Perception & Psychophysics
1987, 42 (5), 411-422

Discriminability and bias in the word-superiority effect

W. K. ESTES and JENNIFER L. BRUNN
Harvard University, Cambridge, Massachusetts

Functions relating four-alternative, forced-choice recognition of target letters to exposure duration were obtained in word (W), pronounceable-nonword (PN), and unpronounceable-nonword (UN) contexts. Advantages were found for letters in W over letters in PN contexts and for letters in both W and PN over letters in UN contexts, with uniformly large effects when displays were terminated by pattern masks but small or absent effects when there was no postmask. Analyses in terms of models that allow separate estimates of discriminability and bias effects showed that both contributed strongly to the advantage of W and PN over UN contexts with generally close parallelism of discriminability and bias measures over types of context, mask condition, and display duration. The findings are interpreted in terms of a two-stage conception of letter processing, with the effects of context localized in the competition among character representations for transition from the early parallel stage (visual buffer) to the later serial stage (working memory). The better recognition of letters in words and pronounceable nonwords is attributed to their advantage in competition for access to working memory, which is conferred by the encoding of familiar letter groups as units.

We were concerned in this experiment with the nature of the word advantage in letter recognition, that is, the fact that a target letter embedded in a word is generally recognized better at brief exposures or under conditions of poor visibility than the same letter embedded in a nonword or presented alone (Baron, 1978; Smith & Spoehr, 1974). Prior to the late 1960s, it was generally assumed that the word advantage was attributable to the ability of readers to use redundancy in the letter sequences of words and knowledge of the statistical structure of the language to fill in gaps in perception with sophisticated guessing. However, the studies of Reicher (1969) and Wheeler (1970) seemed to rule out that interpretation by demonstrating the word advantage in the presence of controls for redundancy. The control was achieved by arranging display sequences so that when a subject's task was to indicate which of two alternative target letters was present in a displayed letter string, either one of the two alternatives would complete a word if the display was a word and either one of the two alternatives would cause the string to remain a nonword if the display was a nonword. However, this control is not always compatible with other properties desired of an experimental design, and, furthermore, a cogent argument has been advanced that the control may not always fully achieve its purpose (Massaro, 1973).

In this experiment, we included comparisons of the Reicher (1969) type, but we placed heavier reliance on the converging results of different types of data analysis, each designed to yield information about the effects of context on an observer's ability to discriminate among alternative targets and about effects on the observer's decision criteria or response biases. The experiment was designed to allow a variety of analyses. We first adapted standard statistical models to the task, then did the same for theoretical models of the signal-detectability family. A major advantage of the former approach is the weakness of the necessary assumptions. But in compensation for requiring stronger assumptions, the latter approach yields theoretically meaningful measures that can enter into quantitative comparisons across conditions and even across experiments. We applied both types of models to an experiment in which we varied over a wide range parameters that have been shown in earlier work to be strongly relevant to the manifestation of the word advantage.

One of the most important design characteristics of the study has to do with the nature of the task. To be able to apply the models of interest, a forced-choice procedure must be employed, using the same set of target alternatives over an extended series of trials. Unfortunately, forced-choice recognition with a constant set of alternatives is the condition that has been most consistently associated with failures to obtain the word advantage (Bjork & Estes, 1973; Massaro, 1973; Thompson & Massaro, 1973). It has been suggested that with this procedure subjects might under some conditions be able to accomplish the recognition task on the basis of decisions made concerning the presence or absence of individual visual features rather than on full identification of letters (Estes, 1975, 1977). We adopted two measures in this experiment to mitigate that possibility. One measure was to use four, rather than the more customary two, choice alter-

This research was supported by Grants BNS 80-26656 and BNS 84-19021 from the National Science Foundation and by a fellowship from the John Simon Guggenheim Foundation to the first author. Reprint requests should be addressed to W. K. Estes, William James Hall, Harvard University, 33 Kirkland Street, Cambridge, MA 02138.

Copyright 1987 Psychonomic Society, Inc.

natives to increase the difficulty of making decisions on the basis of single features, and the other was to require the subjects on each trial not only to make a choice of the target alternative present in the display but also to attempt to report the full letter string presented, a task that obviously required full identification.

To keep the task orientation constant over all of the types of displays to be compared, letter strings constituting words (W), pronounceable nonwords (PN), and unpronounceable nonwords (UN) were randomly mixed over series of trials in which the set of letter alternatives for the recognition task was fixed. Half of the trials of a series were assigned to the W category and half to the combined PN and UN categories so that subjects could not improve their accuracies by guessing letters that completed words when they were unsure of the identity of the target.

A second major design characteristic had to do with conditions of visibility: exposure duration and the presence or absence of a postmask. In preceding work, failures to obtain the word advantage have been most conspicuously associated with presentations of letters for very short exposure durations without a mask (e.g., Johnston & McClelland, 1973). The entirety of the present experiment was replicated under mask and no-mask conditions. Independent groups of subjects were used for the two replications, since substantially different exposure durations had to be used for the mask and no-mask conditions to keep recognition accuracies within the same range. However, in view of the necessary confounding between exposure duration and mask versus no mask, a sequence of three exposure durations was used for each group. Thus it was possible to distinguish the effect of the mask on visibility from effects that were independent of visibility.

METHOD

Subjects

Twenty college-age adults with normal or corrected-to-normal vision served as subjects. All were new to the task and were paid $4/h for their participation.

Apparatus

Stimuli were presented on a Textronix 604 monitor screen interfaced to a PDP 11/10 computer. The screen was rectangular, 10 cm wide × 6.5 cm high. A five-letter display subtended approximately 3.5° arc, and interletter spaces 0.17°.

Stimulus Materials

All stimuli were strings of five uppercase letters centered on the screen in a horizontal row. In the mask condition, the display was followed by a row of &s occupying the same positions as the letters.

Two sets of 32 strings were prepared; each string included 16 words, and 8 each of pronounceable and unpronounceable nonwords. The pronounceable nonwords were orthographically regular and easily pronounceable, the unpronounceable nonwords were orthographically irregular and difficult or impossible to pronounce. In one set, each string contained exactly one of the target letters F, L, N, or P (FLNP letter set) and, in the other set, each string contained exactly one of the target letters D, G, K, or T (DGKT letter set). The target letter was located in Position 1, 2, 4, or 5 of the string, and the positions were balanced over letter sets and string types.

Each set of 32 strings comprised eight subsets, or modules, of four, each including two words: one pronounceable nonword and one unpronounceable nonword, as in the following example.

LEAST
FEAST
NEAST
PKAST

The members of a subset of four always shared a three-letter stem in either the first three or the last three letter positions, and each of the four target letters of a set appeared in one of the strings of the subset.

Design

Independent groups of 10 subjects were assigned to the mask and no-mask conditions, but both groups saw the same displays. Each subject viewed the 32 strings of each letter set at short, intermediate, and long exposure durations. The durations were set at 10, 15, and 20 msec for the no-mask condition and 35, 45, and 55 msec for the mask condition. These durations were selected by means of a pilot experiment to obtain the desired range of levels of accuracy within each condition and the roughly comparable levels of accuracy between the two conditions.

Half of the subjects in each group viewed the FLNP letter set first and the other half viewed the DGKT letter set first; the order of the strings in each set was randomized independently for each subject. A subject first saw the assigned 32-item set at the shortest duration, then at the intermediate duration, and then at the long duration, each time in a different random order.

The plan of presenting the same displays at successive durations was used for two reasons. First, we wished to trace the formation of W and PN units as additional information accrued with increasing duration. (The unit data are being prepared for a separate report, along with a companion study.) Second, we wanted maximal sensitivity for the comparisons of context effects at different overall performance levels. To deal with the resulting confounding of duration with number of repetitions of a given stimulus string, the mask condition was replicated with an additional 6 subjects for whom duration and repetition were combined orthogonally in a Latin square design (the duration control group).

Procedure

A session for a subject comprised 192 experimental trials plus an eight-trial practice series at the intermediate duration prior to the experimental series on each 32-item set. Before each trial block, the subject was informed of the target-letter set for the block. The trials were self-paced, initiated by the subject's keypress. The first event on the trial was the 500-msec display of a fixation dot in the center of the screen. Then the letter string appeared, centered just above the fixation dot for the assigned duration. In the mask condition, the letter display was followed by the pattern mask, consisting of a row of &s, for 50 msec and in the no-mask condition by a blank screen for the same interval. In each case, the screen was blank while the subject made a forced choice among the four possible target letters by pressing one of a set of four labeled keys, then wrote on an answer sheet his or her best guess as to the whole unit shown on the trial (that is, the whole word, pseudoword, or nonword). The subjects were given a short break at the end of each 32-trial block at a given duration, and a somewhat longer break between the first and second sets.

RESULTS

Correct Response Percentages

Only the target-recognition data will be treated in this report. For comparability with previous studies, we note first the principal trends in correct target-recognition percentages, summarized in Table 1. As anticipated on the basis of previous related studies, the differences between W, PN, and UN contexts produced very large effects in the mask condition, but much smaller effects in the no-mask condition. For a statistical assessment, the recognition percentages were subjected to an analysis of variance (ANOVA), with all factors except subjects considered to exert fixed effects. The main effect of letter set was nonsignificant [$F(1,18) = 1.58$]; thus we infer that our conclusions need not be limited to the particular target sets used and we collapse over this variable in all other analyses. The main effect of mask versus no mask was also insignificant, indicating that we had met our objective of roughly equating the overall accuracy levels for the two conditions.

The principal effects of theoretical interest—the W, PN, UN main effect and its interaction with mask/no mask and the duration main effect—were all significant at the .001 level [Fs(2,36) = 46.2, 14.4, and 56.6, respectively]. The interactions with duration were nonsignificant, indicating that the growth of recognition accuracy with exposure duration proceeded essentially in parallel for the three display types and two mask conditions. Planned comparisons between display-type means yielded significant differences between W and PN and between PN and UN for the mask condition ($t = 7.7, p < .01$, and $t = 2.47$, $p < .02$, respectively, with 36 *df*). For the no-mask condition, W-PN approached significance ($t = 2.1$, $p < .05$), and only the W-UN difference reached significance ($t = 2.92$, $p < .01$).

At individual durations, in the mask condition, W versus PN was significant at the .01 level at all durations, but PN versus UN was significant only at Duration 2 ($p < .02$); none of the differences approached significance in the no-mask condition.

In summary, all three independent variables exerted large effects, and in a pattern conforming to expectations from previous studies based on various subsets of the overall design. Our main task now was to employ statistical and model-based analyses of the complete choice data to determine which effects or combinations of effects were attributable to variations in discriminability between letters and which were attributable to variations in the observer's criteria or biases relative to linguistic properties of the letter contexts.

Complete Choice Frequency Data

Statistical analyses. Percentages of choices of each type of response to targets presented in each context are given in the *observed* columns of Table 2. The qualitative effects of context on both discriminability and bias can be grasped intuitively simply by inspection of the table. Correct-response percentages are on the main diagonals of the 4×4 subtables, so discriminability of a target letter from the available alternatives is indexed by the differences between a diagonal element and the off-diagonal elements of its row. Thus, the overall trends are (1) uniform increases in discriminability with exposure duration, and (2) discriminability differences favoring W over PN over UN contexts, the differences being greater and increasing more steeply with duration in the mask condition.

Biases, or criterion settings, that favor letters that complete words or pronounceable nonwords might be expected to show up in either of two ways. Since a bias for choosing, say, a target letter that completes a word in a given context would increase the percentage of occurrences of that response regardless of the stimulus, the effect would be to increase the sum of values in the corresponding column of Table 2. There is very little variation among column sums in the no-mask condition, but in the mask condition, the sums tend to line up in the order W > PN > UN, and do so increasingly with increasing duration. Biases of this kind might also show up as asymmetries between the upper right and lower left entries in a subtable since, for example, a tendency to complete words would increase the frequency of responses in the R1 and R2 columns to PN stimuli over the frequency of responses in the R3 column to W stimuli. The data in Table 2 show little departure from symmetry in the no-mask condition, but marked asymmetry in the mask condition, especially at the longer durations.

The significance of these trends was evaluated by means of a log-linear ANOVA (Bishop, Fienberg, & Holland, 1975; Brier, 1982). The cell frequencies corresponding to the proportions in Table 2 were transformed to logarithms and the resulting values were entered in separate ANOVAs for the mask and no-mask conditions. The model for the analysis was

$$Y_{ijk} = u + u_i + u_j + u_k + u_{ij} + u_{ik} + u_{jk} + e_{ijk}, \quad (1)$$

where Y_{ijk} is log frequency for cell ijk; u is the population mean; the subscripted us denote the main effects and interactions; and e_{ijk} represents normally distributed error.[1] The subscripts i, j, and k index the three levels of exposure duration, the four levels of context (W_1, W_2, PN, UN), and the four response categories, respectively. This mode of analysis was chosen over one based on χ^2 tests because of its flexibility in providing for various contrasts of theoretical interest (Grizzle, Starmer, & Koch, 1969). Row effects are of no interest, since, except for a few missing observations, row frequencies add to a constant, but column effects reflect response biases and the row-column interactions (context × response category) reflect discriminability effects.

The overall Fs support the impressions sketched above regarding effects in both conditions. In the mask condition, the column effect and the row-column interaction were significant at the .001 level [$F(3,18) = 8.47$ and

Table 1
Mean Correct Target Recognition Percentages

Duration	Mask W	Mask PN	Mask UN	No Mask W	No Mask PN	No Mask UN
1	68	52	53	46	42	43
2	84	68	56	62	58	54
3	94	73	67	79	73	70

Table 2
Observed and Predicted Percentages of Target Responses* to Each Display Type

Duration	Display Type	Observed				Theoretical			
		R_1	R_2	R_3	R_4	R_1	R_2	R_3	R_4
		Mask Condition							
1	W	68	13	10	9	69	13	9	9
	W	14	69	8	10	13	69	9	9
	PN	17	15	52	16	16	16	52	16
	UN	13	18	16	53	16	16	16	53
2	W	82	8	5	5	84	7	6	3
	W	6	86	2	5	7	84	6	3
	PN	16	11	69	4	13	13	69	6
	UN	9	19	16	56	15	15	15	56
3	W	94	4	1	1	94	4	2	0
	W	3	94	2	1	4	94	2	0
	PN	17	8	74	1	12	12	74	2
	UN	11	13	10	67	12	12	10	66
		No-Mask Condition							
1	W	43	26	16	15	46	22	16	16
	W	18	49	15	19	22	46	16	16
	PN	15	24	42	18	19	19	42	20
	UN	16	20	21	43	19	19	19	43
2	W	63	13	12	13	63	14	12	11
	W	16	62	14	8	14	63	12	11
	PN	18	11	58	14	14	14	58	13
	UN	16	14	16	54	15	15	16	54
3	W	81	6	6	6	80	8	8	5
	W	9	78	9	4	8	80	8	5
	PN	14	6	73	7	10	10	73	7
	UN	10	10	10	71	10	10	10	71

*R_i, for $i=1$–4, correspond to the Response headings in Table 4.

$F(9,18) = 20.68$, respectively, $MS_e = .228$]. In the no-mask condition, the column effect was nonsignificant ($F < 1$), but the row-column interaction was significant [$F(9,18) = 11.9, p < .001, MS_e = .194$]. The principal overall result, then, is confirmation of a bias favoring responses that complete PN or W units in the mask but not in the no-mask condition.

To address the question of differences in discriminability among target letters presented in different contexts, we turn to the appropriate contrasts. The discriminability of a target letter from alternative targets is measured in this analysis by a contrast within a row of Table 2, taking the form $3m_{ii} - \Sigma_{j \neq k} m_{jk}$, where m_{ii} is the diagonal element (the correct response frequency) and m_{jk} is an off-diagonal element (error frequency). The squared standard error of a contrast is the MS_e from the ANOVA multiplied by the sum of the squared contrast weights $[3^2+(-1)^2+(-1)^2+(-1)^2]$, yielding 3.22 for the mask and 2.34 for the no-mask condition. The contrast for each row of Table 2 is given in Table 3; the values listed are contrasts divided by their standard errors, which may be interpreted as t statistics with 18 *df*. These contrasts confirm the impression that discriminability (1) increases with exposure duration, (2) is higher for the mask than the no-mask condition at all durations, and (3) lines up uniformly in the order W>PN>UN, except for small inversions between PN and UN at Duration 1.

Differences in discriminability between contexts are measured by differences between contrasts, and these yield essentially no evidence of context differences in the no-mask condition. Differences begin to appear, however, at Duration 1 and are significant for all pairs of contrasts at Duration 3 in the mask condition.

The two ways of assessing biases can both be expressed in terms of contrasts on the log response frequencies. One method is to define biases toward completing W, PN, and UN units by contrasts on the column means of Table 2. Weights of 2, 2, −1, and −3 for the R_1, R_2, R_3, and R_4 means, respectively, appear suitable to reflect any bias toward choosing target letters that complete more word-like units. These weights yield bias contrasts of 1.12, 2.78, and 5.70 for Durations 1, 2, and 3, respectively, in the

Table 3
Discriminability Contrasts on Log Response Frequencies

Duration	Condition					
	Mask			No Mask		
	W	PN	UN	W	PN	UN
1	3.19	1.99	2.07	1.85	1.56	1.61
2	4.73	3.45	2.33	3.22	2.82	2.50
3	6.76	4.32	3.01	4.88	4.22	3.90

Note—The values listed are contrasts divided by their standard errors, obtained from ANOVAs. The W contrast in each case is the average taken from the values obtained in the two W rows of the subtable of Table 2.

mask condition and .44, .45, and 1.14 in the no-mask condition. Standard errors of these contrasts, based on 18 *df* in each case, are 1.27 for the mask and 1.08 for the no-mask condition. Thus the bias effect is nonsignificant at all durations in the no mask, but becomes increasingly significant with duration, reaching the .001 level at Duration 3 in the mask condition.

Alternatively, biases may be assessed by contrasts between the six values below and the six values above the main diagonal of each subtable of Table 2:

$$C_{bias} = \sum m_{jk} - \sum m_{kj}, \text{ where } j > k. \quad (2)$$

These contrasts are 2.27, 5.76, and 11.11 for Durations 1, 2, and 3, respectively, in the mask condition and .31, 1.33, and 2.33 in the no-mask condition. In the mask condition, the bias effect is only suggestive at Duration 1, but is significant at the .001 level at Durations 2 and 3; in the no-mask condition, bias is negligible at Duration 1 and barely reaches marginal significance ($p < .05$) at Duration 3. Clearly, the two ways of assessing bias are in satisfactory agreement.

Model-based analyses. The model that will serve as an additional basis for analyzing the data of the experiment is developed within the framework of the choice model of Luce (1963). The model is illustrated in Table 4 in the format applicable to the recognition data for a module of four displays sharing a common stem. A list of four-letter strings that share a common stem is given at the left and the four admissible target alternatives for the trial series in which this set of strings was used, P, N, L, and F, are entered at the tops of the columns. We assumed that when a particular target, say the letter P in the first item listed, appeared in the display shown on a trial, the tendency for the observer to identify it as any one of the four alternatives was given by the product of a discriminability parameter, η, and a bias parameter, β. The parameter η is a measure of the similarity between the target letter and the mental representation of the letter corresponding to the response. We use the term *similarity*, rather than *discriminability*, to conform with Luce (1963); the value of η ranges between 0 and 1 with the value 1 representing complete similarity (no discrimination) and the value 0 complete dissimilarity (perfect discrimination).

Table 4
Model for Analysis of Recognition Data in Four-Item Modules

Display	Response 1 (P)	2 (N)	3 (L)	4 (F)
1 Word (*speak*)	β_1	$\eta_1\beta_1$	$\eta_2\beta_2$	$\eta_3\beta_3$
2 Word (*sneak*)	$\eta_1\beta_1$	β_1	$\eta_2\beta_2$	$\eta_3\beta_3$
3 Pronounceable Nonword (*sleak*)	$\eta_2\beta_1$	$\eta_2\beta_1$	β_2	$\eta_4\beta_3$
4 Unpronounceable Nonword (*sfeak*)	$\eta_3\beta_1$	$\eta_3\beta_1$	$\eta_4\beta_2$	β_3

Note—Illustrative displays and responses are shown in parentheses.

The parameter β represents the observer's bias for choosing a target letter associated with a particular column, independently of the target actually present. The specific definitions of the parameters included in Table 4 are as follows:

Parameter	Description
η_1	Similarity between a letter presented in a W context and an admissible alternative that completes a W with the same stem.
η_2	Similarity between a letter presented in a W context and an admissible alternative that completes a PN with the same stem (or vice versa).
η_3	Similarity between a letter presented in a W context and an admissible alternative that completes a UN with the same stem (or vice versa).
η_4	Similarity between a letter presented in a PN context and an admissible alternative that completes a UN with the same stem (or vice versa).
β_1	Bias for a target letter that completes a word in a given context.
β_2	Bias for a target letter that completes a pronounceable nonword in a given context.
β_3	Bias for a target letter that completes an unpronounceable nonword in a given context.

We should emphasize that our interpretation of the bias parameter differs in an important way from the standard practice associated with applications of Luce's (1963) choice model. In the standard form of the model, this parameter denotes the observer's bias for making a given response independent of the stimulus; thus, for example, β_2 in Column 3 of Table 4 would represent simply the observer's bias for using the letter L as a response regardless of the stimulus (see, e.g., Luce, 1963; Pachella, Smith, & Stanovich, 1978; Townsend, 1971). Our usage, however, differs from the standard by defining the bias as the observer's tendency to make any response that has a specified relation to the context in which the target letter appears (a type of definition introduced by Estes, 1982). In our usage, β_2 is not a bias for the letter L as a response, but, rather, a bias for any admissible letter that completes a pronounceable nonword if assigned to the position of the target letter in the given display.

Response probability in this model is defined as the value in a cell divided by the sum of the values in its row. Thus the predicted probability of the response P to the stimulus display SPEAK in the example of Table 4 would be $\beta_1/(\beta_1+\eta_1\beta_1+\eta_2\beta_2+\eta_3\beta_3)$. Our general procedure in applying the model will be to obtain data matrices corresponding in form to Table 4, then by means of a computer program to find values of the parameters that, according to a least squares criterion, yield the best description of the data. Assuming that the model proves to fit the data satisfactorily, our primary interest will be in relating patterns of parameter values to the issues of interest in the experiment. For example, a facilitating effect of a word context on the discriminability of constituent letters would imply smaller values for η_1, η_2, and η_3 than for η_4 (small values of η denoting high degrees of discriminability); a bias for choosing a letter that completes a high-order unit (W rather than PN, PN rather than UN) in the given context regardless of the target actually present would be reflected in an order of the β values $\beta_1 > \beta_2 > \beta_3$.

The model was fitted to the data of the mask and no-mask groups separately; the six parameters were evaluated for each 16-observation module by minimization of squared differences between observed and theoretical values. The goodness with which the model, with the parameters so evaluated, reproduces the data may be seen in Table 2 and seems quite satisfactory. We conclude that the parameters of the model enable reproduction of the data to a good approximation, and thus the pattern of parameter values becomes the question of primary theoretical interest.

To facilitate interpretation, we have transformed the values of the estimates of the similarity parameter η to a scale of psychological distance by means of the function $d = -2 \log \eta$, which has become fairly standard for this purpose (Noreen, 1979; Nosofsky, 1984; Shepard, 1958), then multiplied the resulting d values by .616 to put the values on the familiar scale of d' of the signal detectability theory (Green & Swets, 1966). The d values so obtained are summarized in Table 5 and may be taken to represent psychological distances between to-be-discriminated target letters in the various context conditions. A large distance signifies a high degree of discriminability and a 0 distance signifies complete lack of discriminability. The overall picture is one of little variation in the d values across W, PN, and UN contextual conditions for the no-mask group, but of a trend at Duration 1 increasing to large effects at Durations 2 and 3 for the mask group.

The specific results on discriminability are larger ds for the mask condition under all combinations of context and duration and quite uniformly smaller ds for the PN-UN column than for those including W in the discrimination. The W-W, W-PN, and W-UN columns exhibit no uniform trend.

To provide some statistical evaluations, an ANOVA was conducted on the d estimates. The effects of mask/no mask and duration proved significant at the .001 level [$F(1,6) = 225.4$, $F(2,6) = 290.2$, respectively], and variation among the contexts reached the .01 level [$F(3,6) = 10.1$], but no interactions approached significance. As in ANOVAs throughout this report, mask/no mask and duration were treated as crossed factors, so that their interactions could be evaluated. If, instead, duration is treated as nested within mask/no mask, none of the conclusions about significance are altered, but the effects of duration within mask and no-mask conditions can be evaluated separately. Both prove significant at $p < .001$ [$Fs(2,12) = 118.9$ and 86.7, respectively]. If, rather than treating duration as a categorical variable with levels 1, 2, and 3, we use the actual durations (35, 45, and 55 for mask and 10, 15, and 20 for no mask) and carry out a multiple regression, the sums of squares for duration are virtually unchanged—8.45 and 6.16 for effect of duration within mask and no mask, respectively, in the ANOVA, and 8.40 and 6.12 in the regression analysis. Evidently, our effort to accomplish a functional match of duration values in the two conditions was quite successful, since the effect of a 10-msec increment in duration under the no-mask condition was virtually equal to the effect of a 20-msec increment in the mask condition.

The bias estimates in Table 5 show uniformly higher bias for letters that complete words and also sharper divergence among biases for W, PN, and UN with duration in the mask than the no-mask condition. An ANOVA on the bias estimates yields significance for differences among contexts [$F(2,4) = 21.0, p < .01$] and marginal significance for the context × duration interaction [$F(2,4) = 7.7, p < .05$].

The results of these model-based analyses agree in major respects with those of the log-linear analysis. In the no-mask condition, there are no effects of context on either discriminability or bias. In the mask condition, discriminability differences favoring target letters presented in W or PN contexts are suggestive at Duration 1 and become increasingly significant with increasing duration,

Table 5
Estimates of Discriminability (d_i) Between Alternative Targets and Biases (β_i) for Target Letters Associated with W, PN, and UN Contexts

	Discriminabilities				Biases		
	W-W	W-PN	W-UN	PN-UN	W	PN	UN
Duration	d_1	d_2	d_3	d_4	β_1	β_2	β_3
			Mask				
1	2.03	1.99	2.01	1.47	.30	.20	.20
2	3.00	2.71	2.94	2.32	.34	.20	.12
3	4.03	3.51	4.63	3.51	.40	.14	.05
			No Mask				
1	.90	1.13	1.15	.96	.27	.23	.24
2	1.83	1.85	1.89	1.66	.27	.24	.21
3	2.90	2.67	2.88	2.69	.28	.23	.20

and a similar trend holds for biases. Although d values for W-UN do not differ significantly from W-PN at any duration, the differences are larger for W-UN in every case (both mask and no mask). Similarly, the difference in bias for targets that complete W versus PN units approaches significance only at Duration 3 of the mask condition, but the difference favors W over PN in every case.

To supplement the overall analysis, the design provides a completely balanced set of comparisons for target letters presented in the various pairs of contexts, as shown in Table 6 (each target occurs equally often in each row of each 2×2 table). As a measure of discriminability we use the cross-product ratio, denoted α in Table 6, which is the basis of commonly used indices of association and correlation in 2×2 tables (Bishop et al., 1975). The natural log of α multiplied by .616 transforms this measure to the same d scale used in connection with Table 5. As a measure of bias, we use the ratio of column products (*CR*), which can be transformed to a scale comparable to β of signal detectability theory by the equation $c = .5 \log CR$. A c equal to 0 signifies no bias, a negative value signifies a bias favoring R_1 over R_2; in all but the W_1, W_2 matrix, a negative c value may be interpreted as a preference for the response that completes the higher order unit of a pair (W over P, PN over UN). An estimator of the standard error of α (Bishop et al., 1975) enables us to perform z tests of the directional pairwise differences in α values. These tests yield significance at the .05 level or better for all pairs except W_1-W_2 versus W-PN, which supports the conclusion that discriminability progressively increases as targets are presented in progressively higher order units. The c values suggest biases toward completing W or PN over UN units, but we know of no suitable statistical test for the observed trend.

Additional information about discriminability in W contexts is available in two pairwise comparisons of special interest in the mask condition data. One of these comes from pairs of words such as *least* and *feast*, which have the property that each of the target letters (L and F) completes a word, so that discrimination between the targets could not be influenced by the predictability of the target letter from the remainder of the word (the Reicher [1969] control). A related comparison comes from pairs of words such as *knees* and *trees* (with target letters k and t), which have the property that substitution of the target letter from either word into the other word of the pair produces a nonword, so that utilization of redundancy should increase the likelihood of correct responding. Analyses of the kind given for Table 6 yield very similar discriminabilities in the two cases (ds = 2.85 and 2.87, respectively), supporting the conclusion from the choice-model analysis that discriminability is influenced by a word context independently of any effects of redundancy on sophisticated guessing or other control processes.

Table 6
Pairwise Analysis of Choice Frequencies by Contexts for Mask Condition

	R_1	R_2		R_1	R_2		R_1	R_2		R_1	R_2
W_1	93	9	W	93	12	W	97	8	PN	91	9
W_2	7	97	PN	10	91	UN	15	51	UN	40	51
α	143.2			70.5			41.2			12.9	
SE	53.2			22.6			14.1			3.8	
d	3.06			2.62			2.29			1.58	
CR	1.34			1.17			.28			.13	
c	.15			.08			−.64			−1.04	

Note—For each matrix, R_1 and R_2 denote the responses of choosing the target letters presented in the upper and lower row contexts, respectively. The parameter α is the cross-product ratio and $d = .616 \log \alpha$; the column ratio (CR) is the product of the R_2 divided by the product of the R_1 frequencies, and $c = .5 \log$ CR. Data are pooled over durations.

Table 7
Frequencies of Errors That Complete W, PN, or UN Units

Stimulus	Result of Error	Mask Condition Mean	Mask Condition %	No-Mask Condition Mean	No-Mask Condition %
W	W	2.21	63	4.06	58
	PN or UN	1.28		2.90	
PN	W	5.22	74	4.44	54
	PN or UN	1.87		3.73	
UN	PN	7.00	66	3.17	44
	UN	3.83		4.00	

Some additional information about bias, or criterion, effects can be obtained from analyses of the specific nature of errors of recognition. In the top row of Table 7 we give the mean frequencies with which errors in recognition of targets presented in Ws were choices of letters that completed Ws (2.21 for the mask condition) or that completed PNs or UNs (1.28 for the mask condition). In the second row the corresponding analysis is given for target letters presented in PNs and responses that complete either Ws (upper values) or PNs or UNs (lower values); in the third row the corresponding analysis is given for targets presented in UNs and responses that complete either PNs or UNs. In the data for the no-mask condition, the percentages of errors that complete the more wordlike units for each display type differ little from chance level (50%), and only for W displays is the difference even marginally significant ($p = .05$) by a chi-square test. In contrast, all of the percentages for the mask condition are substantially above chance and are significantly so at better than the .01 level by chi-square tests. Clearly, the utilization of redundancy in the context to influence the choice of responses when a target letter is imperfectly perceived is almost exclusively limited to the mask condition.

Results for Duration Control Group

Since the ordinal position of a display (that is, first, second, or third presentation of the display) is orthogonal to exposure duration for this group, the data can be analyzed with respect to each of these independent variables separately. For a number of the principal analyses reported above for the main experiment, we compare results of the duration control group with duration versus order as the independent variable (in addition to context in each case) and also compare the duration analyses with the corresponding results for the mask condition of the main experiment.

In Table 8, correct recognition percentages are given for the three contexts with either duration or order as the other independent variable. Differences among contexts are similar to those observed in the main experiment (Table 1) as are the trends over duration, shown in the left side of Table 8. The trends within columns are much shallower, however, with order as the independent variable, shown in the right side of Table 8. Thus the dependence of correct responding on duration observed in the main experiment evidently is due almost wholly to stimulus duration, rather than to any factor, such as number of repetitions of the display, that may be confounded with it in the case of the UN context; this conclusion needs to be only slightly weakened for the other two contexts. The full presentation of choice percentages is shown in Table 9, but only for the duration analysis since the shallow trends over order offer no interesting material for comparison with the main experiment or analysis in terms of the model. Discriminability contrasts on the log response frequencies were, however, computed for both the duration and order analyses and are shown in Table 10. Comparison with Table 3 shows that the pattern of contrasts for the duration analysis is very similar to that found for the mask condition in the main experiment. The only overall difference is a slight accentuation of trends both across contexts and over durations as compared to the main experiment. ANOVAs on the contrasts yield *F*s for context effects that are similar for both analyses (47.3 and 38.9 for the duration and order analyses, respectively, both significant at the .01 level), and both are somewhat larger than the corresponding *F* for the mask condition of the main experiment. With respect to row effects, duration yields an *F* of 41.8, significant at the .01 level, but order yields an *F* of 3.1, entirely nonsignificant.

Bias contrasts were computed on the column means of the log response frequencies as in the main experiment. The resulting values were 1.43, 2.58, and 4.07 for Durations 1, 2, and 3, respectively, and 1.21, 3.27, and 1.46 for ordinal positions 1, 2, and 3, respectively, with standard errors of 1.35 for the durations and .56 for the ordinal values. Clearly, the distinct increase in bias toward wordness observed over duration for the mask condition of the main experiment appears in the control group only for the duration analysis.

The choice model was applied to the context versus duration data of the duration control group in the same manner as for the main experiment and the resulting estimates of discriminability and bias are presented in Table 11. The trends for the discriminability estimates (*d* values) are generally similar to those seen for the main experiment in Table 5. On the average the values of d_2, d_3, d_4 line up in exactly the same order in the two cases. The only apparent difference is that the value of d_1, indexing the W-W discrimination, is a bit higher relative to the other distances in the case of the duration control group. The pattern of bias estimates is also similar in the duration control group and the main experiment, with bias tending to decrease from W to PN to UN and to do so more steeply with increasing duration.

Overall, the results for the duration control group indicate that the various trends apparently associated with display duration in the main experiment were indeed associated almost entirely with duration rather than with any variable confounded with duration. Evidently, number of repetitions of a display had virtually no measurable effect on target recognition in UN contexts but did exert a measurable effect, though not large when compared to duration effects, on recognition in PN and W contexts. As might have been expected on the basis of the results of Salasoo, Shiffrin, and Feustel (1985), there was at least some small tendency for letters in word and wordlike contexts to become increasingly available with repetition.

Table 8
Mean Correct Target Recognition Percentages for Duration Control Group

Duration	Context W	PN	UN	Order	Context W	PN	UN
1	73	52	43	1	78	60	58
2	84	72	66	2	84	71	57
3	90	80	68	3	85	73	60

Table 9
Observed and Predicted Percentages of Target Responses to Each Display Type for Duration Control Group

Display Type	Observed R_1	R_2	R_3	R_4	Theoretical R_1	R_2	R_3	R_4
				Duration 1				
W	75	9	8	5	73	8	10	8
W	5	71	12	12	8	73	10	8
PN	21	12	52	10	17	17	54	12
UN	12	27	17	43	20	20	17	43
				Duration 2				
W	85	4	6	4	85	4	6	5
W	4	83	2	9	4	85	6	5
PN	8	16	72	4	11	11	72	7
UN	12	8	12	66	12	12	11	66
				Duration 3				
W	93	5	1	1	90	5	2	2
W	5	88	2	3	5	90	2	2
PN	12	5	80	3	8	8	80	4
UN	12	12	6	68	13	13	6	68

DISCUSSION

Principal Findings

We used the extensive literature to choose task parameters that would be expected to yield robust word superiority effects. The effort was predictably successful, and we obtained substantial differences in the recognition of targets in W, PN, and UN contexts, in that order. Because of the way the subjects' task was structured, our results are primarily relevant to situations in which the

Table 10
Discriminability Contrasts on Log Response Frequencies for Duration Control Group

	Context				Context		
Duration	W	PN	UN	Order	W	PN	UN
1	3.67	2.25	1.46	1	3.98	2.75	2.50
2	4.89	3.64	3.17	2	4.80	3.58	2.38
3	6.08	4.42	3.22	3	4.79	3.47	2.50

Table 11
Estimates of Discriminability (d_i) and Bias (β_i) for Duration Analysis of Duration Control Group

	Discriminabilities				Biases		
Duration	W-W d_1	W-PN d_2	W-UN d_3	PN-UN d_4	W β_1	PN β_2	UN β_3
1	2.77	1.90	1.79	1.50	.32	.21	.16
2	3.62	2.78	2.86	2.54	.31	.22	.17
3	3.48	3.68	3.37	3.38	.35	.18	.12

individual is attempting not only to recognize individual target letters, but to identify words or other higher order units, as in reading.

Analyses of the choice data in terms of a statistical model enabled us to estimate the effects of context on the perceptibility of target letters separately from the effects of observers' criterion settings or biases. Perceptibility, so measured, was uniformly greater in W than in PN contexts and, except at the shortest duration, greater in PN than in UN contexts, with the differences growing with exposure duration and with all durations being much greater in the mask than in the no-mask condition. Measures of bias showed a preference for letters that complete more regular units (W over PN over UN), with the biases growing with exposure duration and being uniformly larger in the mask condition.

The statistical analyses of perceptibility bear on an aspect that might for brevity be termed absolute discriminability—the ability of the observer to discriminate a target letter from any admissible alternative. The analysis in the framework of the choice model, or signal detectability theory (essentially equivalent for our purposes), brings out another aspect of perceptibility that might be termed pairwise discriminability. By this term, we refer to an observer's ability to discriminate between the target letter included in a display and a specific alternative that might have occurred in its place. Relative discriminability was found to depend appreciably on context in the mask condition only. The nature of the dependence was that lowest discriminability occurred between a displayed letter in a PN and an alternative that would complete a UN, or vice versa, and that higher discriminabilities occurred for W-W, W-PN, and W-UN pairs. The most closely controlled comparisons (Table 6) indicate that discriminability of a target letter from an alternative in a given context increases with the regularity of the unit (W, PN, or UN) completed by each of the letters in that context.

Interpretation of Context and Mask Effects

Our interpretation of the findings will be organized within the framework of a sequence of processing stages that has become rather widely accepted in outline (Coltheart, 1984; Duncan, 1980; Estes, 1978; Mewhort, Marchetti, Gurnsey, & Campbell, 1984). These stages, not necessarily entirely discrete, start with the input of stimulus information into a virtually unlimited-capacity visual buffer, in which featural or spatial frequency information is registered in parallel. Character recognition operates on the contents of this buffer and generates character representations in a relatively abstract code that includes positional information. These representations are held in a slowly decaying character buffer of large, although probably not unlimited, capacity. Finally, the characters that meet criteria set by task demands may be passed through a limited-capacity channel into a response buffer, or working memory, where they are available for verbal report.

We start with the interpretation of mask effects, because mask condition may well be said to be the dominant independent variable in this study. Quantitatively, its effects were very large and robust, and the effects of other variables, especially context, turned critically on presence versus absence of a postmask. On the basis of the accumulated literature on masking (Breitmeyer, 1984), we can assume that the character mask that followed each display in our mask condition must have exerted two distinct kinds of effects. One is peripheral-noise masking, which occurs when stimulus input from the mask summates with stimulus input from the letter display and thus degrades the quality of information available for letter recognition. A patterned postmask, however, also has a central effect, which would operate to interfere with selection of character representations in the character buffer and with their passage into working memory. In all studies that include both mask and no-mask conditions and that require comparable levels of performance in each, it is necessary to use longer exposure durations for the stimulus display in the mask condition to compensate for the degrading effects of noise masking on stimulus quality. Consequently, the observer's task is distinctively different under mask and no-mask conditions. In the no-mask condition, the observer is trying to identify letters under conditions of extremely limited information quality with insufficient featural information about both target and context letters for full identification. Under the mask condition, however, relatively complete featural information has been registered in the first stage of processing for both target and context, and the observer's task is to recognize the target in spite of interference from mask input. Subsets of letters that have frequently occurred in the individual's experience as components of spelling groups, syllables, or words will tend to be encoded as units (Adams, 1979; Smith & Spoehr, 1974) and will be favored for passage into working memory as a consequence of the overall orientation of the processing system toward the identification of higher units.

Now several questions arise. First, why do data show no bias in favor of completing higher units in our no-mask condition? There is no reason to think that the tendency to complete higher units would be shut off in the absence of a mask, so the answer must be that under the no-mask condition the tendency simply cannot become overt. The obvious reason is that the very short exposure durations in the no-mask condition do not allow an incompletely perceived target letter to be accompanied by enough identifiable elements of the context so that the bias toward the completion of syllables and words can operate.

The next question is when, under mask conditions, a target letter embedded in a W or PN context should be better identified than one embedded in a UN context. In our conception, processing in the first stage is equivalent for W, PN, and UN displays, and consequently if a decision can be based only on the presence or absence of a particular feature, which seems to be the case in some two-alternative, forced-choice designs, there should be no word or pseudoword advantage—a frequently observed result (Bjork & Estes, 1973; Estes, 1975; Estes, Bjork, & Skaar, 1974; Thompson & Massaro, 1973). When actual identification of target letters is required, however, the task can only be accomplished by processing of letter representations in working memory, which is evidently the locus of word and pseudoword advantages.

A third question concerns the reason for the parallelism of measures of discriminability and bias. The answer appears to be that they are actually causally linked. Although a word or pseudoword context does not directly influence the registration of features or the recognition of individual letters, it does set the stage for a bias process to select letter groups over unrelated letters for passage to later stages of processing.

Our results on relative discriminability, revealed by the model analysis, suggest a processing account of the way target identification may depend on this selective process. When the mental representation of a display generated in the character or response buffer contains too little information for identification of some individual letter, in particular the target letter in our experiments, the individual may form mental representations of the remaining context together with each of the alternative target letters and choose the one judged most similar to the perceptual representation of the display. As the imperfectly perceived target letter is embedded in increasingly regular units, from UN to PN to W in our study, it becomes increasingly likely that its mental representation will be accompanied by a representation of the surrounding context and increasingly easy also for the individual to generate images of the context containing alternative target letters as a basis for comparison.

It may be noted that our informal account of the way the system employs context to facilitate identification is similar in many respects to the interactive process represented in the more formalized interactive-activation model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). The main difference in emphasis is that, whereas McClelland and Rumelhart conceive of a direct feedback from higher to lower order units during the formation of letter representations, we suggest that the role of context is indirect and distributed over the stages of processing. At an early stage, context constrains the formation of multiletter units. Then, in the later stages of processing under mask conditions, the selective passage of these units makes letters available for report that would otherwise have been lost because of interference from the mask and enables comparisons of mental representations of ambiguous inputs that facilitate discrimination between imperfectly perceived letters.

Some Questions on the Interpretation

In psychological terms, what is the locus of context effects? The answer to the question, posed by Reicher (1969) and many successors, of whether the effects occur at the perceptual level depends on one's definition of perception. If it is limited to the early stage of visual processing, then our answer is negative, which is in full agreement with the conclusion derived by Massaro (1979) from a meticulous quantitative analysis of the independent effects of orthographic context and visual letter information. It appears, rather, that the feature patterns of letters are processed in parallel through the construction of representations of individual letters in short-term visual memory (the visual buffer); the rates and limits of processing are strongly constrained by conditions of visibility, but are uninfluenced by linguistic context. If, however, perception is taken to include the cognitive operations that generate reports of letters or decisions based on their identification, then our answer is affirmative; we assume that these operations occur in working memory, access to which is influenced by word or pseudoword contexts.

A question that might be raised about the role we assume for encoding of letter groups is why attempts to test experimentally for effects of familiarity of letter groups have not been conspicuously successful. McClelland and Johnston (1977), for example, found no significant effect of variation in bigram frequency on the word or pseudoword advantage. However, bigram frequency was varied only within words and pseudowords, so their results indicate only that there is no appreciable effect of increased bigram frequency in strings that necessarily contain familiar letter groups. It may be noted further that in McClelland and Johnston's verbal report data, distributions of letters reported from nonwords exhibited statistical independence, whereas letters reported from words and pseudowords were positively intercorrelated, supporting the assumption that letter groups were encoded as units from words and pseudowords, but not from nonwords. Our answer to the question of why predictability of a target letter from adjacent context letters does not prove to be an important factor is that the role of context in furthering discriminability is not to allow unidentified letters in a display to be filled in by guessing (in agreement with Johnston, 1978), but to assimilate target letters whose features are registered in the first processing stage into units favored for continued processing.

Our proposed interpretation encompasses a variety of phenomena related to the word advantage. One group of these has to do with the implication that W or PN contexts influence decisions based on full-letter identification, but not decisions based on individual feature values. It is probably feasible for subjects to set decision criteria in terms of individual feature values only when the set of target alternatives is small and constant over a block of trials, the condition under which the word advantage has most often failed. But even in a two-alternative, forced-choice situation, the least conducive to context effects, we should expect a greater word advantage with postcues than with precues for target position, as observed by Estes (1977) and Holender (1979), since postcues are more likely to require processing of letter representations in working memory.

Another group of phenomena relates to task settings and instructions. The encoding of letter groups, an important component of our interpretation, is favored by any measures that lead subjects to try to identify words, but is impeded by measures that lead subjects to focus attention on individual letters or letter positions. It has frequently been observed that the word advantage is increased by the former and decreased by the latter (Johnston, 1981; Johnston & McClelland, 1973). An experimental variable that has been little studied in spite of its evident relevance is the nature of the mask stimulus when postmasks are used. In our interpretation, the type of mask most favorable to a word advantage should contain characters that compete with target letters for access to working memory, but that do not interfere with the recognition of familiar letter groups. Both familiar nonletter characters such as &s and artificial characters composed of letter fragments seem to meet these specifications, but the use of letters as components of masks has been found to reduce the word advantage (Johnston & McClelland, 1980). Our interpretation is that the letters in the postmask, which overlaps the target display in position, impede the recognition of familiar letter groups in the target display, thus producing the result observed by Johnston and McClelland that recognition of target letters in words is selectively impaired relative to recognition of single letters when a letter mask replaces an artificial character mask.

The word-superiority effect appears to be a byproduct of the orientation of the visual processing system, in reading and related tasks, toward the identification of larger units than single letters. As a consequence of readers' extensive preexperimental experience, letter groups that are frequent in words tend to be automatically activated as units during the early processing of a letter display, and a selective bias favors these larger units over isolated letters for passage from the early, parallel, to the later, serial, stage of processing. When overall performance is approximately equated for mask and no-mask conditions by adjustment of exposure duration, the combination of longer duration plus a postmask is more conducive to the encoding of multiletter units than is a shorter duration with no postmask. The relevance of unitization to the word-superiority effect is that target letters encoded into groups along with context letters have an increased likelihood of being represented in the later processing stage, in which they are accessible to report and in which other cognitive operations may enhance the accuracy of judgments about target identity.

REFERENCES

Adams, M. J. (1979). Models of word recognition. *Cognitive Psychology*, **11**, 133-176.

Baron, J. (1978). The word-superiority effect: Perceptual learning from reading. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Linguistic functions in cognitive theory* (Vol. 5, pp. 131-166). Hillsdale, NJ: Erlbaum.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.

Bjork, E. L., & Estes, W. K. (1973). Letter identification in relation to linguistic context and masking conditions. *Memory & Cognition*, **1**, 217-223.

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. London: Oxford University Press.

Brier, S. S. (1982). Analysis of categorical data. In G. Keren (Ed.), *Statistical and methodological issues in psychology and social sciences research*. Hillsdale, NJ: Erlbaum.

Coltheart, M. (1984). Sensory memory: A tutorial review. In H. Bouma and D. B. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 259-286). Hillsdale, NJ: Erlbaum.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, **87**, 272-300.

Estes, W. K. (1975). The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General*, **104**, 122-145.

Estes, W. K. (1977). On the interaction of perception and memory in reading. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp. 1-25). Hillsdale, NJ: Erlbaum.

Estes, W. K. (1978). Perceptual processing in letter recognition and reading. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Vol. IX. Perceptual processing* (pp. 163-220). New York: Academic Press.

Estes, W. K. (1982). Similarity-related channel interactions in visual processing. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 353-382.

Estes, W. K., Bjork, E. L., & Skaar, E. (1974). Detection of single letters and letters in words with changing vs. unchanging mask characters. *Bulletin of the Psychonomic Society*, **3**, 201-203.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Grizzle, J. E., Starmer, C. G., & Koch, G. G. (1969). Analysis of categorical data by linear models. *Biometrics*, **25**, 489-504.

Holender, D. (1979). Identification of letters in words and of single letters with pre- and postknowledge vs. postknowledge of the alternatives. *Perception & Psychophysics*, **25**, 313-318.

Johnston, J. C. (1978). A test of the sophisticated guessing theory of word perception. *Cognitive Psychology*, **10**, 123-153.

Johnston, J. C. (1981). Effects of advanced precuing of alternatives on the perception of letters alone and in words. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 560-572.

Johnston, J. C., & McClelland, J. L. (1973). Visual factors in word perception. *Perception & Psychophysics*, **14**, 365-370.

Johnston, J. C., & McClelland, J. L. (1980). Experimental tests of a hierarchical model of word identification. *Journal of Verbal Learning & Verbal Behavior*, **19**, 503-524.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 103-190). New York: Wiley.

Massaro, D. W. (1973). Perception of letters, words, and nonwords. *Journal of Experimental Psychology*, **100**, 349-353.

Massaro, D. W. (1979). Letter information and orthographic context in word perception. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 595-609.

McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception & Psychophysics*, **22**, 249-261.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **88**, 375-407.

Mewhort, D. J. K., Marchetti, F. M., Gurnsey, R., & Campbell, A. J. (1984). Information persistence: A dual-buffer model for initial visual processing. In H. Bouma & D. Bouwhuis (Eds.) *Attention and performance X: Control of language processes* (pp. 287-298). Hillsdale, NJ: Erlbaum.

Noreen, D. L. (1979). *Analysis of some mathematical models for choice and decision.* Unpublished doctoral dissertation, The Rockefeller University, New York.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 104-114.

Pachella, R. G., Smith, J. E. K., & Stanovich, K. E. (1978). Qualitative error analysis and speeded classification. In N. J. Castellan, Jr., & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 169-198). Hillsdale, NJ: Erlbaum.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, **81**, 275-280.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, **89**, 60-94.

Salasoo, A., Shiffrin, R. M., & Feustel, T. C. (1985). Building permanent memory codes: Codification and repetition effects in word identification. *Journal of Experimental Psychology: General*, **114**, 50-77.

Shepard, R. N. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, **65**, 242-256.

Smith, E. E., & Spoehr, K. T. (1974). The perception of printed English: A theoretical perspective. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 231-275). Hillsdale, NJ: Erlbaum.

Thompson, M. A., & Massaro, D. W. (1973). Visual information and redundancy in reading. *Journal of Experimental Psychology*, **98**, 49-54.

Townsend, J. T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, **9**, 40-50.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, **1**, 59-85.

NOTE

1. Although we prefer the framework of the ANOVAs applied to log frequencies, we have also computed χ^2 tests for subsets of cells of Table 2 that correspond to the various contrasts; the resulting pattern of significance closely parallels the one derived from the log linear ANOVA.

(Manuscript submitted April 9, 1986;
revision accepted for publication May 1, 1987.)