Journal of Experimental Psychology
1973, Vol. 99, No. 1, 49–54

ATTENTION AND PROCESSING CAPACITY IN AUDITORY RECOGNITION [1]

JEFFREY J. MOORE AND DOMINIC W. MASSARO [2]

University of Wisconsin—Madison

This research examined whether attention and processing capacity limitations affect the recognition of single tones. The *S*s identified either one or both dimensions (loudness and quality) of a short test tone. Processing time was controlled by following the test tone with a masking tone after a variable silent interval. Performance improved with increases in the silent intertone interval, but was not affected by whether *S*s identified one or both tone dimensions. These results indicated no capacity limitations or selective attention during the recognition of these dimensions. The pattern of the concurrent responses supports these conclusions by indicating that the dimensions were processed independently and in parallel.

The human information-processing system has a limited capacity. It is limited in the amount of information it can process (respond appropriately to) per unit of time. It has been hypothesized by Broadbent (1958) that all sensory inputs converge on a central processor of limited capacity. Selective attention allows the system to process selectively some channels of information at the expense of losing other information. Previous studies have identified a number of dimensions that are processed as competing channels of information, such as spatially separated auditory messages (Moray & O'Brien, 1967), messages similar in voice (Triesman, 1964), and signals presented in different modalities (Lindsay, 1970; Lindsay, Taylor, & Forbes, 1968; Massaro & Kahn, 1973).

Selective attention in these situations has been assumed to occur before recognition is complete, according to the theory of Broadbent (1958, 1971) and Triesman (1960, 1969). They have proposed that a hierarchical series of progressively more complex analyses are performed on a message during its recognition. Selective attention is said to occur within this series, and a message is selected on the basis of the simplest analysis that distinguishes between the inputs. For instance, if a relevant message can be distinguished on the basis of physical cues, such as the speaker's voice, then selective attention would operate after the processing stage at which physical cues to voice are extracted. Given that the speaker's voice is recognized, the relative amount of further processing for semantic content of the irrelevant message would be limited. In this model, selective attention can operate at a preliminary stage of recognition before it is complete. In contrast, Deutsch and Deutsch (1963) and Norman (1968, 1969) have proposed that attention operates after all inputs are analyzed and matched with their meanings in memory, that is, after they are recognized.

The purpose of this experiment was to determine whether two dimensions of a single auditory stimulus are processed as competing channels of information, and if they are, whether *S* can selectively attend to one dimension specified at the beginning of a trial. On each trial of this experiment *S* was required to recognize either the loudness, quality, or both of these dimensions of a test tone. On selective attention trials, *S* was cued in advance to recognize either the loudness or the quality of the test tone. On divided attention trials, *S* was cued to recognize both the loudness and the quality of the test tone. The difference in performance on selective and divided attention trials will index the de-

[1] This investigation was supported in part by U. S. Public Health Service Grant MH-19399-02 to the second author.

[2] Requests for reprints should be sent to Dominic W. Massaro, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

gree to which S can selectively attend to these dimensions of the test tone.

To determine whether S can selectively process one dimension for recognition at the expense of another, it was necessary to hold other stages of information processing constant in the recognition task. Accordingly, the memory and response-selection stages were experimentally controlled so that changes in the selective attention manipulation only could affect processing in the recognition stage. Memory for test tones is critical for their recognition and may vary with the amount of feedback S receives and with the time between trials to rehearse the tones. Since S would have to remember both dimensions for the divided attention condition and only one dimension for the selective attention condition, memory could differ in the two conditions. Therefore, the selective and divided attention conditions were presented randomly from trial to trial to insure that memory for the test tones was held constant under the two attention conditions. The time for response selection may differ as a function of the attention manipulation. In the present experiment one response was required in the selective attention condition, while two were required in the divided attention condition. Accordingly, S was given more time to respond in the divided attention condition.

Given sufficient time, a limited capacity system could optimally process several dimensions of a signal simply by switching between them. Therefore, it is necessary to control the time available for the identification task. Although varying the duration of the test tone varies its clarity or the amount of information which it contains (Massaro, 1972b), it does not sufficiently control the perceptual processing time for that information. A preperceptual image of the stimulus persists, extending the effective duration of perceptual processing. The processing of the preperceptual image can be interfered with by presenting a masking tone after the test tone presentation. Massaro (1972a) has demonstrated that if a short test tone is followed after some silent interval by a masking tone, performance improves with increases in the duration of the silent interval. Information is extracted from a preperceptual image of the tone during the silent interval until halted by the occurrence of the masking tone. In the present experiment, perceptual processing time was experimentally controlled using the recognition masking paradigm.

Previous studies have not conclusively demonstrated that dividing attention between two dimensions of a tonal signal decreases identification performance. Lindsay et al. (1968) reported capacity limitations in a test measuring the discrimination of two auditory dimensions (pitch and loudness) under conditions where either one or two dimensions required attention. However, these findings are confounded with differences in the test stimuli used in the two attention conditions. Only one stimulus dimension was varied in the selective attention condition, whereas both dimensions were varied in the divided attention task. For example, in the selective attention condition when S identified which of two intervals contained the louder tone, the irrelevant pitch dimension was the same in both intervals on a trial. In contrast, in the divided attention condition, both the pitch and loudness differed in the two intervals. In the selective attention condition S identified the loudness of tones equal in pitch, whereas in the divided attention condition S identified the loudness of tones unequal in pitch. The varying pitch dimension could have made the loudness judgment more difficult. For example, Montague (1965) demonstrated that identification performance on pitch and loudness dimensions was reduced by varying one to three additional irrelevant dimensions. It follows that in the Lindsay et al. study an additional varying stimulus dimension in the divided attention condition may account for the lower performance relative to the selective attention condition.

The present experiment examines the possibility of selective attention in the auditory recognition stage, using binaurally presented pure tones varying on both dimensions of amplitude and wave shape in all conditions. Processing time was controlled with a masking paradigm, so that selective and divided attention conditions could be compared over a continuous range of task difficulty.

Method

Subjects. Twenty-two undergraduates fulfilling a course requirement at the University of Wisconsin were employed 1 hr. a day for 5 days. Two *S*s were eliminated from the analysis as they averaged only 3% above chance.

Procedure. On each trial, *S* was presented with a test tone followed by a variable silent interval followed by a masking tone. The *S* was required to identify either the loudness (soft or loud) or the quality (dull or sharp) or both dimensions of the test tone. The loudness of the tone was either soft (68.5 db. SPL) or loud (74.2 db.). The quality of the test tone was either dull (a pure sine wave) or sharp (a pure triangle wave). The frequency of the test tone was held constant at 855 Hz.

For 10 *S*s, the test tone was presented for 20 msec. followed by one of eight silent intertone intervals: 10, 20, 40, 70, 120, 200, 300, or 400 msec. To improve the level of performance for the other 10 *S*s, the test tone was presented for 40 msec., and each of the eight intertone intervals was increased by 10 msec. The masking tone was presented for 120 msec. and consisted of six 20-msec. square-wave segments of different amplitudes, ranging from 67 db. to 75 db. with a steady-state SPL of 73 db. The frequency of the masking tone was 855 Hz.

Each trial began with the presentation of a visual cue that preceded the test tone presentation by 2 sec. and remained on throughout the response interval. This visual cue signified whether *S* was required to identify only loudness, only quality, or both dimensions on that trial. The *S* was given 2 sec. to respond on the selective attention trials (loudness or quality) and 3 sec. to make both responses (loudness and quality) on the divided attention trials. The *S* responded by pressing push buttons labeled "Soft-1," "Loud-2," "Dull-3," and "Sharp-4." Feedback was then given for the dimension(s) of the test tone that had been identified by presenting the appropriate digit(s), 1, 2, 3, or 4. The digits were presented on a visual readout for .5 sec. each. On divided attention trials, feedback was always given on the loudness dimension first. The intertrial interval was also .5 sec. All 96 experimental conditions (3 attention tasks, 2 test tone amplitudes, 2 test tone shapes, and 8 silent intervals) were completely random and programmed to occur equally often.

All experimental events were controlled by a PDP-8L computer. The pure tones were produced by a digitally controlled oscillator (Wavetek Model 155). The tones were presented binaurally over matched headphones (Grason Stadler Model TDH-49). The visual cue and the feedback digits were presented over visual readouts (Industrial Electrical Engineers Readout 800). Up to four *S*s were tested simultaneously in separate sound-attenuated rooms.

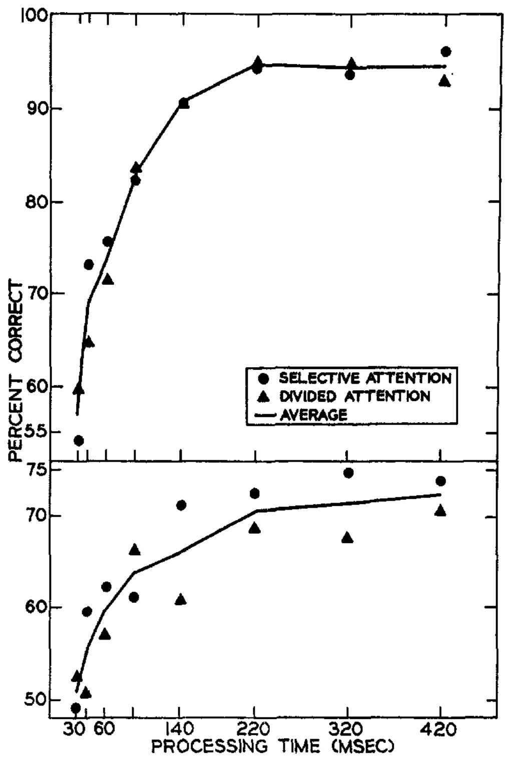

FIG. 1. Percentage of correct identifications of the loudness of the 20-msec. test tone as a function of processing time (test tone duration plus intertone interval) and selective and divided attention. (Top and bottom panels present performance for the five best and five worst *S*s, respectively.)

Practice. To encourage the strategy of listening only to one dimension in the selective attention condition, blocks of 80 trials were given under each of the three attention conditions on the first day. For each attention condition a block of trials was presented without a masking tone. After these three blocks of trials, *S*s were given three blocks of 80 trials with the masking tone present. Data were then collected for 4 days, with an average of 500 trials per *S* per day.

Results

There was very little advantage in performance for the selective vs. the divided attention condition (1.75% averaged over *S*s), $F(1, 18) = 14.3$, $p < .005$. This difference was the same for both stimulus dimensions.

Figures 1 and 2 present percentage of correct identifications of the 20-msec. test

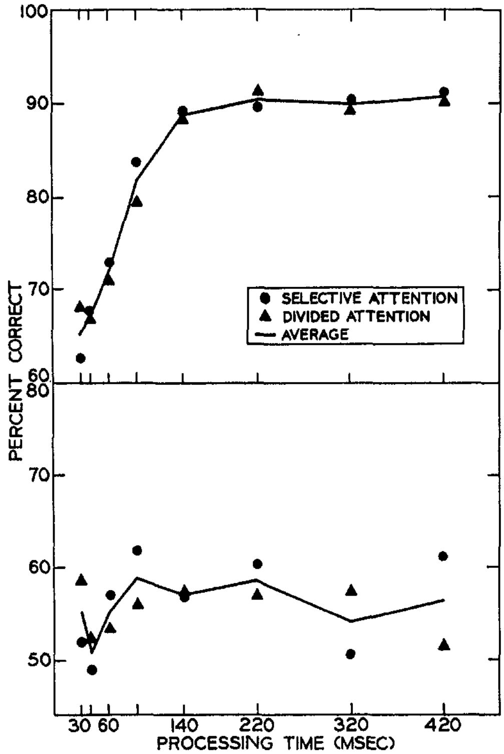

FIG. 2. Percentage of correct identifications of the quality of the 20-msec. test tone as a function of processing time (test tone duration plus intertone interval) and selective and divided attention. (Top and bottom panels present performance for the five best and five worst *S*s, respectively.)

tone in the selective and divided attention conditions as a function of processing time (test tone duration plus intertone interval). These results are also representative of performance in the 40-msec. tone conditions. For all conditions, performance improved significantly with increases in intertone interval, F (1, 126) = 59, $p < .001$. The improvement was greater for loudness identification (26%) than for quality (13%), F (1, 18) = 8.5, $p < .001$. The difference between attention conditions did not change systematically with increasing processing time, ranging between −1.4% and 3.4%. Finally, test tone duration had no significant effect on performance.

The method of collecting data was changed during the course of the experiment to permit examination of performance on both dimensions in a given trial in the divided attention condition. These data are available for 13 *S*s. For each *S* the observed probability of correctly identifying both the loudness and the quality of the test tone was computed. If *S* were identifying the dimensions independently without interference, the observed probability of getting both dimensions correct should be equal to the product of the correct marginal probabilities in the divided attention condition. On the other hand, if identifying one dimension interfered with identification of the other, *S* would tend to be correct on only one dimension on a trial. In that case, the observed joint probability of correctly identifying both would be less than the expected probability predicted from independence. The four expected joint probabilities (of correctly identifying both dimensions, just loudness, just pitch, or neither) were calculated by multiplying the observed marginal probabilities. The observed and expected joint probabilities were compared by a chi-square test for each *S*. Table 1 shows the observed and expected joint probabilities of correctly identifying both dimensions. The results indicate that identifying one dimension did not interfere with identifying the other. The observed joint probability did not significantly differ from the expected joint probability for 10 *S*s, was

TABLE 1

OBSERVED AND EXPECTED PROBABILITIES OF CORRECTLY IDENTIFYING BOTH DIMENSIONS ON A GIVEN TRIAL IN THE DIVIDED ATTENTION CONDITION

S	Probability	
	Observed	Expected
6	50.6	47.8
7	58.2	56.5
10	79.7	78.8
11	35.8	35.6
12	41.4	42.3
13	38.7	36.9
14	44.6	43.2
15	44.7	44.3
16	50.8	49.7
17	53.7	53.8
18	53.7	50.8
19	55.6	54.1
20	82.4	82.2

greater for 3 *S*s ($p < .01$), and never differed by more than 3%.

Discussion

The large improvement in performance with increases in intertone interval demonstrates that processing time is critical for the auditory recognition process. The 1.75% difference in the two attention conditions is too small to be considered meaningful and is negligible compared to the effect of processing time. The present study provided a strong test of selective attention since it covaried a second independent variable that controlled the duration of the perceptual process and consequently the level of performance. Even so, the attention variable did not affect performance. The results demonstrate that two dimensions of an auditory signal can be processed in parallel without a performance decrement, and auditory recognition of one dimension cannot be enhanced by selectively attending to that dimension.

A limited capacity model is not necessary to describe the process of recognizing the quality and loudness of auditory tones. Recognition involves the analysis of the features of the stimulus in order to decide which stimulus was presented. Since no capacity limitations were demonstrated, these results can be modeled more clearly in terms of separate and independent channels of feature extraction and decision for each dimension. Two of our results disconfirm a serial processing model that assumes each dimension is processed serially on the auditory recognition task. Serially processing one dimension at a time should lower the probability of being correct on both dimensions on the same trial (Moray & O'Brien, 1967). In contrast, identification performance in the divided attention condition was accurately predicted by assuming that identification of one dimension was independent of identification of the other. Serially processing one dimension at a time also implies that performance in the divided attention condition would improve at a much slower rate with increases in processing time relative to the selective attention condition. The results indicate that, contrary to this prediction, the rate of improvement with increasing processing time did not differ significantly in the two attention conditions.

It is worthwhile to consider the present results in terms of Garner's (1970) distinction between integral and separable stimulus dimensions. According to Garner, the dimensions of loudness and quality would be integral since the specification of one dimension requires specification of the other. For pure tones, the loudness level must be defined with respect to some wave shape and vice versa. Garner predicts that two integral stimulus dimensions cannot be treated independently. For example, in the present study, listeners could treat the bidimensional stimulus as a gestalt unit and identify it absolutely as one of four alternatives. If the stimuli were analyzed as units, *S* should correctly identify both dimensions on the same trial more frequently than predicted from the marginal probabilities by the independence hypothesis. However, contrary to the unit hypothesis, the observed probabilities of correct concurrent responses were not significantly greater than those predicted from independence. This result indicates that either (*a*) the definition of integrality must be modified or (*b*) some integral dimensions can be processed independently.

The findings of the present experiment contrast somewhat with those from a related attention study by Massaro and Kahn (1973). In the selective attention condition of that study, *S*s were required to identify the quality of a test tone as sharp or dull. The test tone was followed by a masking tone after a variable silent interval. In the divided attention condition, in addition to identifying the test tone, the *S*s also identified the duration (long or short) of a visual stimulus presented simultaneously with the test tone. Identification of the tone in the divided attention condition was 5.4% lower than in the selective attention condition. Test tone recognition improved with increases in the silent intertone interval duration, and the difference between the attention conditions was clear at all but the shortest interval. Therefore, dividing attention between the auditory and visual tasks reduced performance on the auditory task. Furthermore, the difficulty of the auditory recognition task affected performance on the visual task since performance on the duration identification task improved with increases in the silent intertone interval.

The Massaro and Kahn (1973) experiment demonstrated that tone quality recognition was lowered 5.4% in divided attention, whereas the present study found only a 1.75% performance decrement. The difference in the additional processing task in the two experiments might account for the different results. In the Massaro and Kahn experiment, *S*s were required to recognize two sensory inputs in different modalities, whereas *S*s recognized two dimensions of one auditory input in the present experiment. It might be possible that dividing attention between two stimuli is more difficult than dividing attention between two dimensions of one stimulus. Further research is necessary to clarify these differences.

Shiffrin and Gardner (1972) studied limited capacity and attentional effects in recognition of four-letter visual displays. To determine if recognition was limited in capacity and was capable of attentional control, it was necessary to control for two other important processes in the visual recognition task. First, the displays were made so that lateral masking could not differ under the selective and divided attention conditions. Second, performance under a simultaneous display presentation was compared to performance under a sequential display presentation to control for effects of the decision process in the task. The results demonstrated that visual processing of spatial location does not have a limited capacity and is not capable of attentional control by *S*. The four locations were processed independently, as were the two auditory dimensions in the present study.

A number of information-processing theories have been developed around the central concept of limited capacity and attentional effects (Broadbent, 1958, 1971; Deutsch & Deutsch, 1963; Norman, 1968). The point of disagreement of these theories concerns the stages of processing that are limited in capacity. Recent work has shown that whether limited capacity and attentional effects are to be found at a particular stage of processing (recognition) is dependent on the nature of the information-processing task. These results call for development of an information-processing model of recognition centered around perceptual processing that will predict when limited capacity and attentional effects will be found. A theory of attention should directly follow, not precede, a theory of the recognition process.

REFERENCES

Broadbent, D. E. *Perception and communication.* London: Pergamon, 1958.

Broadbent, D. E. *Decision and stress.* New York: Academic Press, 1971.

Deutsch, J. A., & Deutsch, D. Attention: Some theoretical considerations. *Psychological Review*, 1963, **70**, 80–90.

Garner, W. R. The stimulus in information processing. *American Psychologist*, 1970, **25**, 350–358.

Lindsay, P. H. Multichannel processing in perception. In D. E. Mostofsky (Ed.), *Attention: A behavioral analysis.* New York: Appleton-Century-Crofts, 1970.

Lindsay, P. H., Taylor, M. M., & Forbes, S. M. Attention and multidimensional discrimination. *Perception & Psychophysics*, 1968, **4**, 113–117.

Massaro, D. W. Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, 1972, **79**, 124–145. (a)

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

Massaro, D. W., & Kahn, B. J. Effects of central processing on auditory recognition. *Journal of Experimental Psychology*, 1973, **97**, 51–58.

Montague, W. E. Effect of irrelevant information on a complex auditory-discrimination task. *Journal of Experimental Psychology*, 1965, **69**, 230–236.

Moray, N., & O'Brien, T. Signal-detection theory applied to selective listening. *Journal of the Acoustical Society of America*, 1967, **42**, 765–772.

Norman, D. A. Toward a theory of memory and attention. *Psychological Review*, 1968, **75**, 522–536.

Norman, D. A. *Memory and attention.* New York: Wiley, 1969.

Shiffrin, R. M., & Gardner, G. T. Visual processing capacity and attentional control. *Journal of Experimental Psychology*, 1972, **93**, 72–82.

Triesman, A. M. Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 1960, **12**, 242–248.

Triesman, A. M. Verbal cues, language, and meaning in selective attention. *American Journal of Psychology*, 1964, **77**, 206–219.

Triesman, A. M. Strategies and models of selective attention. *Psychological Review*, 1969, **76**, 282–299.

(Received July 17, 1972)