

Form of redundancy as a determinant of tachistoscopic word recognition*

ROBERT L. COLEGATE† and CHARLES W. ERIKSEN
University of Illinois, Champaign, Illinois 61820

Single letters and redundant CVCs and CCCs were presented at threshold durations and followed by a pattern mask. Performance on three-letter stimuli was greater than on single letters. The increased performance was consistent with a state model of perceptual independence. Conditional redundancy increased the frequency with which all three letters were accurately reported as compared to correlational redundancy. Meaning and pronounceability had little effect on performance.

Perhaps one of the most striking features of written language is the high degree of redundancy existing within words. Missing letters or typographical errors typically cause little apparent difficulty for the skilled reader and often go unnoticed. It has been suggested that ample opportunity exists for improving communications by recoding words into fewer letters. Although there are obvious practical difficulties with recoding schemes, there may be other more compelling reasons for not improving the efficiency by reducing the number of symbols.

Garner (1962) has suggested that one function of redundancy in written language is to improve discrimination by providing multiple sources of information. Missing letters or typographical errors go unnoticed only when the context, the remainder of the word, or both provide overlapping information. A question of some interest, however, is to what extent does the way in which overlapping information is distributed among the elements of a visual pattern effect the discrimination of one member from a set of alternatives. Perfect correlations, where each of two letters provides the same information, is an effective form of redundancy, but one which is not found in the English language. Conditionals that can be expressed as rules of the type "CK may occur in the middle or at the end of a word but can never begin a word" or "No more than two vowels ever occur in sequence within a word" are another way in which multiple sources of information can reduce uncertainty. However, each of the multiple sources

does not provide the same information when redundancy takes this form.

One way in which multiple sources of the same information improve discrimination is described by a multiple look model of perceptual independence (Garner & Flowers, 1969; Eriksen, 1966; Doherty & Keeley, 1969). Two tachistoscopic presentations of the same letter result in a gain in recognition accuracy over that for a single letter which is closely approximated by a statistical definition of independence.

The improved discrimination derives from a decision process which differentially weights input arising from two or more states of sensitivity of the perceptual system. It is assumed that there are random variations in the states of the sensory channels associated with small areas of the retina. It is further assumed that the sensitivities of these areas are uncorrelated. Two presentations of the same letter increase the probability that at least one letter will fall on an area in a momentarily high state of sensitivity. S selects the response arising from the highest state.

This formulation would suggest a simple system where the redundancy within a word is compatible with characteristics of the visual perceptual system. When viewing conditions are such that error factors are of consequence, the sensitivity of the visual system is not homogeneous across some spatial extent. Overlapping information in the form of perfect correlations across spatial locations can improve discrimination without increasing sensitivity. However, as was suggested above, most of the important forms of redundancy within words are conditionals.

An extension of the notion of varying sensitivity of the sensory channels interacting with intraword redundancy in the form of conditionals would still seem tenable if certain conditions were met. One is that the perceiver has a rather complete knowledge of the

conditional relationships. This is probably true for literate adults up to a point (Garner, 1962). The second condition is that a single glance is sufficient for the perceiver to derive information from the stimulus (independent of redundancy) about the identity of more than one letter. This condition is met in that the number of letters recognized from a tachistoscopically presented multiletter display is not linearly related to exposure duration. An exposure duration that results in a 75% recognition accuracy level for a single letter is sufficient for two letters to be recognized at an average accuracy level of 69% (Eriksen & Lappin, 1967).

The number of forms of redundancy found among the letters of actual words present difficult problems for investigating the effect of one particular form on discrimination. These problems were avoided by constructing artificial sets of stimuli where overlapping information was provided in the form of either perfect correlations between letters or in the form of conditionals where two letters uniquely determined a third. Single trigrams were presented tachistoscopically and followed at one of several delays by a patterned mask in order to assess recognition performance at several levels. Ss worked with only one set of stimuli during each phase of the experiment, memorized the members of each set, and were acquainted with the relationships existing among letters. The number of alternative responses per letter position was the same for each set of stimuli, as was the number of members (three-letter combinations) in each set. Thus, any differences in recognition accuracy between sets of stimuli can be attributed to differences in the form of redundancy.

METHOD

Subjects

Four undergraduate students at the University of Illinois (one female) served as paid Ss.

Apparatus

The stimuli were displayed with a Scientific Prototype Model GA three-field tachistoscope which had been modified by replacing the standard bulbs with Sylvania F4T5/CWX fluorescent lamps.

Stimuli

A population of consonant-vowel-consonant (CVC) stimuli was generated in the following way. Three letters were selected for each of three positions designated left, middle, and right. The letters that could occur in

*This investigation was supported by United States Public Health Service Research Grant MH-1206, United States Public Health Service Research Career Program Award K6-MH-22014, and Predoctoral Fellowship 1-F01-MH-47201-01.

†Requests for reprints should be made to Robert Colegate, who is now at the Department of Psychology, Norfolk State College, Norfolk, Virginia 23504.

Table 1

CVC-LMC	Stimulus Set		CVC-SL
	CVC-COND	CCC-COND	
HAD	HAN	HFN	HWW
HAG	HUG	HJG	PWW
HAN	HOD	HQD	RWW
POD	PAD	PFD	WAW
POG	PUN	PJN	WOW
PON	POG	PQG	WUW
RUD	RAG	RFG	WWD
RUG	RUD	RJD	WWG
RUN	RON	RQN	WWN

these positions were H, P, and R; A, O, and U; and D, G, and N, respectively. All possible combinations of three letters, preserving the correspondence between letters and positions, yields 27 CVCs. The two sets of stimuli drawn from this population and an additional two variant sets are described below and listed in Table 1.

A set of nine CVCs was selected from this population such that the left and middle letters were perfectly correlated. The identity of the letter in the right position was orthogonal with respect to the letters in the left and middle positions. This set was designated CVC left and middle positions correlated (LMC). A second set was selected from the larger population such that the identities of letters in any two positions were uncorrelated with each other; however, the identities of letters in any two positions taken jointly fully determined the identity of the letter in the remaining position. This set was designated CVC conditional (COND).

Since seven members of Sets CVC-LMC and CVC-COND have a unique meaning and all members of these sets are pronounceable as a single syllable, it was considered desirable to determine recognition accuracy for stimuli that do not exhibit these properties. Thus, a third set of nine stimuli was constructed by replacing the letters A, O, and U of Stimulus Set CVC-COND with the letters F, Q, and J, respectively. This set was designated CCC-COND.

It was necessary to assess recognition accuracy in the absence of redundancy. Thus, a set of nine single-letter stimuli was constructed, in which each of the letters H, P, and R; A, O, and U; and D, G, and N were used once in the left, middle, and right positions, respectively. In order to make this set of stimuli comparable with regard to adjacent contours, the letter W was placed in the two unused positions. This set was designated CVC single letter (SL).

Stimuli were constructed by mounting Paratype No. 11316 18-pt Futura bold capital letters on white vinyl cards. One member of each set was placed on each card. The letters

were black, .2 deg of visual angle high x .12 deg wide and spaced .33 deg center-to-center.

The masking field contained a diagonal black and white striped pattern and had an average luminance of 4 mL. The stripes were .15 deg of visual angle in width. The adaptation field was dark at all times except for a transilluminated fixation cross subtending .38 deg at a luminance of .5 mL. The middle letters of the stimuli appeared .6 deg above the fixation cross.

Procedure

All Ss were given a list of the members of Set CVC-SL, which was learned rapidly. This set was used to establish and maintain a 75% correct duration threshold throughout the experiment. S fixated on the cross and, when ready, pressed a switch to turn on the stimulus field and reported one of the nine letters. During the first session, exposure durations were initially long, but they rapidly decreased to a relatively stable value as Ss became proficient at the task. All other stimuli were presented at the duration that yielded 75% accuracy on the CVC-SL set. Throughout the

experiment, this duration was adjusted before each session as required to compensate for improvement through practice. The range of exposure durations was 1.5 to 3.5 msec.

The second session was used to learn the members of S's first stimulus set and to provide practice on reporting the stimuli when followed by the mask at one of several fixed intervals. Each S was given a list which was memorized to a criterion of two errorless recitations. The nature of the information overlap was explained. Stimulus set order was counterbalanced between Ss. List learning (approximately 5 min of the 50-min session) was omitted for the S who started with Stimulus Set CVC-SL. For Sets CVC-LMC, CVC-COND, and CCC-COND, S reported three letters on every trial in sequential order, with the left position first and the right position last. For Set CVC-SL, S reported one letter on every trial.

Five experimental sessions of eight blocks each, with 18 trials per block, were given for each set. Each stimulus occurred twice within a block in randomized order. In each session, blocks of trials were run at

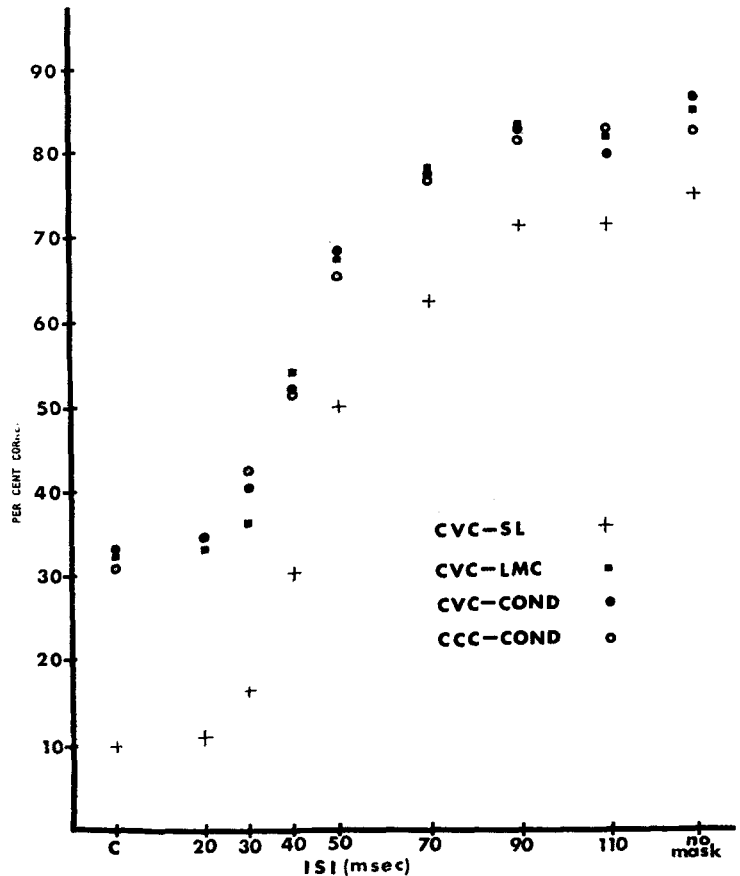


Fig. 1. Proportion of letters correct as a function of stimulus set and ISI.

interstimulus intervals (ISI) of concurrent, 20, 30, 40, 50, 70, 90, and 110 msec between the offset of the stimulus and the onset of the mask. The duration of the mask was 10 msec. The order of ISIs was counterbalanced within Ss. Each experimental session was preceded by 5 min of dark adaptation and warm-up. In addition, one block was run without the mask.

Following the five experimental sessions on the first stimulus set, Ss were given a new stimulus set, one practice session, and five experimental sessions. This procedure was continued until each S had completed all stimulus sets. Thus, each S completed 25 sessions, 5 practice and 20 experimental.

RESULTS AND DISCUSSION

The proportion of correct letter recognition responses averaged across positions and Ss are shown for each stimulus set in Fig. 1 as a function of ISI. The significance of differences was evaluated with a three-way classification analysis of variance on factors of Ss, stimulus sets, and ISI. The analysis showed a significant variation between Ss [$F(3,21) = 11.408, p < .01$], stimulus sets [$F(3,9) = 127.62, p < .01$], and ISI [$F(7,21) = 195.00, p < .01$]. As is evident in Fig. 1, performance on all trigram sets (CVC-LMC, CVC-COND, CCC-COND) was approximately equal but at a higher level than on the letter set. The ISI effect was monotonic for all sets, a typical result from most backward masking studies.

Both interactions involving ISI were significant: ISI by Ss [$F(21,63) = 3.027, p < .01$] and ISI by Stimulus Set [$F(21,63) = 2.339, p < .01$]. The latter interaction reflects a decreasing difference in performance between the trigram sets and the letter set, primarily over the range of ISIs from concurrent to 50 msec.

Performance on the letter set was limited by correct perception of the letter position as well as by recognition of the letter. The probability of a correct response, given that S could not discriminate which position was filled with a response letter, was 1/9. The comparable value for the trigram sets was 1/3. For ISIs where the difference between the trigram stimulus sets and the letter stimulus set was decreasing, position errors (e.g., a letter appropriate to the left or right position was reported in response to a stimulus displayed in the middle position) on the letter set were also decreasing.

If S has no information about the stimulus and guesses randomly, the probability of an error is 8/9, and the probability of a position error is 6/9.

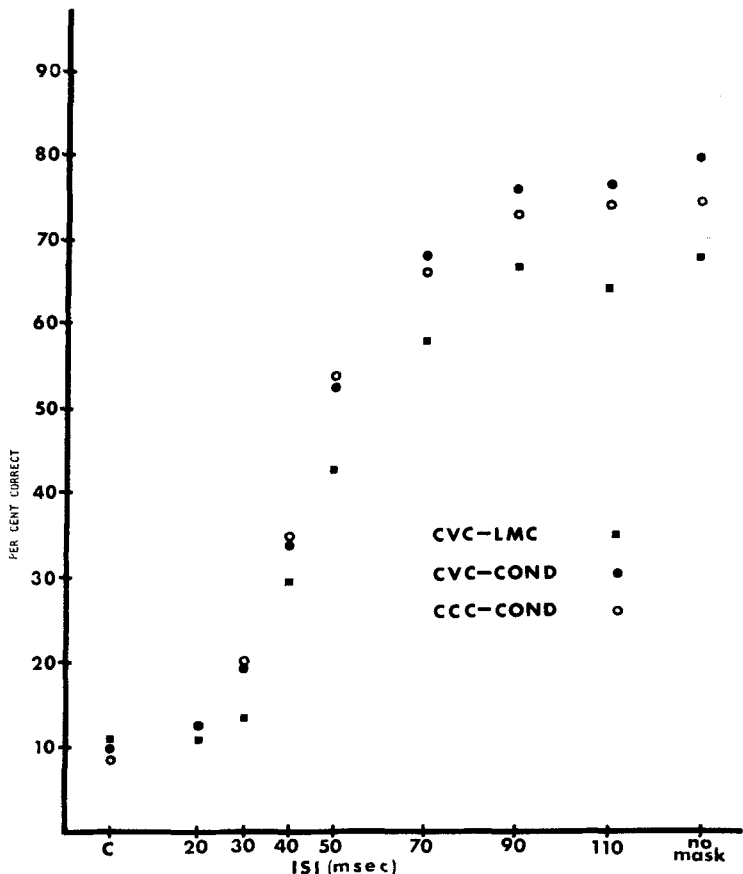


Fig. 2. Proportion of trials on which all letters were correct as a function of stimulus set and ISI.

The ratio of position errors to total errors ranged from .72 for the concurrent condition to .18 for the no-mask condition. The value of .72 is very close to the expected value of .75 for chance performance. Thus, this interaction was indicative of a response constraint occurring in the letter set.

Inspection of the data suggests that the interaction of Ss by ISI was due to one S who showed a faster recovery from masking for all stimulus sets, beginning his recovery about 10 msec before the other three Ss. It would seem reasonable that this S was more adept at discriminating the degraded stimuli at the shorter ISIs.

In order to examine more closely the effect of form of redundancy, the trigram conditions were analyzed further. The proportion of trials on which all three letters were correct are shown in Fig. 2 for each set as a function of ISI. The significance of differences was evaluated with a three-way classification analysis of variance on factors of Ss, stimulus sets, and ISI. The analysis showed significant variation between Ss [$F(3,42) = 7.620, p < .01$], stimulus

sets [$F(2,6) = 10.680, p < .01$], and ISIs [$F(7,21) = 179.070, p < .01$]. The interaction of Ss by ISI was also significant [$F(21,42) = 1.98, p < .05$]. Again, this was attributed to the same S's recovery from masking approximately 10 msec before the other Ss.

The significant difference between stimulus sets was largely due to the performance decrement of CVC-LMC relative to CVC-COND and CCC-COND.

The proportion of correct letter recognition responses on Set CVC-LMC are presented by position and ISI in Fig. 3. Performance on letters in the left and middle positions is at least 10% more accurate than for those in the right position for ISIs of 40 msec or greater. The same data for the right position of Set CVC-SL are also plotted in Fig. 3 to facilitate comparison. The difference between performance on the right position of the sets is very large for shorter ISIs but decreases to zero for the 110-msec ISI and no-mask conditions.

The data given in Fig. 3 for Set CVC-LMC qualitatively conform to expectations from the multiple look

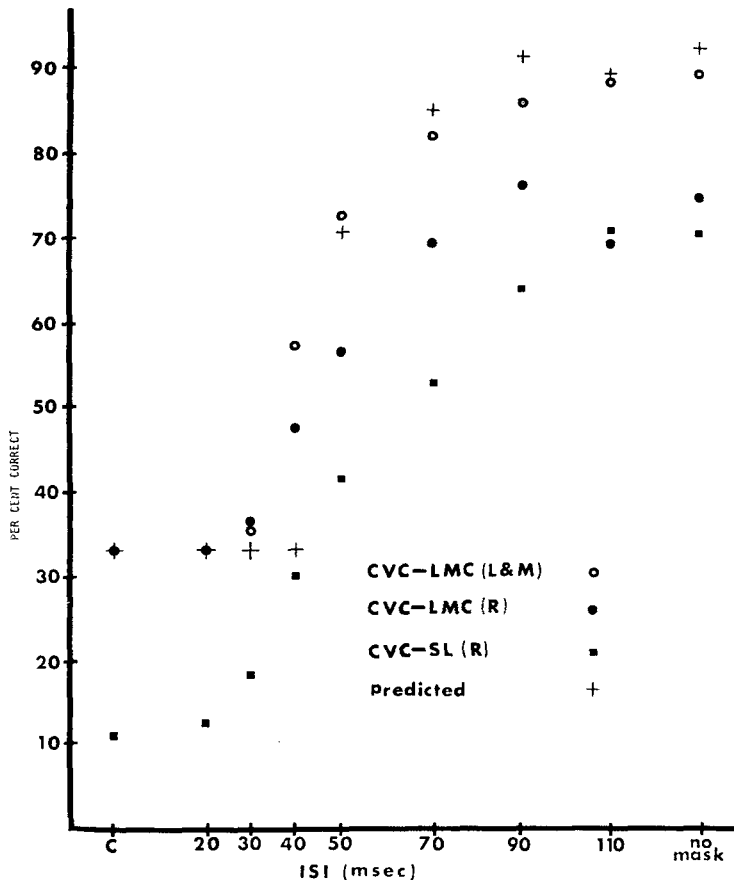


Fig. 3. Proportion of letters correct for Stimulus Set CVC-LMC as a function of letter position and ISI.

model. Accuracy for the correlated left and middle positions is appreciably higher than for the right. However, there is an apparent quantitative discrepancy. The right position of the CVC-SL set should approximately equal the same position for the CVC-LMC set. This relationship holds only for the 110-msec ISI and no-mask conditions. In addition, accuracy for the left and middle positions of the CVC-LMC set, when corrected for chance, should approximately equal $L + M - LM$,

where L and M are proportion of letters recognized in the left and middle positions of Set CVC-SL. The results of such computations (plotted as +s in Fig. 3) underestimate performance for the short ISIs but seem reasonably close for the 70-, 90-, and 110-msec ISIs and no-mask condition.

This pattern of results can be simply interpreted by considering the accuracy differences on the right positions as a result of a response constraint operating on the CVC-SL

set. The frequent position errors for this set indicated that either the filler letters (Ws) were not always discriminable from the letter to be reported or that on some trials a filler letter looked more like a response letter than the one actually presented. Thus, except for ISIs of 70 msec or greater and the no-mask condition where position errors were infrequent, performance on the CVC-SL set was artificially low. The obtained underestimates of performance on the left and middle positions of CVC-LMC for ISIs of 50 msec or less are consistent with this interpretation.

It should be emphasized that, except for the accuracy data in Fig. 3 for the right position of CVC-LMC, the trigram data represent not simply discriminability, but the result of an interaction between discriminability and redundancy. This point can be clarified by considering the data in Fig. 2 which show a large difference between CVC-COND and CVC-LMC for the proportion of trials on which all letters were correct. It can be assumed that, on any trial, from zero to three letters were perceived, or alternatively, a correct decision was made about the letters in none, some, or all positions. Table 2 illustrates the probability of correctly reporting all letters in the remaining positions for all possible states of the perceptual system. The expected outcome is the same for both stimulus sets for all possible states of the perceptual system except for the case where two letters are perceived. The overall probability of reporting the remaining letter of a CVC-COND trigram, given that any two were perceived, is unity, while the comparable value for a trigram from Stimulus Set CVC-LMC is 7/9.

Since the zero-, one-, or three-letter cases lead to the expectation of no difference between performance on these sets for this measure, the obtained differences strongly suggest that two letters were perceived on a large proportion of trials. If a correct decision was made about the letters in

Table 2
Conditional Probability of Correctly Reporting All Letters, Given that Zero, One, Two, or Three Letters Were Perceived, as a Function of Form of Redundancy

	Stimulus Set					
	CVC-COND			CVC-LMC		
Letter in position perceived	None			None		
Probability of reporting all letters	1/9			1/9		
Letter in position perceived	Left	Middle	Right	Left	Middle	Right
Probability of reporting all letters	1/3	1/3	1/3	1/3	1/3	1/3
Letters in positions perceived	Left and middle	Left and right	Middle and right	Left and middle	Left and right	Middle and right
Probability of reporting all letters	1	1	1	1/3	1	1
Letters in positions perceived	All			All		
Probability of reporting all letters	1			1		

the left and middle positions of CVC-COND, the letter in the right position was maximally constrained. The letter in the right position was orthogonal with respect to those in the left and middle positions for CVC-LMC trigrams and was not constrained.

DISCUSSION

Performance on Set CVC-LMC was reasonably well predicted from performance on Set CVC-SL by a two-state model of perceptual independence. The slight overestimate of joint hit rate is characteristic of the two-state model. As Eriksen (1966) and Garner and Flowers (1969) have pointed out, the expansion of the two-state model to a multistate model by assessing confidence ratings will appreciably reduce the overestimate of joint hit rate.

The gain in accuracy due to correlational redundancy found in the present experiment is interesting in that the previous investigations of perceptual independence employed multiple presentations of the same letter. That is, not only were the stimuli correlated, but the multiple opportunities to observe the stimulus were obtained by repeating the same letter. In the present experiment, a comparable gain in accuracy due to multiple opportunities to observe the stimulus was obtained without repeating the same letter. Apparently, the correlation between stimuli was sufficient.

When compared to the correlational form of redundancy, conditional

redundancy is an optimum form in that it increases the frequency of correct responding for the visual pattern as a whole.¹ It would thus appear that the form of redundancy found in actual words is functional for rapid reading. This suggestion must be tempered with some additional considerations. It is not clear that the outcomes from an experiment involving tachistoscopic exposure can be generalized to tasks where speed of discrimination is the appropriate measure. Flowers and Garner (1971) have argued convincingly from data derived from speeded classification tasks that redundancy increases speed of discrimination only when the stimuli have low contrast or some energy limitation.

However, reading behavior involves a chain of brief fixations interspersed with short voluntary eye movements. Volkman (1962) has reported elevated thresholds for acuity during voluntary saccades. Further, portions of a line of text fall on the periphery of the retina during a fixation. It would thus seem that elevated thresholds during a saccade and the reduced acuity of peripheral vision would be expected to momentarily decrease discriminability. Under these conditions, redundancy can be expected to improve discrimination.

Recognition accuracy was but slightly affected by meaning and pronounceability. This contrasts with the findings of Reicher (1969), where a large "word" effect was found for a task requiring forced choice recognition of a single letter that was

part of a word or anagram. Apparently, when the number of stimulus alternatives are few and well learned, as in the present experiment, differences in recognition between words and nonwords all but disappear.

REFERENCES

- DOHERTY, M. E., & KEELEY, S. M. A Bayesian prediction of four-look recognition performance from one-look data. *Perception & Psychophysics*, 1969, 5, 362-364.
- ERIKSEN, C. W. Independence of successive inputs and uncorrelated error in visual form perception. *Journal of Experimental Psychology*, 1966, 72, 26-35.
- ERIKSEN, C. W., & LAPPIN, J. S. Independence in the perception of simultaneously presented forms at brief durations. *Journal of Experimental Psychology*, 1967, 73, 468-472.
- FLOWERS, J. H., & GARNER, W. R. The effect of stimulus element redundancy on speed of discrimination as a function of state and process limitation. *Perception & Psychophysics*, 1971, 9, 158-160.
- GARNER, W. R. *Uncertainty and structure as psychological constructs*. New York: Wiley, 1962.
- GARNER, W. R., & FLOWERS, J. H. The effect of redundant stimulus elements on visual discrimination as a function of element heterogeneity, equal discriminability, and position uncertainty. *Perception & Psychophysics*, 1969, 6, 216-220.
- REICHER, G. M. Perceptual recognition as a function of meaningfulness of stimulus materials. *Journal of Experimental Psychology*, 1969, 81, 275-280.
- VOLKMAN, F. C. Vision during voluntary saccadic eye movements. *Journal of the Optical Society of America*, 1962, 52, 571-578.

NOTE

1. This relationship was anticipated by W. R. Garner (1962, p. 166).

(Received for publication June 26, 1972.)