to record the responses. A. B. Kristofferson, L. E. Marks, and two unknown referees provided extremely helpful comments on an earlier draft of the manuscript. Wendy L. Idson is now at the Department of Psychology, University of Texas, Austin, Texas 78712. Requests for reprints should be sent to Dominic W. Massaro, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

discover how long a percept lasts (Efron, 1970a,b,c). Efron presented subjects with a short light and sound on each trial. The subjects' task was to judge when the onset of the light occurred simultaneously with the onset of the sound. They were then asked to judge when the onset of the light occurred at the offset of the sound. The point at which the onset of the light and the onset (offset) of the sound appeared to be simultaneous was taken as the subjective onset (offset) point of the sound, allowing computation of perceived duration. The principal finding of these studies was that, while accurate perception of duration occurred for stimuli longer than 130 msec, for sounds less than this length the actual duration was overestimated to produce a perceived duration of at least 130 msec. The perceptual offset latencies of stimuli having physical durations shorter than this critical value were found to increase with decreases in physical duration, yielding a minimal perceptual duration (cf., Allan, 1976, for a critical review of this concept). Efron interpreted these results as indicating that the percept of a short tone persists after its physical offset.

A more directly relevant study was conducted by Gol'dburt (1961), who found that perceived duration varies directly with processing time. Gol'dburt (1961) conducted a backward-masking experiment, varying both the duration of the target and the intertone interval between target and mask. There were three related findings of interest. First, Gol'dburt found that the presence of the mask decreased the perceived duration of the target. Second, this effect was greatest with short targets and decreased as target duration was lengthened. Third, the effect was found to decrease with increases in the time between the target and the mask.

Efron's (1970a,b,c) and Gol'dburt's (1961) findings suggest that perceived duration varies consistently with certain temporal parameters. Massaro and Idson (1976) attempted to explore this suggestion with a paradigm which is explicitly designed to tap changes in the perception of a stimulus over time. They used a backward-masking recognition paradigm, in which one of two target tones of differing durations was presented on each trial, followed after a variable, silent interval by a masking tone of one of three durations. The usual finding in backward-masking recognition studies, when the subject is asked to identify pitch (Hawkins, Thomas, Presson, Cozic, & Brookmire, 1974; Massaro, 1970), timbre (Massaro, 1972b), loudness (Moore & Massaro, 1973), or sound lateralization (Massaro, Cohen, & Idson, 1976), is that performance improves as a monotonic, negatively accelerated function of the intertone interval, asymptoting at a level comparable to that found when a masking tone is not presented. In Massaro and Idson's studies, the subject was asked to identify the target tone as being long or short. Under these conditions, performance averaged over the two targets produced results quite comparable to those found for other perceptual attributes. The top panel of Fig. 1 presents the average percentage of correct test-tone identifications, as a function of intertone interval and masking-tone duration. Average performance improved as a negatively accelerated function of the intertone interval. This result was obtained for all three masking-tone durations. Thus, the average masking function for duration reflects similar processes to those occurring for other stimulus dimensions. With other attributes, the alternative targets yield essentially equivalent performance. In contrast, markedly different results were obtained for identification of the short and long target tones. While identification of the long target conformed to previous results in backward masking, increasing with increases in the silent intertone interval, identification of the short target was better at short than at long intervals. Moreover, contrary to the typical findings in these studies, though identification of the short target was quite good on no-mask trials, identification of the long target was quite poor. The bottom panel of Fig. 1 presents the percentage of correct identifications of each of the target tones as a function of the intertone interval. Performance was initially equal on the two targets at the 5-msec interval. At intervals of 25 and 45 msec, the short target was identified far more accurately than the long target, while at longer intervals performance on the short target fell to a level below that found for the long target. These results are obviously at variance with those typically found in backward-masking recognition studies.

To account for these results, Massaro and Idson (1976) proposed an extension of a more general model of auditory recognition (Massaro, 1972a, 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 auditory recognition process, information in this store is read out continuously over the course of approximately 250

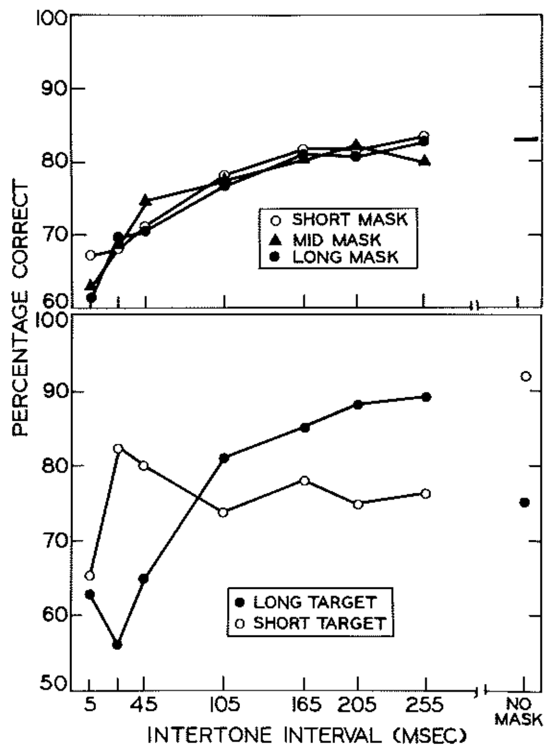

FIG. 1. Top panel: Percentage of correct responses to each of the target tones as a function of the duration of the silent ITI; bottom panel: percentage of correct identifications of the target, as a function of both duration of the mask and ITI (from Massaro & Idson, 1976).

msec. 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 backward-masking recognition tasks with increases in the 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.

The perception of duration is assumed to occur in essentially the same manner. However, the added assumption is made that the perceived duration of the target increases with the silent processing time after target presentation. That is, both discriminability and perceived duration are assumed to increase with additional processing time. If sufficient time is allowed for complete processing of the target, its perceived duration will be directly related to its 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 function of processing time. 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 ensures that with long intertone intervals the perceived duration of the target will approach its asymptotic duration. With short intertone intervals both the long and short targets will have relatively short perceived durations, while with long intertone intervals both targets will have relatively long perceived durations. Therefore, with 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 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 which were obtained (cf., Fig. 1). The model also assumes that the mask adds a component, proportional to its duration, to the perceived duration of the target. This assumption was supported by the strong target-by-mask interaction found in these studies. To the extent that the target and mask were both short (long), performance was better than if the target was short (long) and the mask was long (short). This model was found to be capable of describing quite well the data obtained for pure tones (Massaro & Idson, 1976) and for both long (180/240 msec) and short (60/80 msec) vowels (Massaro & Idson, in press).

Similar results have also been found in a visual-masking paradigm. Cantor and Thomas (1976) presented a target stimulus for 30 or 50 msec, followed after an interstimulus interval (ISI) of 0, 30, 70, or 110 msec by a 500-msec presentation of a masking field. The target stimuli were either two circles of differing sizes or two different abstract forms. In separate sessions, the subjects were asked to judge the duration, size (circles), or form (abstract forms) of the target stimuli. Accuracy of discrimination of size and form increased monotonically with increases in both stimulus duration and ISI. In addition, the perceived duration of the form stimuli increased monotonically with increases in both stimulus duration and ISI. These results are largely predictable from the model presented above. The monotonic increase in accuracy with increases in stimulus duration and ISI and in perceived duration of the form stimuli with increases in processing time replicate Massaro and Idson's (1976, in press) results and can be interpreted in a similar manner. It should be noted, however, that Thomas and Cantor (1976; Cantor & Thomas, 1976) have developed a somewhat different model to account for these results. In many respects, their model is comparable to that presented here; they also assume that perceived duration increases with increases in processing time. The key difference is that in Thomas and Cantor's model the duration percept is assumed to be derived from the time spent processing the nontemporal information in the stimulus. Because processing time for nontemporal information will increase with increases in the stimulus-onset-asynchrony, and perceived duration is assumed to be directly related to processing time, Cantor and Thomas's model can also predict all of the obtained results. Though it should be possible to distinguish between the two models empirically, their overall similarity reinforces the idea of analogous dynamic mechanisms in visual and auditory information processing.

Taken together, these results (Cantor & Thomas, 1976; Efron, 1970a,b,c; Gol'dburt, 1961; Massaro & Idson, 1976, in press; Thomas & Cantor, 1975, 1976) suggest an intriguing difference between duration and other stimulus attributes. When overall discriminability is considered, the results for duration do not differ significantly from those obtained for attributes such as pitch. This suggests that duration can be handled within the general context of Massaro's (1975) model of auditory recognition. Yet when identification of the individual targets is considered, the results for duration are quite different. These results are well described by Massaro and Idson's (1976) model, assuming that the duration percept grows over time. In contrast to other stimulus attributes, perceived duration appears to be function not only of the parameters of the physical stimulus, but also of the processing time available and the duration of a second stimulus that is presented before the first is completely processed.

In order to evaluate this hypothesized difference between duration and other stimulus attributes, it is necessary to consider not only the accuracy of the subject's response, but also the perceptual experience leading to that response. The previous work measured only accuracy and inferred changes in the duration percept from variations in the discriminability of the two targets. The intent of the current research is to provide more explicit evidence on the nature of the perceptual experience of duration, by employing a dependent measure which should directly tap the subjective durations of the stimuli. In order to do this, we used a backward-masking task such as that employed in the earlier studies. Each trial presented one of two target tones differing in duration, followed after a variable, silent interval by a masking tone, which could assume any of three possible durations. However, rather than categorizing the target as long or short, the subjects were asked to estimate the duration of the target along a continuous scale from short to long. The use of such a scale has the important advantage of allowing the subject to respond directly in terms of his or her perceptual experience, rather than in some sense mapping a continuum of perceived durations into a binary choice.

If the model is valid, certain definite patterns should appear in these ratings. Because the objective duration of the tone contributes to its perceived duration, the long target should be rated as longer than the short target at each experimental condition. The perceived duration of the target is also assumed to increase with increased processing time. In this case, the rated durations of both targets should increase as the intertone interval is lengthened. Third, if the duration of the mask does in fact contribute a component to the perceived duration of the target, then perceived duration should increase with increases in the duration of the masking tone. Consequently, having subjects directly estimate perceived duration should allow a direct evaluation of the model.

The predicted results for duration are quite different from those expected for other attributes. Consider, for example, a study in which the subject is asked to rate the perceived pitch of the target in a backward-masking task. Discriminability between the two targets is assumed to increase with the intertone interval (ITI), but no overall pitch change should occur. As the ITI is increased, the high target should be consistently rated as higher and the low target as lower, the results which are obtained (Idson & Massaro, Note 1). This contrasts markedly to the predictions for duration, where both targets should come to be rated as longer when the ITI is increased.

It should be noted that, while the use of a rating scale has often been questioned, its validity has been repeatedly established in work in information-integration theory and functional measurement (Anderson, 1970, 1974). As Anderson argues, if a quantitative model can be found which provides a good fit to the data, the results can be taken to provide a joint validation of the model and the response scale. If the model fits the results, the most likely alternative is that the model describes how the observer makes his or her judgments and that the response scale was used appropriately. A less likely alternative is that the model is wrong, but the scale was biased in the precise manner needed to compensate for the inadequacies of the model. In the present context, the likelihood of this latter alternative would be quite low, as the model has been previously supported (Massaro & Idson, 1976, in press) using an entirely different dependent measure. Consequently, a reasonable degree of confidence can be placed in the rating scale, which will in turn allow a finer level of analysis of the perceptual experience of duration.

Accordingly, the present experiments replicate Massaro and Idson's (1976) stimulus conditions, but ask for a direct estimate of perceived tone duration. In order to replicate the stimulus conditions of the earlier studies exactly, Experiment 1 employs target tones of 60 and 80 msec and masking tones of 50, 70, and 90 msec, whereas Experiment 2 employs durations of 50 and 90 msec and of 30, 70 and 110 msec.

EXPERIMENT 1

Method

Subjects. The subjects were eight University of Wisconsin undergraduates who were paid $2.00/hr for participating an hour a day for 5 consecutive days.

Apparatus and stimuli. Four subjects could be tested simultaneously in separate sound-insulated booths. All experimental events were controlled by a PDP-8L computer. The tonal stimuli 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 responses were recorded on a continuous response scale, 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 pressed the pushbutton to indicate that a response had been made. The placement of the pointer was indicated by the voltage passed by the potentiometer to an analog-to-digital converter. The voltage given by the linear potentiometer allows the 5.5-cm scale to be mapped into 50 intervals from left to right, which were assigned the integer values from 0 to 49.

Two target tones were employed in the study, having durations of 60 msec (short) and 80 (long) msec. Three masking tone durations, symmetrical around the durations of the target tones, were used: 50 (short mask), 70 (mid mask), and 90 msec (long mask). All tones had frequencies of 700 Hz. In order to avoid having perceived loudness function as a cue to duration, the intensities of the target and masking tones were sampled randomly and independently on each trial from a set of seven possible intensities: 75 to 81 db SPL, in 1-db steps. All tones were turned on and off at the zero crossing and had essentially instantaneous rise times.

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. Each session consisted of 320 test trials. All of Day 1 and the first 20 trials of the eight subsequent experimental sessions were treated as practice, though the subjects were not informed of this. On seven out of eight of the trials one of the two target tones was followed after a variable, silent intertone interval (ITI) of 5, 25, 45, 105, 165, 205, or 255 msec, by one of the three possible masking tones. On one out of eight of the trials, no masking tone was presented. All 48 experimental conditions (2 target durations × 3 mask durations × 7 ITIs plus the no mask) were completely random, with the constraint that each condition would occur once in every 48 trials. (The masking tone's duration was a dummy variable under the no-mask condition.)

On each trial, the subject's task was to rate the perceived duration of the target tone, ignoring the masking tone. It should be noted that the subject could begin making the rating responses as soon as the target was presented, but that the reaction time always exceeded the longest ITI. The subject responded by placing the pointer at a location along the response scale, labeled short on the left end and long on the right end. Though only two target tones occurred in the experiment, the subjects were instructed that they would hear a range of target-tone durations. They were asked to map the range of durations which they heard onto the response scale, utilizing the entire length of the scale. Subjects were instructed that the longest tone they would hear would represent the right-most endpoint of the scale, and the shortest tone would represent the left-most endpoint. Subjects were also instructed to attempt to have equal changes in duration represented by equal distances along the scale. In a postexperiment questionnaire, no subject reported difficulty in using the scale in the required manner. The subjects were given 3 sec from the onset of the target in which to make their responses. Following the response interval, a 250-msec display of an asterisk was presented over a visual display of light-emitting diodes (Monsanto Model MDA-III) to indicate that the response interval was over and a new trial was about to begin. The intertrial interval was 1 sec.

Results

The ratings of target duration, indicated by placement of the pointer along the response scale, were recorded as integers between 0 and 49, with 0 representing the shortest and 49 the longest possible tone. The mean ratings of duration were computed for each subject on each day at each target duration by mask duration by ITI condition. These ratings were then submitted to two separate analyses of variance on subjects, days, targets durations, mask durations, and ITIs. In one analysis, the no-mask condition was treated as an ITI of ∞ and entered the analysis as a level of the ITI factor. In the other analysis, the no-mask was excluded and the ITI factor consisted of only the seven levels corresponding to the seven ITIs. This latter analysis was performed so as to allow an evaluation of the effects of the duration of the masking tone independent of the results on the no-mask trials, for which the duration of the mask was a dummy variable. All of the results given below which involve mask duration as a factor are drawn from this second analysis. All other results are taken from the analysis in which the no-mask was included.

In both analyses, the main effect of days and all interactions involving days as a factor were nonsignificant, indicating that performance did not vary appreciably across the 4 experimental days. Consequently, all of the data will be presented in terms of the average results across the 4 experimental days.

The top panel of Fig. 2 presents the rated durations of each of the two target tones as a function of the ITI. As the figure shows, the long target was consistently rated as being longer in duration than the short target, $F(1, 7) = 15.40$, $p < .01$. For both target tones, rated duration increased as a negatively accelerated function of ITI, between intervals of 25 and 255 msec, $F(7, 49) = 3.39$, $p < .01$. This result indicates that the perceived duration of the target was increasing over the time available for processing that target. However, much longer ratings were observed for both targets at the 5-msec ITI than at any other interval. This result is consistent with the hypothesis that at extremely short ITIs the target and mask are integrated into a single composite sound (Massaro, 1975; Massaro *et al.*, 1976). The integration of the target and mask into a composite at the 5-msec ITI would produce a perceived duration which is longer at this interval than that of the target alone at any other ITI (Massaro & Idson, in press). On no-mask trials, the long target was again rated as longer than the short target. However, both targets had shorter rated durations on no-mask trials than at any ITI longer than 45 msec. Though the functions across the ITI were similar for both targets, the target duration-by-ITI interaction was significant, $F(7, 49) = 3.85, p < .005$.

The bottom panel of Fig. 2 presents the average rated target duration as a function of both ITI and duration of the mask. For all three mask durations, rated target duration increased as a monotonic function of the ITI between intervals of

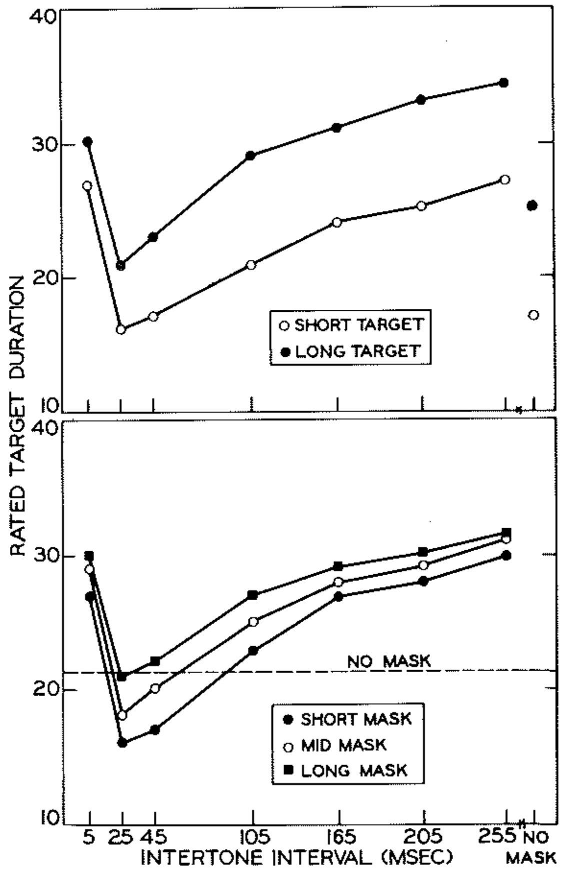

FIG. 2. The top panel presents the rated durations for each of the two target tones, as a function of the duration of the silent ITI. The bottom panel presents the rated target durations, as a function of both duration of the mask and ITI (Experiment 1).

25 and 255 msec. The targets show the same increased ratings at the 5-msec ITI for all three mask durations. Although similar functions were found for all three masking tone durations, the target was rated as longest when followed by the long mask and shortest when followed by the short mask, at all ITIs. Both the main effect of the duration of the mask, $F(2, 24) = 17.92, p < .001$, and the interaction of mask-by-ITI, $F(14, 98) = 2.99, p < .05$, were significant.

The overall interaction of target-duration by mask duration can be seen clearly in Fig. 3, which presents the rated durations of each of the two targets as a function of the duration of the mask. The duration of the mask had a large effect upon rated target duration. For both the long and the short targets, the target tone was rated as shortest when followed by a short mask and longest when followed by a long mask. This effect was greater for the short than for the long target. For the short target, rated duration increased from 20.92 with the short mask to 24.62 with the long mask, while for the long target the increase was only from 28.00 to

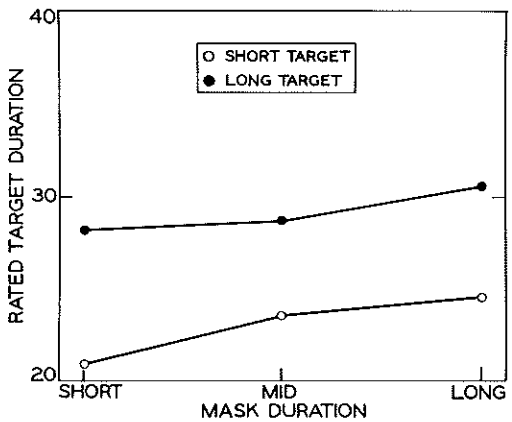

FIG. 3. Rated target duration for each of the two target tones, as a function of duration of the mask. The no-mask trials have been eliminated (Experiment 1).

30.71. This is confirmed by the significant interaction found between target duration and mask duration, $F(2, 14) = 5.92, p < .025$. The three-way interaction of target duration, mask duration, and ITI was nonsignificant in the analysis.

Figure 4 presents the rated durations of each of the two target tones as a function of ITI, for two representative subjects (TI and RA). It is apparent from the figure that the group data accurately reflect the performance of the individual subjects.

Discussion

The results of Experiment 1 strongly support the proposed model. The model assumes that the percept of target duration grows over the time from the onset of the target. As a result, both the actual duration of the target and the length of the ITI should influence the subjective duration of the target. In accord with this, the long target was rated as having a longer perceived duration than the short target under each experimental condition. For both targets, rated duration increased monotonically between ITIs of 25 and 255 msec, confirming the assumption that perceived duration grows over time and can be terminated by the onset of the mask. The high ratings of target duration at the 5-msec interval conform to the hypothesis (Massaro, 1975; Massaro *et al.*, 1976; Massaro & Idson, 1976) that at intervals this short the target and the mask are integrated into a composite sound for which processing time is essentially unlimited. The second assumption of the model is that the mask adds a constant, proportional to its duration, to the perceived duration of the target. As would be expected from such an assumption, both targets were rated as shortest when followed by the short mask and longest when followed by the long mask. Moreover, on no-mask trials the contribution of the mask is absent, and the targets were rated as shorter than they were on masking trials with long ITIs. The target on no-mask trials was rated as longer than the same target on masking trials with short ITIs. These results reflect the independent contributions of processing time and presence of a masking tone to the perceived duration of the target (cf. Massaro & Idson, 1976, Experiment 5, for

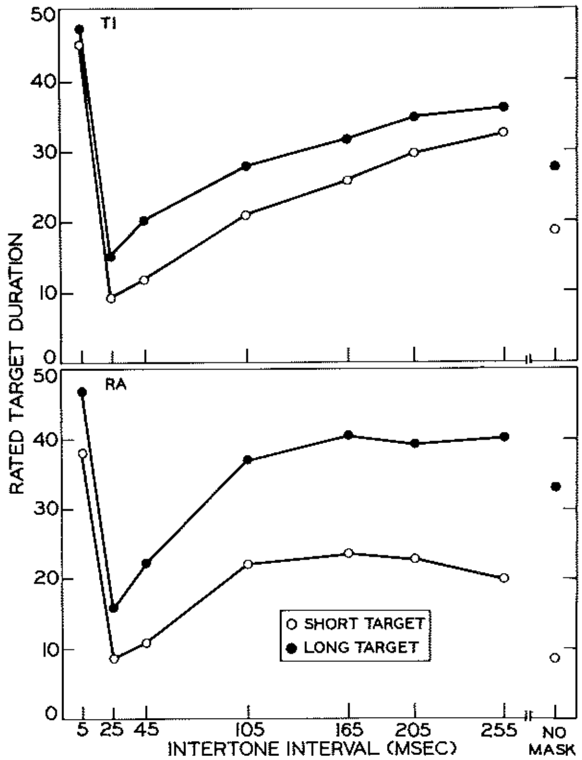

FIG. 4. Rated duration of each of the two targets as a function of ITI, for two individual subjects (Experiment 1).

a detailed discussion of the relationship between no-mask trials and trials on which a masking tone is presented).

EXPERIMENT 2

Method

Subjects. The subjects were nine University of Wisconsin undergraduates who were paid $10.00 for participating in the 5-hr experiment.

Procedure. The procedure was identical to that of Experiment 1, with two exceptions. First, the durations of the short and the long target tones were changed to 50 and 90 msec, respectively. Second, the durations of the masks were changed to 30, 70, and 110 msec, for the short, mid, and long masks, respectively.

Results

The mean ratings of target duration were computed for each subject on each day at each target duration, mask duration, and ITI condition. As in Experiment 1, these ratings were submitted to two separate analyses of variance on subjects, days, target durations, mask durations, and ITIs.

The main effect of days was nonsignificant in both analyses, as were all interactions involving days as a term. Consequently, all of the data will be presented in terms of the average results for the 4 experimental days.

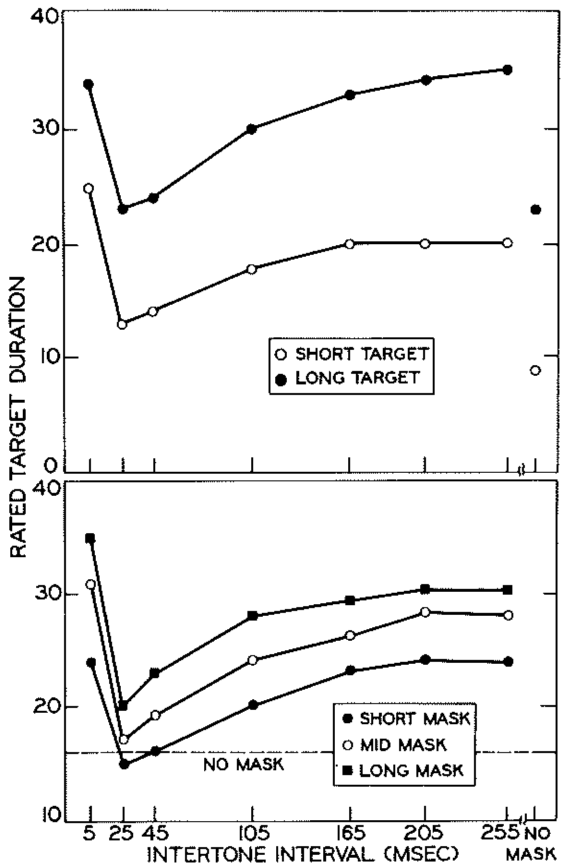

FIG. 5. Top panel: rated durations for each of the two target tones as a function of duration of the silent ITI; bottom panel: rated target durations as a function of both duration of the mask and ITI (Experiment 2).

Figure 5 presents the rated durations of each of the two target tones, as a function of ITI. As the figure shows, the long target was rated as longer than the short target, at all ITIs. As in Experiment 1, rated target duration increased for both targets as a negatively accelerated function of ITI, between intervals of 25 and 255 msec. At a 5-msec ITI, much longer ratings were again observed than those seen at any other masking intervals. The rated duration of the long target increased more than did that of the short target. In addition, the rated duration of the long target increased between ITIs of 165 and 255 msec, while the rated duration of the short target asymptoted at an interval of 165 msec. In contrast to Experiment 1, the rated durations of both targets were shorter on no-mask trials than at any ITI on masking trials. Supporting these conclusions, the analysis revelaed significant effects of target duration, $F(1, 18) = 47.20, p < .001$; ITI, $F(7, 56) = 4.53, p < .001$, and the interaction of these factors, $F(7,56) = 2.52, p < .05$.

The bottom panel of Fig. 5 presents the average rated target duration as a function of both duration of the mask and ITI. Replicating the results of

Experiment 1, the three masking-tone durations yielded qualitatively similar functions across the values of ITI. Rated target duration increased monotonically with ITI between intervals of 25 and 255 msec, with the longest ratings occurring at the 5-msec interval. Once again, rated target duration increased with the duration of the mask at all ITIs. The target was rated as having the longest duration when followed by the long mask and the shortest duration when followed by the short mask. These findings were confirmed by the significant main effect of the masking-tone duration, $F(2, 16) = 28.05$, $p < .001$, and the nonsignificant interaction of mask by ITI found in the analysis.

Figure 6 presents the rated durations of each of the two targets, as a function of the duration of the mask. As the figure shows, rated target duration increased with mask duration for both the long and the short targets. This effect was again somewhat smaller for the long than for the short target tone.

Figure 7 presents the rated durations of the two targets as a function of the ITI, for two individual subjects. The data for these two subjects (NJ and JJ) are typical of those obtained for seven of the nine subjects tested. The two remaining subjects (CS and AN), whose data are presented individually in Fig. 8, showed quite different results. For subject CS, the duration of the short target increased as rapidly as that of the long target. Moreover, this subject rated both targets on no-mask trials as having the shortest possible duration, yielding ratings on the no-mask trials which were substantially shorter than those given on any masking condition. This contrasts with the other subjects, whose ratings of target duration on no-mask trials fell within the same range as their ratings on masking trials. It seems probable that the difference between Experiments 1 and 2 on the no-mask trials reflects the contribution of the data from Subject CS. With her data eliminated, the rated durations on the no-mask trials were longer than the ratings for the same targets on masking trials at ITIs of 25 or 45 msec, though shorter than the ratings at any other interval. These results replicate those of Experiment 1 directly. Subject AN yielded results which are even more deviant with respect to

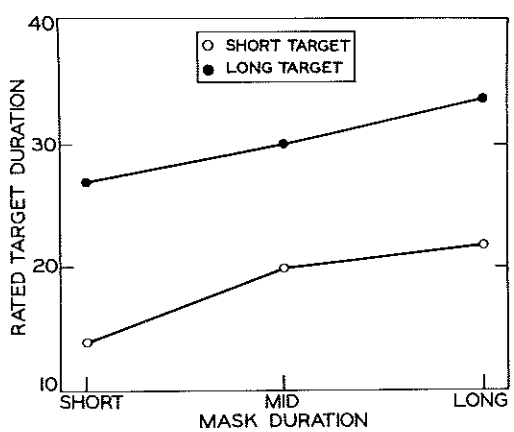

FIG. 6. Rated target duration for each of the two targets as a function of duration of the mask. The no-mask trials have been eliminated (Experiment 2).

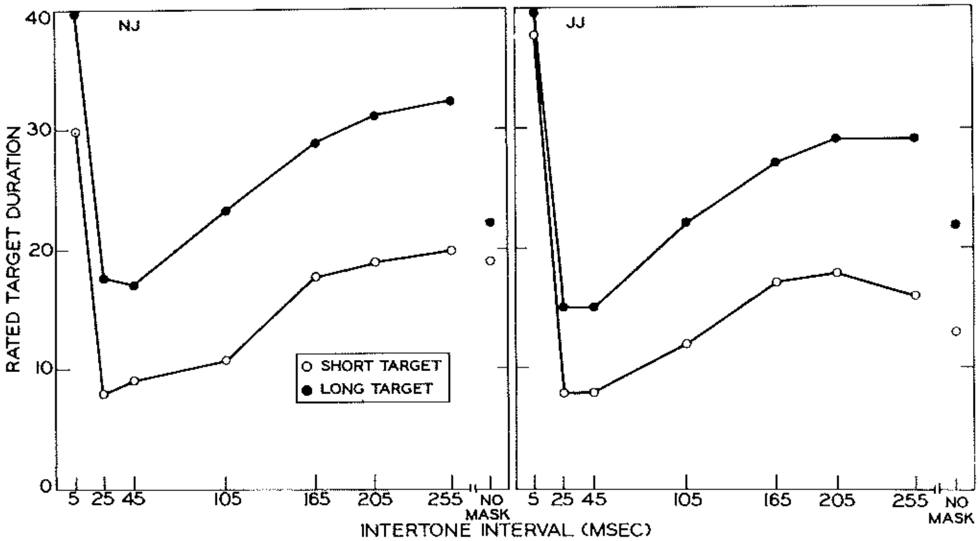

FIG. 7. Rated duration of each of the two targets as a function of ITI, for two typical subjects (Experiment 2).

the group data. Her ratings of the duration of the long target increased only mininally with ITI and actually decreased across intervals for the short target. No immediate explanation for the results of subjects CS and AN suggests itself. However, the great similarity observed in the data of the other seven subjects indicates that their group data do, in fact, represent a reliable phenomena.

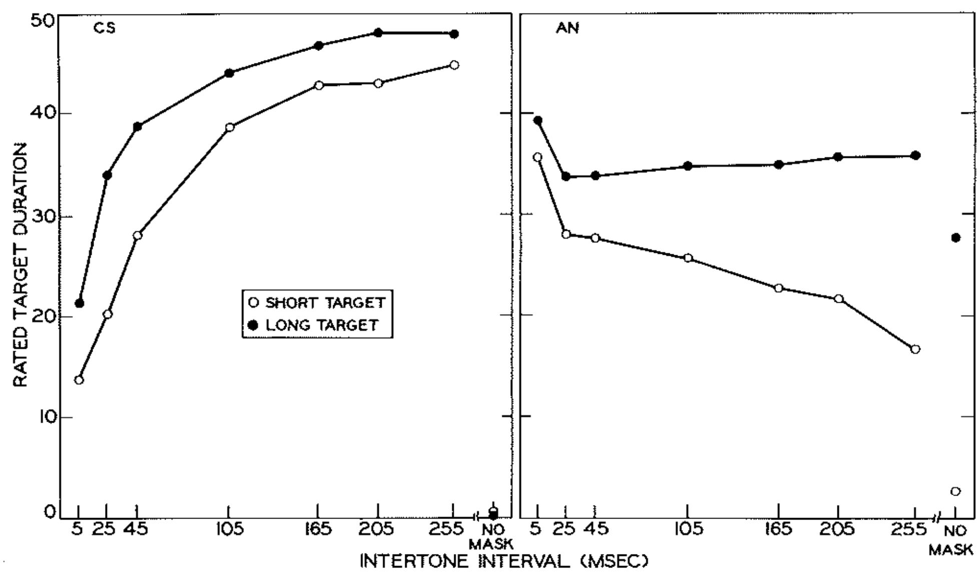

FIG. 8. Rated duration of each of the two targets as a function of ITI, for the two atypical subjects (Experiment 2).

Discussion

The results of Experiment 2 largely replicate those of Experiment 1. The long target was again rated as being longer in duration than the short target, at all ITIs. Replicating Experiment 1, rated target duration increased monotonically with increases in ITI for both target durations. These results provide additional confirmation of the model's assumption that perceived target duration is a function of actual target duration and available processing time. The second assumption, that the mask adds a constant (proportional to its duration) to the perceived duration of the target, also gained additional support. Both targets were again rated as shortest when followed by the short mask and longest when followed by the long mask. Moreover, both targets were again rated as shorter on the no-mask trials, when the mask component was absent from perceived duration, than on masking trials with long ITIs.

GENERAL DISCUSSION

Taken together, the results of Experiments 1 and 2 provide qualitative support for Massaro and Idson's (1976) model of perceived duration. Supporting the major tenet of the model, the perceived duration of a tone appears to be a function of the three factors. Perceived duration increased directly with the actual duration of the target, the available processing time, and the duration of the mask.

A more precise evaluation of the model's ability to account for the present results is possible, by developing a quantitative description of the model outlined in the introduction. Massaro's (1975) model of auditory recognition can be formalized as

$$d' = \alpha\,(1-e^{-\theta t}) \qquad [1]$$

where d' indexes stimulus discriminability, α the maximal amount of stimulus information available given infinite processing time, and θ the rate of extraction of information over time t. Massaro and Idson's (1976) model assumes that discriminability of duration changes in essentially this manner. Furthermore, the perceived duration (PD) of the target is assumed to increase as a negatively accelerated function of processing time. Consequently, the perceived duration (PD) of the target can be conceptualized as an exponential growth function of time:

$$\mathrm{PD} = \alpha\,(1-e^{-[\theta_D t_D + \theta t_I]}), \qquad [2]$$

where α is the asymptotic value of perceived duration; θ_D is the rate of growth of PD during the time of target presentation, t_D; and θ 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 perceived duration of the target by a constant, proportional to the actual duration of the mask. Incorporating this assumption into Eq. 2 yields:

$$\mathrm{PD} = \alpha(1-e^{-(\theta_D t_D + \theta t_I)}) + K(t_m), \qquad [3]$$

where t_m is the mask duration and K is a constant of proportionality. If PD_S and PD_L represent the perceived durations of the short and the long targets, respectively, then the growth of the duration percept over time is given by:

$$PD_S = \alpha_S(1-e^{-(\theta_D t_D + \theta_S t_I)}) + K_S(t_m) \quad [4]$$

$$PD_L = \alpha_L(1-e^{-(\theta_D t_D + \theta_L t_I)}) + K_L(t_m), \quad [5]$$

where θ_D represents the growth rate during target presentation, t_D: θ_S and θ_L represent the rates of growth of the short and long targets over the silent interval, t_I, to asymptotes α_S and α_L, respectively; and 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 the perceived durations (cf. Massaro & Idson, 1976, for the derivation of d' in this model).[2]

Several points deserve further discussion. Consider first the contribution of the mask. In order for the mask to contribute to the perceived duration of the target it must be processed. Logically, the same processes apply to perception of the target as to perception of the mask, with the perceived duration of the mask increasing over approximately 250 msec. In the quantitative model, the contribution of the mask is represented by a constant, because the backward-masking studies allow sufficient time for complete processing of the mask. Under these conditions, the perceived duration of the mask should be directly related to its actual duration and can be appropriately represented as a constant. If processing time for the mask was varied, the component t_m in Eq. 3 would be replaced by an exponential growth function comparable to that given for the target tone in Eq. 3.

The underlying process responsible for the contribution of the mask to the perceived duration of the target cannot be determined from the present data. Several possibilities suggest themselves. First, the processing system could integrate all of the stimulus energy occurring within a critical time interval. Such a mechanism would be analogous to that offered by Hawkins *et al.* (1974) to explain target–mask interactions in recognition masking of pitch. Although this mechanism probably accounts for the extremely long ratings at the 5-msec ITI, it cannot easily account for the relatively constant effect of masking duration for ITIs between 25 and 255 msec. Second, it is possible that the mask interferes with a representation of the target in synthesized auditory memory, in addition to terminating processing of the target (cf. Massaro, 1975, Chap. 24). In this case, the masking tone could modify the observer's memory for the target tone so that, for example, the target may be remembered as being longer than it was perceived when a long masking tone is presented. Third, a decision process could be at work

[2] The earlier version of the model (Massaro & Idson, 1976) assumed a single value of θ, representing the rate of growth of PD, which differed for the long and short targets: $PD_L = \alpha_L(1-e^{-\theta_L t}) + K_L(t_m)$; $PD_S = \alpha_S(1-e^{-\theta_S t}) + K_S(t_m)$. In order for the value of θ to differ for the long and short targets, however, the duration percept would need to be processed differentially from the onset of the targets. Yet the stimulus events corresponding to the short and long targets do not differ until the short target is turned off. Compartmentalizing the growth rate into a value for the time of target duration and the interval following target offset, as in Eqs. 4 and 5, remedies this difficulty.

in which the subject adopts a different criterion for each mask duration (cf., Massaro & Idson, 1976). Subjects may be biased to rate the perceived duration of the target tone in the direction of the perceived duration of the masking tone. Finally, the subjects could simply confuse the target and mask and inadvertently rate the mask duration. This last explanation seems somewhat unlikely, because the same target–mask interactions are found when the mask varies in frequency as well as duration (Idson & Massaro, Note 2), in which case the subjects could discriminate the two tones on the basis of pitch. Resolution of this issue clearly depends upon further experimentation.

One might be concerned with whether the subjects in our task consciously attended to just the target tone or also attended to the masking tone. Informal questioning of the observers indicated that the subjects did not consciously use the duration of the masking tone to aid their rating of the target tone's duration. However, it might be worthwhile to vary the instructions and payoffs given the subjects in order to influence the amount of weight given the masking tone in the rating judgment. If instructions and payoffs were relatively unsuccessful in eliminating the effect of the masking tone, its effects would appear to be localized at a preconscious state of processing.

Finally, consider the applicability of the model to stimuli having longer durations. In backward masking, the primary task is to discriminate between target tones. For attributes such as pitch, the information relevant to this discrimination is available to the processing system virtually from target onset, as the alternative tones differ immediately in frequency. In contrast, because the time span for the short target is a subset of the time span for the long target, no useful information for target identity will be available until the time span of the short target has been exceeded. Thus, the critical factor for duration is the intertone interval rather than the stimulus-onset-asynchrony. It would be expected that the model could apply to stimuli of substantially longer durations than those employed in the current studies, provided that the durations of the two targets were relatively similar (cf. Massaro & Idson, in press, for one example of this).

In order to fit the proposed model to the obtained data, the assumption is made that noise in the sensory system creates a distribution of perceived durations for each experimental condition. The noise distribution is assumed to be normal with a variance equal to 1. For the binary-choice data (Massaro & Idson, 1976, in press), the probabilities were transformed into relative standard-deviation units, which were taken to provide an index of perceived duration. This was necessary in order to test the predictions given by Eq. 4 and 5. Analogously, a similar transformation must be performed on the present ratings of target duration. We assume that the ratings in the current study can be taken to provide a direct index of the perceived duration of a target tone under a particular experminental condition. Then, if the further assumption can be made 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, p. 246). This allows a direct test of the predictions of Eq. 4 and 5. The assumption of an interval scale does not seem unreasonable. Moreover, by analogy to intergration theory, if the model provides a good fit to the data it will jointly validate both the model and the interval response scale. In order to fit the observed data to Eqs. 4 and 5, a linear transformation of the form

$$\text{Transformed data} = x\ (\text{observed data}) + y \qquad [6]$$

was employed in the following way. Values of x and y were estimated simultaneously with the estimates of the free parameters in Eqs. 4 and 5 in order to minimize the squared deviations between the predicted and observed values, using the reiterative minimization routine STEPIT (Chandler, 1969). The 5-msec ITI was excluded from the fit, because the intergration process which occurs at this interval requires a different estimate of α at this ITI (Massaro, 1975; Massaro & Idson, 1976). In order to ensure that the transformed scores were in the same range for the rating data as for the binary choice data, the value of α_L was set to that obtained in the comparable Massaro and Idson (1976) study.

Figure 9 presented the predicted and observed values of perceived duration for the long and the short target tones in Experiment 2. The two atypical subjects were excluded and the group data for the remaining seven subjects only were fitted. The predicted scores are represented by the connected filled squares for the long target and the connected unfilled squares for the short target. The observed points are given by the filled and unfilled circles for the long and the

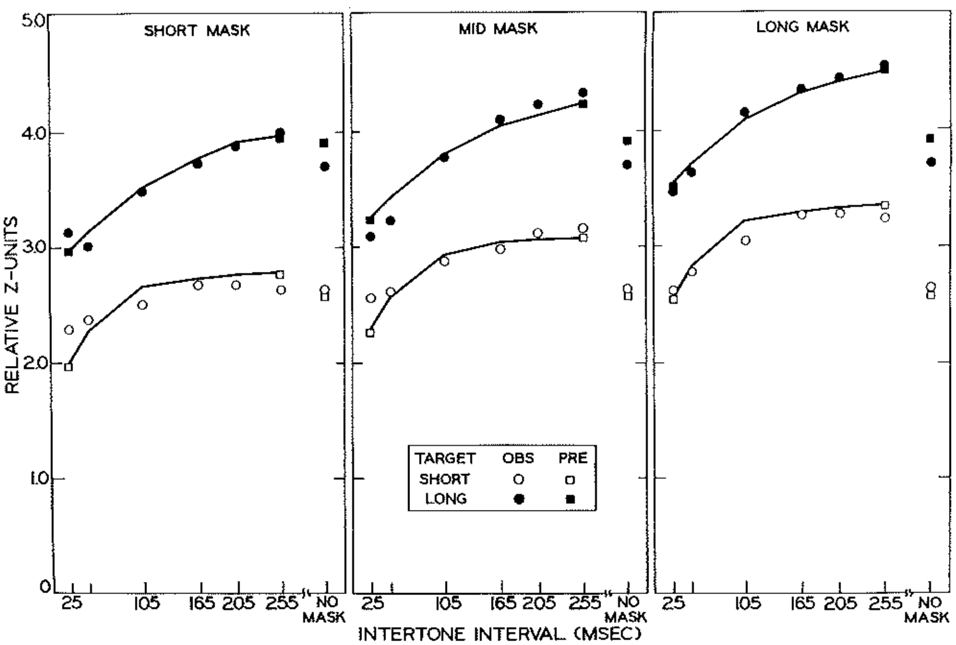

FIG. 9. Observed and predicted transformed values of rated target duration, as a function of both duration of the mask and ITI. The left-hand panel presents the data for the short masking tone, the middle panel for the mid masking tone, and the right-hand panel for the long masking tone. The no-mask values do not differ as a function of the masking-tone duration, since the latter is a dummy variable under the no-mask condition (Experiment 2).

short targets, respectively. The predicted values were computed by estimating the parameter values α_S, θ_L, θ_S, θ_D, K_L, K_S, x, and y by minimizing the squared deviations between the predicted and the observed values. As the figure shows, the model provides a relatively good fit to the data. Thirty-eight data points (two target durations × three mask durations × six ITIs + two no-mask targets) were fitted with just eight parameters. The averaged squared deviation between the predicted and the observed values was .014. The value for α_L was set at 3.90. The parameter estimate for α_S was 2.58. These estimates are in accord with the model, the long target having a greater potential perceived duration than the short target. However, these estimates indicate a lower discriminability than that seen for the two targets in the two alternative forced-choice task. In Experiment 1 of Massaro and Idson's (1976) paper, which is identical to the current Experiment 2 in all respects except for the dependent measure employed, α_S and α_L were estimated at 1.85 and 3.90, respectively. The difference in discriminability in the rating and binary-choice experiments is quite reasonable. In the binary-choice study, subjects learned the two targets prior to the test trials and received feedback throughout the study, while the rating task provided neither a learning session nor feedback. Discriminability would, therefore, be expected to be poorer in the rating than in the binary-choice task.

The parameter estimates for θ_L and θ_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 θ_D, the growth of the duration percept during target onset, was 11.15. The parameter estimates for K_L and K_S were 6.83 and 6.96, respectively, indicating that the mask had a slightly greater effect on the short than the long target. The parameter estimates for x and y were .075 and 1.75, respectively.

The quantitative formulation of the model, given in Eqs. 4 and 5, was also fitted to the group data from Experiment 1. Table 1 presents the predicted and observed transformed values of perceived duration. As Table 1 shows the model is providing quite a good fit to the obtained data. Fitting 38 points with eight parameters, the average squared deviation between the predicted and observed values was .004. The value of α_L was set to 3.64, and the parameter estimate for α_S was 3.40. This is in accord with the model, the long target having a greater potential perceived duration than the short target. The difference between α_S and α_L is smaller for the 60- and 80-msec tones used in Experiment 1 than the difference between α_S and α_L and the 50- and 90-msec tones in Experiment 2. This result confirms the psychophysical fact that decreasing the physical difference between the durations of the two target tones will decrease the degree to which they can be discriminated. The estimated values for the rates of processing were $\theta_D = 25$, $\theta_L = 16.68$, and $\theta_S = 25$. The values for K_L and K_S were estimated at 1.46 and 3.21.

The results presented in a recent paper by Kristofferson (1977) contrast with the observations made here. Observers judged duration in a single-response Donders Type-C speeded reaction-time task. On each trial, one of four possible durations was presented, and the subject responded as quickly as possible to either the two shortest or two longest target durations. The time intervals were periods of silence surrounded by 10-msec tones, and the four test durations were arrayed around a midpoint of 1,150

TABLE 1

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

	Mask duration					
	Short		Mid		Long	
ITI	Observed	Predicted	Observed	Predicted	Observed	Predicted
			Short target			
25	3.28	3.16	3.35	3.22	3.43	3.29
45	3.33	3.32	3.38	3.38	3.43	3.45
105	3.43	3.51	3.51	3.58	3.57	3.64
165	3.55	3.55	3.59	3.62	3.59	3.68
205	3.56	3.56	3.68	3.63	3.64	3.69
255	3.61	3.57	3.67	3.63	3.65	3.69
No mask[a]	3.38	3.40				
			Long target			
25	3.45	3.43	3.48	3.46	3.53	3.49
45	3.45	3.53	3.54	3.56	3.61	3.59
105	3.64	3.68	3.70	3.71	3.75	3.73
165	3.76	3.73	3.75	3.76	3.80	3.79
205	3.80	3.75	3.81	3.78	3.85	3.81
255	3.83	3.76	3.83	3.79	3.61	3.81
No mask[a]	3.61	3.64				

[a] 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 (Experiment 1).

msec. The reaction-time distributions showed that long and short responses were time-locked to different stimulus events. Short responses appear to have been time-locked to stimulus offset, whereas long responses were time-locked to stimulus onset. Kristofferson interprets these results to mean that the beginning of the target stimulus triggers an internal time interval which is compared to the target duration. The internal time interval is set between the durations of the long and short target durations. If the stimulus ends before the internal time interval is complete, a response of short is made. A response of long is initiated if the internal time interval terminates before the stimulus.

A direct application of this model to the backward-masking task would predict no effects of processing time and masking duration on the judged target durations. This follows from the fact that the response is time-locked to stimulus onset or offset and, therefore, other later stimulus events should have no influence on performance. However, the results contradict this expectation, and the model can be easily modified to account for backward masking. A response might be time-locked to a particular event but also be dependent on adequate processing time after the event has occurred. This interpretation of Kristofferson's model, therefore, also allows the possibility of backward-masking effects in judgments of duration.

A final implication of these results is worth considering. In studying duration, we found that an understanding of how the stimulus is processed was not possible unless the nature of the percept of that stimulus was considered. Only when changes in apparent duration were considered did the interactions obtained in the binary-choice data (Massaro & Idson, 1976, in press) become predictable. Yet it was not until the processing-time analysis was developed that the concept of consistent changes in the perceived duration of a constant stimulus became plausible. Thus, only by a consideration of the perceptual experience of duration and the nature of its processing was it possible to arrive at a reasonably complete interpretation. The interaction between experience and process seen for duration may apply to other areas in perception as well. For example, Idson and Massaro (1976) have recently demonstrated that the structure of an auditory stimulus has a profound impact upon its processing. Instances such as these suggest that a full understanding of certain types of perception may require considering both perceptual experience and perceptual process (e.g., Posner, Nissen, & Klein, 1976).

REFERENCES

Allan, L. G. Is there a constant minimum perceptual duration? *Quarterly Journal of Experimental Psychology*, 1976, **28**, 71–76.

Anderson, N. H. Functional measurement and psychophysical judgment. *Psychological Review*, 1970, **77**, 153–170.

Anderson, N. H. Information integration theory: A brief survey. In D. H. Krantz, R. E. Atkinson, R. D. Luce, and P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2). San Francisco: W. H. Freeman, 1974.

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*, 1970a, **8**(4), 231–234.

Efron, R. The measurement of perceptual duration. *Studium Generale*, 1970b, *23*, 550-561.

Efron, R. The minimum duration of a perception. *Neuropsychologia*, 1970c, **8**, 57–63.

Garner, W. R. *The processing of information and structure*. Potomac: Lawrence Erlbaum Associates, 1974.

Garner, W. R., Hake, H. W., & Eriksen, C. W. Operationalism and the concept of perception. *Psychological Review*, 1956, **63**, 317–329.

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.

Idson, W. L., & Massaro, D. W. Cross-octave masking of single tones and musical sequences: The effects of structure on auditory recognition. *Perception & Psychophysics*, 1976, **19**, 155–175.

Kristofferson, A. B. A real-time criterion theory of duration discrimination. *Perception & Psychophysics*, 1977, **21**, 105–117.

Marks, L. E. *Sensory processes: 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. Preperceptual images, processing time, and perceptual units. *Psychological Review*, 1972a, *79*, 124–145.

Massaro, D. W. Stimulus information versus processing time in auditory pattern recognition. *Perception & Psychophysics*, 1972b, **12**, 50–56.

Massaro, D. W. *Experimental psychology and information processing*. Chicago: Rand McNally, 1975. McNally, 1975.

Massaro, D. W., Cohen, M. M., & 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. Temporal course of perceived vowel duration. *Journal of Speech and Hearing Research*, in press.

Moore, J. J., & Massaro, D. W. Attention and processing capacity in auditory recognition. *Journal of Experiment Psychology*, 1973, **99**, 49–54.

Posner, M. I., Nissen, M. J., & Klein, R. M. Visual dominance: An information processing account of its origins and significance. *Psychological Review*, 1976, **83**, 157–170.

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.

REFERENCE NOTES

Note 1: Idson, W. L., & Massaro, D. W. *Backward masking of perceived pitch.* Unpublished manuscript, 1976.

Note 2: Idson, W. L., & Massaro, D. W. *Target–mask similarity in the perception of duration.* Manuscript in preparation.