Multidimensional letter similarity derived from recognition errors

G. C. GILMORE, H. HERSH, A. CARAMAZZA, and J. GRIFFIN
The Johns Hopkins University, Baltimore, Maryland 21218

In order to provide a reliable measure of the similarity of uppercase English letters, a confusion matrix based on 1,200 presentations of each letter was established. To facilitate an analysis of the perceived structural characteristics, the confusion matrix was decomposed according to Luce's choice model into a symmetrical similarity matrix and a response bias vector. The underlying structure of the similarity matrix was assessed with both a hierarchical clustering and a multidimensional scaling procedure. This data is offered to investigators of visual information processing as a valuable tool for controlling not only the overall similarity of the letters in a study, but also their similarity on individual feature dimensions.

In order to understand how a particular stimulus is recognized, it is important to know the salient properties of the stimulus. To this end, experimental psychologists carefully measure and control such variables as the size and the luminance of a visual display. One property which has not been commonly subject to this careful scrutiny is the feature composition of the stimulus. This is unfortunate, for feature or distinctive figural attributes are considered fundamental to perception and ongoing information processing operations.

Feature lists have been proposed (Briggs & Hocevar, 1975; Geyer, 1970; Gibson, 1969; Laughery, 1971) which purportedly specify the important attributes of uppercase English letters. These lists have, however, enjoyed only limited use because of a reluctance by investigators to employ lists of questionable validity (Holbrook, 1975).

Townsend (1971a, 1971b) generated an alphabetic confusion matrix which has proven to be of some value to investigators in testing the power of visual recognition process models (Geyer & DeWald, 1973; Holbrook, 1975). This source of information on the similarity among letters has, however, rarely been used in empirical studies of the factors affecting letter recognition performance (Eriksen & Eriksen, 1974; Estes, 1972, 1974). This may be attributed to at least two factors: (1) The reliability of Townsend's confusion matrix is questionable due to the very low expected values in the off-diagonal elements of his matrix. Assuming a uniform distribution of confusions, the expected value of an off-diagonal element was only 3 responses out of 150. Thus, random

fluctuation could have significantly affected the values of the elements. (2) Since Townsend's study, many investigators have begun to use computer-based systems for the presentation of stimuli. The letters generated by these systems tend to be rather different from the typographically produced letters employed by Townsend. The reluctance of investigators to employ a similarity metric generated with this latter set of letters is understandable, since it is problematic whether this metric may be applied to letters of a different type font.

The purpose of the present study was to generate a highly reliable confusion matrix of uppercase letters displayed on a CRT, which could be used: (1) to establish a subjectively derived metric for describing the similarity of uppercase letters; (2) to analyze the errors of classification in an attempt to infer the important structural characteristics (features) of the letters; and (3) to eventually use this metric as a control variable in visual perception studies.

METHOD

Apparatus
The experimental procedure and the display device were under the control of a Digital Equipment Corporation PDP-11/20 computer. The CRT display, a DEC VR14 with a fast-decay phospher, was located in a lightproof room adjacent to the computer facility. The stimulus characters were generated in the center of the display by the selective illumination of dots within a 5 by 7 dot matrix. The matrix subtended a visual angle of 14' × 20' at a viewing distance of 64 cm. Under the viewing conditions employed, the letters appear as green continuous-line figures against a dark background. Representations of the generated letters are presented in Figure 1. The characters in this figure were constructed such that the dot size and interdot distances are proportional to the original displays.[1]

Subjects
Twenty students at The Johns Hopkins University participated and were paid $2/h for their services. Visual acuity in all subjects was normal or corrected to normal.

Design

The stimuli were presented in blocks of 72 letters. Each letter of the alphabet was presented three times within a block, with the order of presentation randomized by the computer just prior to the presentation of the block. Responses were collected from each subject on a total of 20 blocks of trials spread over 4 days of data collection. Thus, each uppercase letter was presented a total of 60 times to an individual (excluding practice trials).

One day of practice and calibration preceded the 4 experimental days. To insure that all subjects were operating at a suboptimal level, stimulus duration was adjusted after each block of trials so that the proportion of correct responses was held at .5. Stimulus duration ranged from 10 to 70 msec, with a mean duration across subjects of 33 msec.

Procedure

A subject was led into the room containing the CRT display and was asked to sit on a stool facing the screen. After attaching a head microphone, the subject placed his head on a chinrest to insure that the viewing distance was held at 64 cm. The subject was then left in the room for 5 min to dark-adapt.

The subject was instructed to fixate on the point which appeared in the center of the CRT screen at the beginning of each trial, and to identify the letter which appeared shortly therafter. Subjects were informed that their response accuracy was being measured, and they were asked to make their best guess on those trials on which they were not sure of the identity of the stimulus letter.

Each trial commenced with the presentation of the fixation point for 1 sec. To minimize potential masking effects, the stimulus letter was presented 1 sec after the fixation point disappeared, and was followed by a blank field. The subject then verbally identified the stimulus. The response was picked up by the microphone and transmitted to the experimenter, who proceeded to enter the response into the computer through a terminal keyboard.[2] Immediately after the experimenter had recorded the response, the fixation point for the next trial appeared. Each trial period lasted approximately 5 sec. After each block of trials, the subject was given a 3-min rest period while the computer calculated his performance. A practice block of 26 trials was presented at the beginning of each session. An experimental session, including the dark adaptation, instructions, a practice block, and five experimental blocks, lasted approximately 50 min.

RESULTS

The responses for each subject were summed across blocks into a Stimulus by Response confusion matrix, where each element indicated the number of times a stimulus letter, s_i, was identified as a response letter, r_j. Individual differences in response frequencies among subjects were analyzed by an inverted principle-components factor analysis before obtaining a composite confusion matrix. The first component, indicative of communality among subjects, accounted for 91% of the total variance, leaving only 9% for stable individual differences and subject unreliability. In addition, the subjects did not appear to fall into any obvious subgroupings based on response patterns. It was therefore deemed appropriate to average over the 20 subjects to obtain a composite matrix.

Since each stimulus letter was presented a total of 1,200 times, the composite confusion matrix was normalized by this value and each element was then interpreted as a proportion. The resulting matrix is shown in Table 1. Overall, the proportion of correct responses was .51, reasonably close to the projected correct response rate of .5.

To facilitate an analysis of the perceived structural characteristics of the uppercase font used in this paper, the confusion matrix was decomposed, according to Luce's choice model (Luce, 1963), into a symmetrical matrix and a response-bias vector. Since the choice model assumes that the probability of choosing any letter, r_j, as a response to any stimulus letter, s_i, is greater than .0, it was necessary to estimate the values of any empty off-diagonal elements. (Less than 5% of the off-diagonal elements were empty.) An iterative, least-squares procedure was used to solve for the similarity matrix and bias vector by substituting a variable probability value in the empty elements of the confusion matrix. With the value of these elements set at .001, the sum of the squared errors in reconstructing the original confusion matrix proportions from the similarity matrix and bias vector was only .076.[3] It thus appears that the choice model solution, as shown in Table 2, provides a very good fit to the confusion data. It is interesting to note that the three letters with the largest bias values, I, J, and L, contain few physical elements and can, in fact, be contained in other letters.

The values of the elements in the matrix of Table 2 represent proximities along a scale relating the set of 26 letters on the basis of perceived (structural) similarity. Although the matrix represents only the overall similarity among the letters, several analytic techniques are available for assessing underlying structures which may form the basis for the similarities. One such nonspatial technique is hierarchical clustering (Johnson, 1967), where the level in the hierarchy at which two elements cluster is monotonically related to the similarity of the elements: i.e., the higher in the hierarchy at which two letters cluster, the weaker the similarity relationship.

The results of clustering the 26 uppercase letters are shown in Figure 1, using both the minimum and maximum procedures. According to Johnson (1967), if the data conform to the ultrametric inequality, then the two solutions will be the same, and thus a hierarchical relationship exists among the data. The closeness of the two solutions implies that a nonspatial representation is appropriate for these data. The two solutions show that the letters H and M are the most similar letters for the font used, while L appears to be a rather unique letter. Two major clusters are evident: one appears to be related to the angularity of the physical representation (e.g., X, N, W), while in the other cluster, curved letters (e.g., D, G, O, Q) are predominant.

Although the results of the clustering procedure clearly depict the levels of similarity among the let-

Table 1
Confusion Matrix

Response

Stimulus	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	.278	.018	.007	.008	.004	.010	.039	.129	.004	.005	.027	.005	.044	.051	.031	.023	.016	.224	.017	.005	.011	.007	.014	.013	.003	.009
B	.022	.155	.022	.078	.062	.024	.119	.029	.012	.015	.023	.022	.011	.022	.061	.043	.024	.085	.048	.012	.034	.028	.013	.012	.010	.014
C	.005	.007	.545	.017	.058	.020	.179	.001	.008	.003	.008	.022	.001	.003	.043	.010	.013	.009	.005	.005	.004	.004	.002	.006	.003	.021
D	.011	.039	.045	.379	.032	.018	.084	.007	.013	.006	.010	.018	.002	.007	.105	.060	.020	.033	.011	.013	.028	.041	.003	.007	.003	.008
E	.004	.038	.039	.012	.457	.057	.073	.012	.028	.004	.042	.061	.004	.004	.005	.016	.008	.034	.027	.008	.014	.008	.004	.016	.007	.019
F	.003	.002	.009	.002	.025	.721	.010	.004	.015	.001	.013	.003	.008	.003	.003	.106	.001	.023	.008	.023	.002	.005	.002	.006	.003	.003
G	.010	.015	.063	.021	.030	.015	.518	.006	.003	.013	.013	.023	.003	.010	.069	.013	.041	.030	.031	.007	.012	.013	.009	.009	.007	.016
H	.033	.008	.000	.003	.001	.007	.014	.484	.007	.009	.040	.007	.063	.117	.008	.010	.005	.038	.003	.004	.022	.004	.087	.023	.003	.003
I	.000	.000	.001	.001	.002	.004	.000	.002	.865	.017	.000	.013	.000	.001	.000	.005	.000	.001	.003	.073	.004	.001	.000	.001	.001	.007
J	.000	.001	.001	.001	.000	.000	.005	.002	.007	.934	.004	.000	.002	.000	.005	.003	.001	.001	.002	.002	.019	.006	.002	.000	.000	.002
K	.013	.007	.008	.003	.039	.042	.014	.031	.023	.006	.562	.032	.015	.032	.005	.009	.002	.062	.004	.012	.008	.011	.017	.030	.009	.003
L	.001	.002	.003	.001	.004	.000	.003	.002	.013	.003	.008	.943	.000	.001	.003	.000	.000	.003	.000	.003	.006	.000	.001	.001	.002	.001
M	.035	.008	.000	.011	.003	.008	.011	.391	.009	.009	.027	.004	.110	.165	.018	.010	.008	.056	.006	.006	.011	.003	.053	.024	.010	.004
N	.032	.004	.003	.007	.003	.004	.016	.292	.008	.003	.028	.088	.060	.288	.010	.007	.008	.048	.004	.003	.024	.004	.100	.032	.004	.003
O	.016	.033	.036	.108	.007	.002	.156	.007	.006	.013	.004	.005	.002	.002	.393	.013	.077	.018	.015	.002	.040	.023	.009	.005	.002	.006
P	.010	.017	.006	.012	.012	.212	.014	.008	.006	.002	.013	.002	.004	.004	.013	.564	.001	.043	.004	.028	.002	.010	.002	.003	.008	.003
Q	.023	.027	.026	.059	.015	.006	.183	.012	.004	.011	.014	.003	.000	.012	.207	.020	.184	.047	.023	.005	.042	.026	.017	.008	.006	.011
R	.041	.017	.008	.013	.035	.056	.040	.029	.012	.007	.073	.013	.014	.021	.019	.109	.018	.364	.013	.011	.006	.012	.023	.021	.007	.018
S	.011	.068	.019	.029	.026	.019	.130	.008	.008	.013	.015	.006	.005	.011	.036	.019	.028	.050	.392	.012	.015	.025	.006	.018	.013	.018
T	.001	.001	.002	.002	.003	.018	.004	.003	.053	.002	.002	.002	.001	.002	.002	.026	.001	.001	.002	.844	.002	.002	.001	.000	.025	.002
U	.003	.007	.002	.019	.006	.002	.037	.027	.013	.028	.002	.013	.007	.011	.042	.004	.021	.004	.001	.002	.655	.050	.026	.010	.005	.002
V	.000	.030	.003	.007	.003	.003	.004	.021	.003	.005	.006	.002	.004	.003	.010	.015	.007	.002	.000	.006	.033	.631	.031	.009	.171	.010
W	.027	.006	.005	.005	.003	.004	.012	.354	.008	.013	.034	.010	.058	.131	.009	.009	.012	.043	.003	.001	.043	.008	.180	.014	.006	.003
X	.016	.006	.003	.005	.003	.003	.007	.042	.004	.013	.191	.004	.055	.054	.004	.008	.003	.052	.006	.005	.008	.015	.030	.437	.014	.013
Y	.001	.004	.001	.002	.000	.008	.001	.000	.005	.003	.003	.002	.000	.001	.001	.013	.002	.004	.003	.028	.002	.055	.002	.007	.853	.002
Z	.006	.006	.036	.014	.047	.019	.018	.009	.032	.033	.011	.013	.003	.002	.008	.008	.008	.016	.021	.012	.003	.012	.006	.031	.010	.618

Table 2
Interletter Similarities and Response Bias of Each Letter Generated by the Choice Model from the Original Confusion Matrix

Response Bias

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
.668	.415	1.215*	.946	.947	2.038	1.601	1.524	3.478	3.259	1.494	3.370	.420	.889	1.229	1.844	.471	1.181	.681	2.656	1.680	1.831	.624	1.044	2.427	1.045

Similarity

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	1.000																									
B	.096	1.000																								
C	.015	.043	1.000																							
D	.029	.228	.061	1.000																						
E	.011	.182	.095	.047	1.000																					
F	.012	.021	.021	.066	.011	1.000																				
G	.052	.149	.200	.096	.095	.020	1.000																			
H	.178	.056	.002	.007	.011	.009	.018	1.000																		
I	.004	.009	.004	.011	.006	.003	.006	.006	1.000																	
J	.004	.010	.002	.007	.012	.012	.012	.006	.012	1.000																
K	.047	.043	.014	.080	.003	.001	.037	.068	.007	.007	1.000															
L	.004	.017	.011	.012	.024	.002	.002	.006	.014	.002	.022	1.000														
M	.224	.072	.004	.023	.015	.028	.024	.006	.010	.013	.081	.006	1.000													
N	.143	.044	.008	.021	.010	.008	.033	.495	.006	.003	.074	.005	.559	1.000												
O	.067	.182	.085	.276	.014	.008	.230	.017	.004	.013	.010	.006	.013	.029	1.000											
P	.038	.091	.014	.058	.027	.005	.005	.026	.008	.003	.019	.002	.081	.025	.013	1.000										
Q	.085	.151	.058	.130	.038	.007	.014	.079	.005	.008	.016	.004	.013	.056	.469	.014	1.000									
R	.301	.160	.019	.056	.085	.070	.149	.080	.006	.005	.149	.011	.080	.140	.049	.151	.112	1.000								
S	.041	.232	.021	.046	.063	.023	.021	.141	.008	.008	.017	.004	.141	.026	.020	.059	.094	.067	1.000							
T	.005	.010	.005	.009	.017	.026	.005	.008	.073	.002	.007	.003	.005	.008	.005	.003	.039	.006	.009	1.000						
U	.013	.048	.005	.046	.017	.003	.008	.036	.010	.029	.007	.011	.043	.033	.081	.005	.086	.010	.008	.003	1.000					
V	.006	.093	.006	.035	.012	.006	.013	.034	.002	.007	.014	.002	.004	.013	.030	.021	.040	.010	.010	.005	.063	1.000				
W	.087	.053	.010	.034	.013	.008	.034	.595	.007	.012	.076	.008	.394	.003	.034	.013	.078	.123	.016	.003	.097	.047	1.000			
X	.041	.033	.009	.011	.015	.008	.017	.068	.003	.006	.153	.003	.166	.003	.117	.010	.017	.083	.025	.004	.017	.022	.073	1.000		
Y	.004	.017	.003	.002	.002	.006	.004	.003	.003	.002	.008	.010	.010	.002	.003	.015	.009	.011	.009	.031	.004	.132	.009	.016	1.000	
Z	.017	.030	.047	.014	.008	.011	.030	.010	.020	.014	.008	.005	.013	.005	.014	.008	.028	.036	.040	.007	.004	.018	.013	.039	.006	1.000

Table 3
Interletter Euclidean Distances from Multidimensional Scaling Solution

Euclidean Similarity

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	0.000	0.579	1.078	0.772	0.931	0.979	0.708	0.556	1.386	1.402	0.728	1.257	0.486	0.421	0.750	0.791	0.615	0.357	0.809	1.347	0.930	1.085	0.532	0.772	1.327	1.174
B		0.000	0.701	0.439	0.653	0.830	0.429	0.833	1.119	1.109	0.788	1.152	0.652	0.763	0.485	0.633	0.384	0.457	0.454	1.031	0.793	0.672	0.681	0.771	0.949	0.756
C			0.000	0.639	0.494	1.004	0.470	1.350	1.217	1.328	1.124	0.982	1.233	1.289	0.715	1.014	0.761	0.903	0.754	1.264	1.153	1.204	1.228	1.311	1.437	0.827
D				0.000	0.837	0.992	0.514	1.077	1.192	1.259	1.133	1.210	0.944	1.013	0.338	0.708	0.439	0.753	0.770	1.121	0.780	0.836	0.931	1.189	1.210	1.066
E					0.000	0.697	0.581	1.124	1.068	1.327	0.715	0.844	1.008	1.073	0.922	0.874	0.873	0.635	0.680	1.097	1.220	1.221	1.069	1.007	1.286	0.646
F						0.000	1.048	1.185	1.136	1.707	0.847	1.273	1.046	1.143	1.224	0.468	1.173	0.680	0.967	0.833	1.446	1.262	1.211	1.122	1.136	1.106
G							0.000	0.992	1.254	1.130	0.901	1.002	0.882	0.911	0.411	0.953	0.357	0.633	0.534	1.321	0.866	1.000	0.839	0.981	1.329	0.773
H								0.000	1.184	1.166	0.624	1.148	0.224	0.149	1.005	1.038	0.887	0.637	1.117	1.286	0.813	1.138	0.257	0.631	1.294	1.212
I									0.000	1.073	1.081	0.846	1.123	1.259	1.288	1.132	1.315	1.183	1.439	0.619	1.066	1.254	1.171	1.298	1.192	1.131
J										0.000	1.264	1.179	1.108	1.220	1.082	1.578	1.067	1.343	1.296	1.398	0.749	1.025	0.986	1.129	1.279	0.973
K											0.000	0.954	0.566	0.612	1.135	0.960	1.003	0.484	0.906	1.151	1.156	1.242	0.692	0.524	1.220	0.861
L												0.000	1.168	1.177	1.193	1.369	1.206	1.092	1.367	1.295	1.144	1.602	1.150	1.324	1.713	1.101
M													0.000	0.222	0.898	0.879	0.767	0.508	0.938	1.150	0.774	0.942	0.221	0.498	1.089	1.061
N														0.000	0.938	0.992	0.805	0.545	1.025	1.327	0.837	1.123	0.257	0.612	1.306	1.190
O															0.000	0.969	0.205	0.807	0.755	1.330	0.620	0.841	0.812	1.118	1.306	1.028
P																0.000	0.939	0.625	0.918	0.784	1.169	0.969	1.022	1.089	0.985	1.186
Q																	0.000	0.680	0.618	1.339	0.654	0.797	0.685	0.943	1.237	0.941
R																		0.000	0.656	1.116	1.010	1.037	0.627	0.661	1.156	0.926
S																			0.000	1.328	1.162	0.869	0.968	0.817	1.075	0.646
T																				0.000	1.258	1.069	1.271	1.291	0.841	1.194
U																					0.000	0.859	0.628	1.093	1.277	1.193
V																						0.000	0.943	0.985	0.598	0.998
W																							0.000	0.582	1.188	1.078
X																								0.000	1.003	0.845
Y																									0.000	1.038
Z																										0.000

MAXIMUM
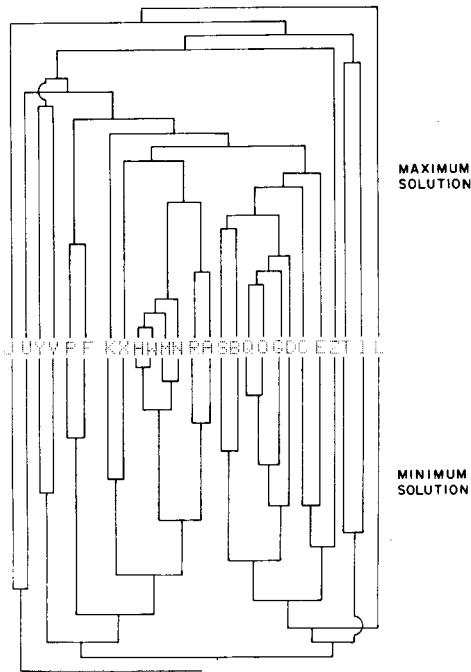SOLUTION

MINIMUM
SOLUTION

Figure 1. Hierarchical clustering representations of the letter similarities.

ters, the resulting hierarchies are nondimensional: i.e., they do not reveal whether a common dimensional structure underlies the perceptual space. To obtain this information, the similarity matrix was decomposed into a dimensional representation through the use of the multidimensional scaling procedure TORSCA (Young & Torgerson, 1967). The value of this procedure is that the similarity among the letters can be specified not only on an overall basis, but also in terms of individual feature dimensions. Multidimensional scaling thus gives a fine-grain representation of similarity.

Solutions of various dimensionalities were obtained in an attempt to find a spatial representation which adequately characterized the 26 letters. An acceptable fit (stress = .077) was obtained with a five-dimensional, Euclidean metric solution. The distances separating the letters in this multidimensional space are given in Table 3, and the coordinate value of each letter on the five dimensions is contained in Table 4. The relationship among the letters, depicted in the clustering solution of Figure 1, remain well represented in the five-dimensional solution. For instance, H and M have very similar dimensional values, while the letter L has rather extreme values on three of the dimensions.

Given the postulation by several investigators (Geyer, 1970; Gibson, 1969; Laughery, 1971) of feature lists, it would be of interest to name the dimensions in Table 4. The physical continua which correspond to

the five dimensions are not readily apparent, however. While the first dimension separates the two letter groups noted in the cluster analysis and the second dimension seems to distinguish among the letters according to their widths, the remaining dimensions elude even such a loose specification.[4] Thus, we are faced with a situation where a perceptual structure can be specified but the physical dimensions related to it cannot be labeled. This is a rather interesting reversal of a common psychophysical problem.

DISCUSSION

The empirical confusion matrix presented here, and the letter similarity values derived from it, should be of value in investigations of the letter recognition process. For instance, several models of letter identification (Gardner, 1973; Rumelhart, Note 1) have postulated effects based on the similarity among letters in a visual display and in memory. The information presented here should permit a careful testing of these hypothesized effects.

Before these measures are used, one must, of course, have confidence in the reliability and the validity of the measures. The fact that the proportion in each cell of the confusion matrix is based on 1,200 presentations indicates that stimulus sampling was great enough to ensure a reliable response measure. Also, the fact that the choice model provides such a good

Table 4
Dimension Values from Multidimensional Scaling Solution

	Dimension				
	1	2	3	4	5
A	.251	.492	.157	−.102	.024
B	−.167	.136	.042	−.088	−.120
C	−.648	.120	.138	.350	.123
D	−.445	.265	−.008	−.218	.161
E	−.254	−.048	.354	.428	.031
F	−.082	−.181	.782	−.075	.067
G	−.308	.325	−.046	.241	−.010
H	.612	.181	−.105	−.063	.128
I	.001	−.746	−.080	.033	.526
J	.013	−.357	−.887	.203	−.039
K	.441	−.064	.261	.309	−.038
L	.014	−.206	−.065	.671	.648
M	.466	.113	−.054	−.124	−.005
N	.549	.289	−.047	−.052	.073
O	−.357	.387	−.267	−.084	.080
P	−.117	−.018	.511	−.417	.097
Q	−.218	.402	−.218	−.044	−.057
R	.162	.215	.307	.033	−.021
S	−.267	.215	.148	.131	−.481
T	−.113	−.750	.283	−.371	.252
U	.013	.111	−.610	−.207	.274
V	−.214	−.152	−.307	−.503	−.388
W	.430	.197	−.248	−.076	.020
X	.496	−.023	−.010	.096	−.427
Y	−.039	−.601	.023	−.518	−.516
Z	−.218	−.302	−.052	.449	−.401

fit to the data suggests that the matrix contains, at least in relation to the choice model, a rather small noise component. The face validity of the measures can be evaluated by simply examining the letters in Figure 1. It makes sense, given their physical characteristics, that H, M, N, and W would have a high-scaled similarity, and it is reasonable that the letters I, L, and Z should be scaled as highly unique representations.

Another source, supporting the validity of these measures, is a series of experiments reported by Gilmore (1976). He employed the multidimensional similarity metric reported here, in evaluating target identification performance when the target was flanked by noise letters. He noted that performance in each experiment could be accounted for by postulating two distinct similarity effects. One was attributable to the overall similarity between the letters, while another seemed to be mediated by the similiarity of the letters on a single "critical feature" dimension. Obviously, this interpretation of the results would not have arisen if the scaled properties of the stimuli were not known. Thus, regardless of whether or not Gilmore's postulations are correct, he has demonstrated that the similarity metric can be a valuable research tool.

REFERENCE NOTE

1. Rumelhart, D. E. *A multicomponent theory of confusion among briefly exposed alphabetic characters.* (Center for Human Information Processing Report No. 22) San Diego: University of California, 1971.

REFERENCES

BRIGGS, R., & HOCEVAR, D. J. A new distinctive feature theory for upper case letters. *The Journal of General Psychology,* 1975, **93**, 87-93.

ERIKSEN, B. A., & ERIKSEN, C. W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics,* 1974, **16**, 143-149.

ESTES, W. K. Interactions of signal and background variables in visual processing. *Perception & Psychophysics,* 1972, **12**, 278-286.

ESTES, W. K. Redundancy of noise elements and signals in visual detection of letters. *Perception & Psychophysics,* 1974, **16**, 53-60.

GARDNER, G. T. Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology,* 1973, **4**, 130-155.

GEYER, L. H. A two channel theory of short-term visual storage (Doctoral dissertation, SUNY at Buffalo, 1970). *Dissertation Abstracts International,* 1970, **31**, 5660B. (University Microfilms No. 71-7165)

GEYER, L. H., & DEWALD, C. G. Feature lists and confusion matrices. *Perception & Psychophysics,* 1973, **14**, 471-482.

GIBSON, E. *Principles of perceptual learning development.* New York: Appleton-Century-Crofts, 1969.

GILMORE, G. C. The interactions among letters in a visual display (Doctoral dissertation, The Johns Hopkins University, 1975). *Dissertation Abstracts International,* 1976, **36**, 3645B. (University Microfilms No. DAH76-01540)

HOLBROOK, M. B. A comparison of methods for measuring the interletter similarity between capital letters. *Perception & Psychophysics,* 1975, **17**, 532-536.

JOHNSON, S. L. Hierarchical clustering schemes. *Psychometrika,* 1967, **32**, 241-254.

LAUGHERY, K. R. Computer simulation of short-term memory: A component decay model. In G. T. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. VI). New York: Academic Press, 1971.

LUCE, R. D. Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. I). New York: Wiley, 1963.

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

TOWNSEND, J. T. Alphabetic confusion: A test for individuals. *Perception & Psychophysics,* 1971, **9**, 449-454. (b)

YOUNG, F. W., & TORGERSON, W. S. TORSCA, A FORTRAN IV program for Shepard-Kruskal multidimensional scaling analysis. *Behavioral Science,* 1967, **12**, 498.

NOTES

1. The letters depicted in Figure 1, which were constructed to resemble the original CRT letters, are now available as transfer lettering. This will allow future investigators to present these letters with a conventional tachistoscope. The material, Letraset Serial No. S51406, can be ordered through any authorized Letraset dealer. The investigator is advised to specify whether or not a clear carrier film printed over the individual letters is desired. A minimum order of 50 sheets is required.

2. Care was taken to avoid phonetic confusions by having the subject repeat potentially confusable letters as the first letter of a word. For instance, the subject may have responded, "M, as in man."

3. In Townsend (1971a), the sum of squared errors in the comparable (CI) condition was .23, a factor of three greater than the error found in this study.

4. It is interesting to note that Townsend (1971a, 1971b) also failed to discern any physical dimensions emerging from his multidimensional scaling solution.