Behavior Research Methods, Instruments, & Computers
1989, 21 (6), 579-586

An upper- and lowercase alphabetic similarity matrix, with derived generation similarity values

DAVID B. BOLES and JOHN E. CLIFFORD
Rensselaer Polytechnic Institute, Troy, New York

A full upper- and lowercase visual similarity matrix is presented for a standard set of computer characters, implemented on the Apple-Psych system. The 2,704 (52×52) letter pairs were rated by 12 subjects each. From the ratings, generation and veridical similarity values are derived, and they are tabled for use in research on mixed-case letter matching. In addition, the results of multidimensional scaling and cluster analyses are presented, which give complementary, simplified descriptions of the data.

The impact of visual similarity on pure-case letter comparisons (e.g., between "A" and "A") is well documented. Letters that are commonly confused with one another (e.g., "GQ") produce longer reaction times (RTs) in same-different matching tasks than do letters that are rarely confused (Bagnara, Boles, Simion, & Umiltà, 1983). This effect is a source of evidence that pure-case letter pairs are matched physically (visually), and, indeed, the task of matching such pairs is known as the *physical* matching task (Posner, 1978).

Mixed-case letter comparisons (e.g., between "A" and "a") have come to be known as *name* matches, because introspection and early experimental results suggested that such matches are made by extracting and comparing phonetic representations of the letter names (Dainoff & Haber, 1970; Posner, Boies, Eichelman, & Taylor, 1969; Thorson, Hochhaus, & Stanners, 1976). However, more recent work indicates that phonetic or name codes do not typically underlie such matches (Besner, Coltheart, & Davelaar, 1984; Boles, 1986; Boles & Eveland, 1983). Thus, Boles and Eveland (1983) showed that phonetically similar different pairs (e.g., "Gd") produce no increment in RT, contrary to what is predicted if name codes are used. In addition, mixed-case matches fail to produce a right-visual-field superiority with lateralized presentation, again contrary to prediction if left hemisphere phonetic processes are involved. In contrast, letter pairs judged same-different on the basis of rhyme do show such a superiority (Boles, 1986). The implication is that rhyme matches draw on phonetic processes, which are lateralized to the left hemisphere, while name matches do not.

In further investigations, Boles and Eveland (1983) and Boles (1986) formulated and tested a visual generation model of mixed-case matching. According to the model, one or both members of a presented mixed-case pair rapidly generate a visual representation of the opposite case in memory. For example, if both members generate the opposite case, the pair "Aa" might result in the generation of "aA" in memory. The generated letter or letters are then compared to one or both of the original letters, supporting a type of visual match.

Evidence for the generation model was produced by manipulations of generation similarity. That is, different pairs were constructed in which the generated letters were visually similar to the original letters (e.g., "Gq" might generate "gQ"). RTs to such generation similar pairs proved to be substantially longer than to dissimilar pairs. The effect could not be attributed to visual similarity between the original letters themselves (called *veridical* similarity).

Unfortunately, a serious impediment to further pursuit of this work has been the absence from the literature of a visual similarity matrix using all crossings of both upper- and lowercases of a standard font. This is an impediment because (1) generation similarity as used by Boles and Eveland (1983) and Boles (1986) is a derived index based on the similarities of the upper- and lowercase forms of a letter pair, and (2) in constructing generation similar and dissimilar letter sets for comparison, veridical similarity (e.g., between "G" and "q" for the pair "Gq") should be matched between the sets if this variable is to be controlled. Thus, a similarity matrix is needed that tables similarity values for both pure-case and mixed-case pairs. Existing matrices are generally based on confusions made in recognizing single letters, and so, in effect, they represent pure-case similarities (e.g., Gilmore, Hersh, Caramazza, & Griffin, 1979; Townsend, 1971; van der Heijden, Malhas, & van den Roovart, 1984).

In the absence of a matrix using both pure-case and mixed-case pairs, the experimenter must collect ratings for the particular font in use and compute the generation and veridical similarity indexes. The process is very time-consuming, and, for practical reasons, it may require limitation to a small set of candidate letter pairs. Each candidate pair requires the rating of three or four actual

We would like to thank Michael E. Walker and Angela Ebreo for statistical consulting and assistance, and Susan Savage for entering data for the analyses. We would also like to thank Lester Krueger for comments on an earlier draft. Reprint requests should be sent to David B. Boles, Department of Psychology, Rensselaer Polytechnic Institute, Troy, NY 12180.

pairs (for the candidate pair "Gq," the pairs "Gq," "GQ," "gq," and, if completeness is desired, "gQ"). Thus, Boles and Eveland (1983) restricted ratings to only 50 candidate pairs.

To overcome this methodological limitation, we now report a full similarity matrix using the upper- and lowercase versions of a standard set of characters. The characters are those used in the Apple-Psych system of experimental software, which is a system freely available and widely distributed in the public domain for Apple II-series computers. It also has received recent description in a number of articles (Barnes & Burke, 1988; Boles, 1988; Osgood, 1988). Thus, the matrix can be used by anyone implementing the Apple-Psych system and desiring to conduct letter-confusion research. To add to the utility of the results, we also report derived generation similarity values, so that these need not be computed.[1]

We must emphasize that our main purpose in constructing and reporting the matrix is methodological. The comparison of such models of visual confusion as template overlap, featural, spatial frequency, or choice models (Gervais, Harvey, & Roberts, 1984; Loomis, 1982), while important, is beyond the scope of this paper. Nevertheless, we will report the outcomes of multidimensional scaling (MDS) and cluster analyses of the similarity ratings. This is to provide a description of the data that is simplified from the 2,704 (52×52) original letter pairs, and to point out structural aspects of the data that those taking a theoretical approach to letter similarity may want to consider.

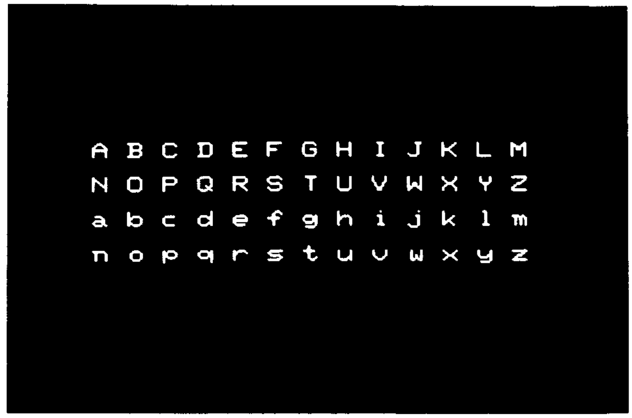

Figure 1. The Apple-Psych high-resolution letter set.

METHOD

Subjects

Thirty-two undergraduate volunteers participated, receiving extra course credit. All had normal or corrected vision.

Apparatus and Stimuli

Letter pairs were generated and responses collected, using Apple-Psych software on an Apple IIe microcomputer. Letters were displayed in Apple-Psych size "1" on an Amdek 300 monitor with P31 phosphor, in high-resolution mode. All 2,704 ordered pairs of upper- and lowercase letters (a 52×52 matrix) were used. The letter set is illustrated in Figure 1. Letters were constructed of illuminated pixels, generally in a 6×7 (horizontal × vertical) matrix. Letters with descenders (e.g., "y") involved an 8th row. The letter set is standard in the Apple-Psych system and is identical to the SYSTEM.CHARSET provided with Apple Pascal. Size varied somewhat with the letter, but to take a representative pair like "Zm," the capital letter subtended 0.4° both horizontally and vertically, and the lowercase letter 0.4°×0.3°. Letters within a pair were immediately adjacent, with the gap between them amounting to about 1.4′. The subjects used a chinrest to maintain a fixed viewing distance of 0.48 m.

Procedure

The full matrix of pairs was randomly divided into eight subsets of 338 pairs each. Each subject was given three subsets to rate, with the subsets randomly selected and presented in random order, under the constraint that each subset was rated by 12 subjects. The pairs within a subset were also presented in random order.

On a given rating trial, the subject saw the heading "LOW=1 HIGH=5" at the top of the screen, and a letter pair centered toward the bottom of the screen. The subject then pressed one of five keys, labeled 1-5, on the computer console, to indicate the rated similarity. They were instructed to rate the pairs on overall visual similarity, with "1" representing low similarity and "5" high similarity. The letters remained on the screen until the subject responded.

RESULTS AND DISCUSSION

Mean similarity values were calculated for each pair over subjects, with each pair having been rated by 12 subjects. These values are presented in Appendix A. Analysis indicates a high correlation between the two possible orders of the pairs, $r = +.88$, indicating that the matrix is substantially, though not perfectly, symmetric.

Summary data are shown in Table 1, indicating the number of pairs falling into each of eight similarity bands, along with the mean (across pairs) of the standard devia-

Table 1
Summary Data of the Pairs Classified by Pair Type and Similarity Band

	Pair Type				
	Mixed	Upper	Lower	All Pairs	
Band	*N*	*N*	*N*	*N*	*SD*
1 00-1.50	608	7	48	663	.61
1.51-2 00	530	198	256	984	87
2.01-2.50	128	279	188	595	99
2.51-3 00	44	109	83	236	1.03
3.01-3.50	16	46	51	113	99
3 51-4.00	11	10	20	41	.88
4 01-4.50	14	1	4	19	.75
4.51-5.00	1	26	26	53	19

Note—N = number of pairs in band SD = mean standard deviation tions (across subjects) for the pairs within each band. Most pairs were rated fairly low in similarity, the mode being in the 1.51-2.00 band. Nearly all of the pairs in the highest band, of course, were physically identical pairs (e.g., "AA"). It is to be noted that not all of the identical pairs produced perfect similarity values of 5.00, undoubtedly due to occasional keying errors by the subjects. Not surprisingly, the mean standard deviation is closely related to the scalar position of the band: Standard deviations were lower for the extremes (.61 and .19) than for the middle bands (1.03 and .99).

From these veridical similarity values, generation similarity can be computed. First, however, there is a complexity in the process that must be considered. Notably, the generation model makes rather unspecific claims as to what is generated and compared. The presented pair "Ab," for example, might generate only "a" (which is then compared to "b"), or only "B" (compared to "A"), or both "a" and "B" (respectively compared to "b" and "A"). Obviously, which model is adopted has a bearing on how generation similarity should be calculated. However, Boles and Eveland (1983) avoided the issue altogether by using as stimuli both forms of an original source pair (e.g., "Ab" and "aB") and calculating generation similarity in a joint manner (e.g., as the mean similarity of "A" to "B" and "a" to "b"). Regardless of the form of the model that is adopted, the joint generation similarity value describes the average generation similarity of the two forms of the source pair.

In short, we have followed this practice in computing the present generation similarity values. We have also collapsed over the two letter orders (e.g., "Ab" and "bA") since, as noted above, the matrix is substantially symmetric. To take an example, the generation similarity of the pair "Ab" was calculated as the mean similarity of "AB" (2.25), "ab" (2.66), "BA" (2.42), and "ba" (2.25), or a value of 2.40. The result of the computations is Appendix B, giving the generation similarity for each pair (upper entry) and, for convenience in constructing stimulus sets, the veridical similarity as well (lower entry). The tabled veridical similarity values were calculated under the same assumptions—that is, that both forms of a pair would be presented and that order is largely irrelevant. For example, the veridical similarity of "Ab" (1.52) reflects the mean similarity of "Ab" (1.42), "aB" (1.75), "bA" (1.50), and "Ba" (1.42).

It is instructive to compare the present generation and veridical similarity values to those reported by Boles and Eveland (1983) for 50 letter pairs. In spite of the fact that somewhat different computer fonts were used in the two instances, the correlation between generation similarity values is $r = +.84$, and that between veridical similarity values is $r = +.74$. Evidently there is substantial similarity between the two fonts.

To better understand the subjects' ratings of similarity between letters, the similarity matrix in Appendix A (symmetrized by arithmetic averaging) was analyzed, using the nonmetric multidimensional scaling (MDS) program KYST-2A (Kruskal, Young, & Seery, 1977). This program uses Torgerson's iterative procedure to fit Kruskal's (1964a, 1964b) model. An examination of the stress plot for the 2- through 5-dimensional solutions indicated the adequacy of fit of three dimensions. The decision to accept three dimensions was based on the form of the skree plot, with a pronounced elbow existing at three dimensions. In terms of a reduced number of dimensions, ours is a more conservative decision than that made by Gilmore et al. (1979), who accepted five dimensions on the basis of a low stress value (.077). Our stress value at three

Table 2
3-Dimensional Coordinates from the MDS Solution

	Dimension		
Letter	1	2	3
A	− 89	−.10	− 28
B	−.64	−.77	15
C	− 33	−.62	−.33
D	− 51	−.82	00
E	− 78	− 32	−.17
F	−.84	− 48	.41
G	−.35	− 90	−.32
H	−.88	26	.07
I	− 81	.08	.74
J	−.42	−.16	71
K	−.85	.54	.26
L	−.89	.05	.39
M	−.72	45	−.61
N	−.74	44	− 37
O	− 12	− 78	− 34
P	− 25	− 64	.29
Q	−.32	−.86	−.53
R	−.76	−.47	.08
S	− 32	−.53	−.90
T	− 72	− 10	72
U	− 39	09	−.32
V	−.37	.77	−.38
W	−.54	.68	−.56
X	−.71	.85	−.25
Y	− 46	.74	.17
Z	− 75	−.12	−.85
a	1.12	−.29	−.38
b	.54	−.66	.25
c	.58	− 22	−.34
d	.66	−.49	.48
e	.78	− 47	−.50
f	.29	−.27	1.06
g	1.08	−.36	.28
h	.42	.30	.34
i	.46	.56	1.05
j	.66	16	.80
k	.03	.97	.48
l	−.21	.19	1 08
m	.82	.63	− 48
n	.79	.31	−.36
o	.77	−.65	−.28
p	88	−.48	.17
q	.90	−.68	.26
r	.80	.10	.12
s	.76	−.10	−.86
t	.38	16	1.00
u	.52	.32	−.26
v	.39	.81	− 14
w	.35	.80	−.70
x	.41	1.21	− 07
y	.71	.56	.13
z	.49	.32	−.92

dimensions was considerably higher (.178). Accordingly, the dimensional difference between the studies is more apparent than real, since if we had ignored the skree plot and selected a lower stress value, the solution would have included more dimensions. In any case, Table 2 gives the coordinates for the 3-dimensional solution.

Clearly, Dimension 1 is closely related to the upper- versus lowercase status of the letters. All of the 26 uppercase letters show negative values on the dimension, while 25 of the 26 lowercase letters show positive values. The exception was "l," which some subjects may in fact have interpreted as "I."

Dimension 2 also appears interpretable. It yields strongly negative values for a number of curved letters (e.g., B, D, G, O, Q) and strongly positive values for a number of letters containing straight slanted lines (e.g., V, X, Y, k, v, w, x). While some slanted letters produced negative values (e.g., A, Z), their magnitudes were small. Letters composed only of horizontal and vertical lines, neither curved nor slanted, produced intermediate values. Thus, it seems appropriate to label Dimension 2 as curved versus slanted.

Dimension 3 appears to capture acute angular versus vertical letters. The former are composed of acute angles, and they tend to have negative values on the dimension (e.g., S, Z, s, w, z). Vertical letters have a predominantly vertical component and show positive values (e.g., I, J, T, f, i, l, t).

The symmetrized ratings were also subjected to three hierarchical clustering methods: complete linkage (Johnson, 1967), and the weighted (WPGMA) and unweighted (UPGMA) pair-group methods using arithmetic averages (Sokal & Michener, 1958). Of these, WPGMA gave the most satisfactory results in terms of interpretability. The dendogram for the results of this analysis is given in Figure 2.

One can differentiate the clusters at either a proximal or a distal level, but general groupings can be described by examining those clusters emerging at the "20" level or higher. From left to right on Figure 2, the first such grouping (Z, z, N, S, s) captures a number of the letters composed of acute angles, while the second (E, F, L) contains letters showing both strong horizontal and vertical components. The third group (B, D, C, c, O, o, Q, G) contains mostly uppercase curved letters, along with their lowercase visual analogues, and the fourth (K, k, X, x, H) contains vertically bisected letters that are symmetrical or nearly so about a horizontal axis. The fifth group contains largely diagonal letters (W, w, M, V, v, U, u, y, Y, A). The sixth has letters with a strong vertical component (f, t, J, j, I, l, i, T), and the seventh has small humped letters (n, h, r). The eighth and last group contains mostly curved lowercase letters (a, e, g, q, d, b, P, p, R). The difference between the third and eighth groups may be that in the third, lowercase letters have clustered with their look-alike uppercase analogues (C, c, O, o), whereas in the eighth, the lowercase letters are mostly nonanalogues. The exceptions, "p" and the physically similar letters "P" and "R," may have clustered with the rest of the eighth group because of the similarity of "p" with "q," "b," and "d."

Comparing the outcomes of the MDS and cluster analyses, it is apparent that they were sensitive to somewhat differing aspects of letter similarity. The MDS analysis was considerably more sensitive to the case (upper vs. lower) of the letters. The cluster analysis was more sen-

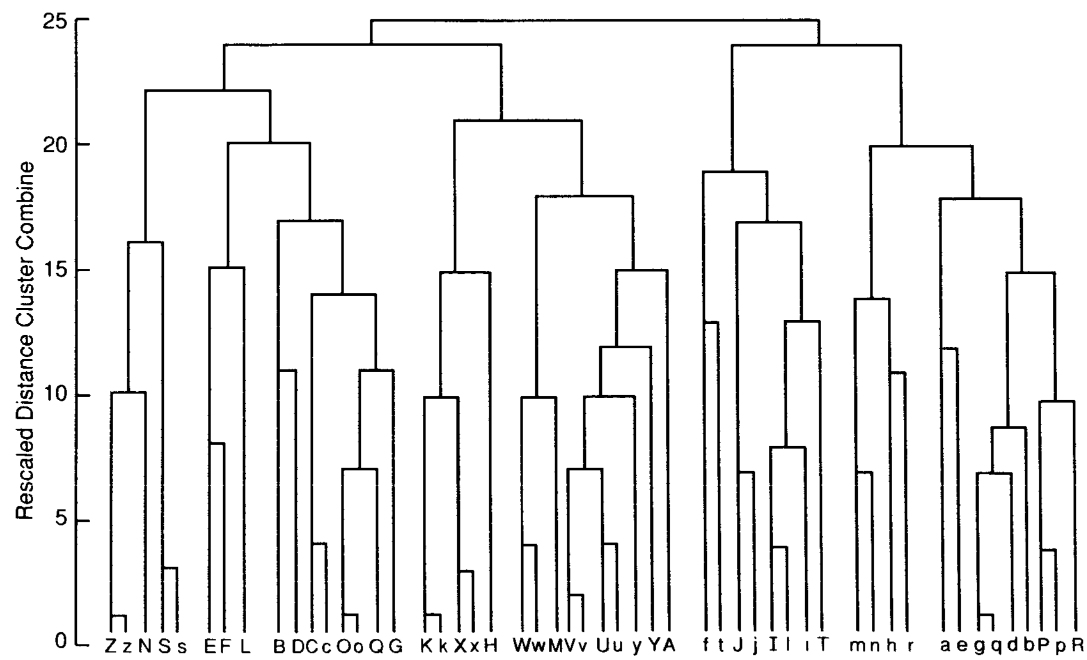

Figure 2. The WPGMA cluster analysis.

sitive to featural similarities. Specifically, clustering revealed eight groupings based on various combinations of acute angles, horizontal, diagonal and vertical lines, curves, and bisections, with some influence of case. The MDS analysis produced case as a nearly pure dimension, but other dimensions were sensitive to a reduced number of features, including acute angles, diagonal and vertical lines, and curves.

The differential outcome is interesting, because instead of considering dimensional and clustering analyses as complementary in nature, attention has often been directed toward considering one type of analysis as more appropriate than the other. Indeed, Pruzansky, Tversky, and Carroll (1982) provided statistical procedures for classifying a data set as favoring "spatial" (e.g., MDS) or "tree" (e.g., clustering) analyses. In contrast, our results suggest that useful information can be gained from combining approaches rather than selecting one over the other. In this, our conclusion differs from those of Gilmore et al. (1979) and Townsend (1971), both of whom concluded that a MDS solution yielded dimensions that were not readily interpretable. Yet even a cursory glance at Dimension 1 in Table 2 demonstrates ready interpretability. It is of course precisely on the methodological variable of upper- versus lowercase that our study differs from the previous ones.

In conclusion, the structural aspects of Table 2 and Figure 2 may be of interest to theorists wanting to model letter-similarity judgments. On the other hand, researchers interested in the effects of visual similarity on same-different judgments should find the data in Appendixes A and B to be of greatest interest.

REFERENCES

BAGNARA, S., BOLES, D. B., SIMION, F., & UMILTA, C. (1983). Symmetry and similarity effects in the comparison of visual patterns. *Perception & Psychophysics*, **34**, 578-584.

BARNES, S., & BURKE, R. S. (1988). What is Apple-Psych? *Behavior Research Methods, Instruments, & Computers*, **20**, 150-154.

BESNER, D., COLTHEART, M., & DAVELAAR, E. (1984). Basic processes in reading: Computation of abstract letter identities. *Canadian Journal of Psychology*, **38**, 126-134.

BOLES, D. B. (1986). Confusion and visual field effects in rhyme and name letter matching. *Perception & Psychophysics*, **39**, 123-128.

BOLES, D. B. (1988) Voice recognition with the Apple-Psych system *Behavior Research Methods, Instruments, & Computers*, **20**, 158-163

BOLES, D. B., & EVELAND, D. C (1983). Visual and phonetic codes, and the process of generation in letter matching. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 657-674.

DAINOFF, M. J., & HABER, R. N. (1970). Effect of acoustic confusability on levels of information processing. *Canadian Journal of Psychology*, **24**, 98-108.

GERVAIS, M. J., HARVEY, L. O., & ROBERTS, J. O. (1984) Identification confusions among letters of the alphabet. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 655-666

GILMORE, G. C., HERSH, H., CARAMAZZA, A., & GRIFFIN, J. (1979). Multidimensional letter similarity derived from recognition errors. *Perception & Psychophysics*, **25**, 425-431.

JOHNSON, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, **32**, 241-254.

KRUSKAL, J. B. (1964a). Multidimensional scaling optimizing goodness of fit to a non-metric hypothesis. *Psychometrika*, **29**, 1-27

KRUSKAL, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, **29**, 115-129

KRUSKAL, J. B., YOUNG, F. W., & SEERY, J. B. (1977). *How to use KYST 2A, a very flexible program to do multidimensional scaling and unfolding*. Murray Hill, NJ: AT&T Bell Laboratories

LOOMIS, J. M (1982). Analysis of tactile and visual confusion matrices. *Perception & Psychophysics*, **31**, 41-52.

OSGOOD, G. (1988). Generalizing the Apple-Psych system. *Behavior Research Methods, Instruments, & Computers*, **20**, 155-157.

POSNER, M. I. (1978). *Chronometric explorations of mind*. New York. Erlbaum.

POSNER, M. I., BOIES, S. J., EICHELMAN, W. H., & TAYLOR, R. L (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monographs*, **79**(1, Pt. 2).

PRUZANSKY, S., TVERSKY, A., & CARROLL, J. D. (1982). Spatial versus tree representations of proximity data. *Psychometrika*, **47**, 3-19.

SOKAL, R. R., & MICHENER, C. D. (1958). A statistical method for evaluating systematic relationships *University of Kansas Science Bulletin*, **38**, 1409-1438

THORSON, G., HOCHHAUS, L., & STANNERS, R. F. (1976). Temporal changes in visual and acoustic codes in a letter-matching task *Perception & Psychophysics*, **19**, 346-348

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

VAN DER HEIJDEN, A. H. C., MALHAS, M. S. M., & VAN DEN ROOVART, B. P. (1984). An empirical interletter confusion matrix for continuous-line capitals. *Perception & Psychophysics*, **35**, 85-88.

NOTE

1. A diskette formatted in Apple Pascal is available on request at no cost to anyone desiring facilitated access to the similarity ratings for modeling or other purposes. Send the request to David B. Boles, Department of Psychology, Rensselaer Polytechnic Institute, Troy, NY 12180.

(Continued on following page)

APPENDIX A
Mean Similarity Values, Each of Two Decimal Places

	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	467	225	225	242	216	275	216	266	175	208	242	192	233	275	216	275	208	292	200	200	275	266	200	233	242	167
B	242	500	250	333	292	283	275	225	242	183	250	225	200	175	292	350	208	350	275	200	225	175	150	192	233	158
C	250	258	467	258	258	167	342	192	200	225	208	225	208	167	383	216	325	208	275	208	308	225	183	183	175	216
D	233	350	308	483	308	225	308	225	225	233	183	225	216	225	342	292	292	225	167	216	250	208	175	183	183	200
E	200	316	316	266	492	342	216	225	225	216	233	275	275	242	192	308	183	275	300	242	242	167	242	225	175	292
F	250	258	208	242	392	500	225	216	242	175	250	242	225	216	183	350	192	316	192	250	192	167	175	158	192	225
G	258	233	292	250	292	233	500	233	200	200	192	200	200	167	366	216	350	216	275	200	216	183	200	175	175	225
H	275	225	175	208	292	250	208	500	292	216	250	308	233	275	167	283	200	208	208	233	242	250	258	250	225	242
I	175	233	216	200	250	292	200	250	500	300	250	266	208	200	233	258	183	225	183	316	242	192	200	175	208	183
J	208	225	266	208	250	200	208	192	275	500	233	225	208	200	233	233	225	233	200	258	258	242	200	167	283	175
K	233	208	208	200	250	200	183	233	233	183	475	266	225	266	183	183	208	275	183	216	192	258	216	350	216	233
L	208	225	266	275	300	275	208	258	275	292	225	500	216	242	183	216	192	216	167	258	325	225	216	192	225	225
M	250	192	175	200	225	200	183	242	250	167	216	242	475	350	200	208	183	225	208	216	216	216	366	250	167	216
N	216	216	200	225	200	192	208	233	216	233	275	216	342	500	216	183	216	233	192	242	242	242	258	308	233	333
O	242	275	366	300	242	208	350	208	192	250	216	183	158	208	500	225	383	233	208	183	375	208	192	167	175	233
P	275	316	225	283	275	333	275	233	250	258	225	266	183	200	275	500	242	358	275	208	225	175	183	216	233	242
Q	216	233	333	292	200	216	316	175	167	242	192	150	192	208	350	216	500	216	266	208	308	167	216	175	183	216
R	266	292	275	216	258	266	233	258	242	200	300	208	216	200	233	408	216	492	242	200	233	167	200	242	175	242
S	192	258	258	192	233	200	208	192	133	183	183	183	158	183	250	275	216	266	500	192	208	142	183	258	208	325
T	225	266	216	225	242	292	175	266	292	283	258	250	200	242	183	250	200	250	158	500	216	192	167	167	233	208
U	258	225	342	250	242	216	250	258	225	333	216	250	250	266	342	216	258	208	216	192	500	383	308	216	258	242
V	233	142	225	200	208	175	192	208	216	200	167	200	266	225	233	208	183	167	192	192	350	500	242	275	342	208
W	208	200	150	200	242	192	200	258	183	167	225	233	333	283	175	225	192	225	175	216	292	333	492	266	250	216
X	250	183	200	167	208	183	167	292	200	167	325	200	242	308	175	175	167	192	192	200	208	300	242	500	308	242
Y	216	183	233	192	150	200	158	250	192	225	225	167	242	208	200	266	175	208	175	266	300	375	275	275	500	175
Z	183	242	167	200	258	275	225	183	216	183	175	233	216	316	216	183	192	200	258	225	233	192	183	192	175	500
a	250	175	150	150	142	125	125	125	108	167	108	125	133	158	158	175	150	125	158	117	142	150	125	125	117	142
b	150	250	175	216	167	133	208	175	167	183	142	200	142	158	225	258	208	192	142	167	158	150	125	125	150	133
c	133	216	366	158	225	158	216	175	133	150	158	175	158	117	233	200	175	158	208	150	275	192	183	150	200	150
d	158	167	183	242	167	167	150	133	175	183	150	167	150	158	158	250	167	183	167	167	183	142	150	150	158	133
e	150	158	167	158	266	158	258	133	117	150	125	133	125	117	208	183	167	150	216	125	167	192	150	167	125	167
f	233	167	158	158	167	350	158	158	183	208	150	150	142	183	133	283	125	150	150	208	167	133	108	150	142	133
g	125	150	167	150	142	150	225	133	125	158	133	142	142	117	192	192	175	175	175	125	150	117	125	150	150	125
h	175	175	183	175	142	158	183	308	167	175	233	216	183	242	208	192	192	175	133	208	225	192	142	158	175	158
i	125	125	125	125	167	158	125	150	333	158	150	200	133	150	133	133	158	142	150	233	167	117	133	133	158	133
j	167	133	208	133	158	158	158	175	216	400	183	175	133	150	150	142	133	175	150	192	183	192	142	133	175	158
k	125	142	158	142	192	167	117	150	183	142	450	233	158	192	133	158	150	200	142	142	175	167	192	258	192	158
l	150	167	183	216	233	183	150	225	400	233	192	242	200	183	133	200	158	167	142	342	175	208	183	142	183	167
m	167	150	133	150	192	183	142	167	133	125	150	125	316	200	150	133	100	133	133	158	150	142	192	142	150	142
n	208	175	175	183	117	158	167	183	142	158	133	133	192	283	183	158	183	142	133	133	208	183	167	142	183	133
o	150	158	250	200	125	133	167	117	117	150	150	125	117	125	442	242	375	150	175	133	216	142	142	158	142	125
p	125	158	158	167	125	183	175	125	125	133	117	167	150	125	216	433	192	208	142	158	133	167	150	117	150	142
q	133	150	167	192	150	125	142	125	125	150	150	158	125	142	192	275	233	208	133	175	117	125	117	117	158	108
r	192	125	233	125	175	183	167	150	142	175	167	158	167	175	133	225	108	233	142	175	150	192	142	158	175	150
s	125	167	158	125	225	175	142	142	125	142	117	117	158	150	158	158	150	158	400	142	133	158	158	142	133	258
t	125	167	175	125	142	175	133	175	167	192	175	167	125	150	117	175	117	150	142	316	183	167	167	158	175	183
u	158	175	216	216	142	133	175	192	142	258	158	183	150	192	200	158	183	133	142	125	400	325	242	167	283	167
v	192	158	192	133	125	133	158	150	142	200	167	142	175	175	183	133	158	125	125	150	333	450	250	225	292	150
w	150	158	142	133	233	150	150	167	158	150	125	150	283	208	142	133	117	142	133	125	225	242	392	158	175	125
x	133	133	142	117	158	100	125	150	150	133	283	158	150	125	108	125	125	125	133	133	133	200	175	433	167	158
y	167	158	158	150	158	125	167	200	133	208	192	150	158	142	192	158	150	125	142	142	250	216	167	216	300	167
z	142	133	158	133	150	208	142	158	142	125	167	175	183	275	125	142	142	158	208	142	158	158	175	167	125	442

Note—left column = first letter in pair. Top row = second letter.

Appendix A (Continued)

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	
225	142	175	158	158	175	117	258	125	133	150	158	167	192	150	167	167	167	133	192	167	167	175	125	150	133	A
142	275	167	200	192	158	175	192	142	150	142	150	133	125	192	216	175	158	183	158	142	142	133	117	158	142	B
158	225	458	258	242	150	142	175	133	200	125	142	133	183	216	158	167	216	133	133	225	200	125	125	150	142	C
125	258	242	242	192	125	158	183	125	183	133	200	183	183	192	175	175	142	133	125	192	142	133	117	158	133	D
158	167	208	117	250	192	142	150	142	133	183	183	175	167	133	167	125	133	216	167	167	150	175	158	150	192	E
133	158	183	142	150	366	125	208	192	150	158	175	158	158	150	208	150	200	142	192	133	133	150	117	133	158	F
167	225	266	200	225	150	216	142	150	150	117	133	175	167	192	183	192	183	158	167	200	150	133	125	150	133	G
125	192	150	175	142	158	158	333	175	150	208	200	175	216	133	133	158	158	133	142	192	125	200	175	192	200	H
133	175	142	167	133	183	117	200	292	167	192	417	150	175	125	175	158	158	150	192	117	142	142	125	117	125	I
125	158	175	175	150	225	175	158	158	358	183	242	142	158	158	192	125	175	158	258	142	216	150	133	233	150	J
125	175	133	150	133	183	133	183	125	142	417	200	158	158	142	117	142	192	133	150	158	192	125	275	150	175	K
133	200	192	158	150	150	133	192	158	175	258	266	117	150	125	142	125	175	133	167	183	175	142	133	158	175	L
142	167	150	133	142	133	125	200	142	142	200	183	350	225	133	125	125	200	150	133	158	200	283	167	142	167	M
133	158	158	125	133	117	133	250	158	125	225	208	200	325	125	133	125	158	150	133	183	175	225	208	167	292	N
192	250	275	216	216	150	167	175	158	183	133	167	142	150	425	192	200	175	158	150	250	183	142	158	158	158	O
167	283	208	258	216	258	192	192	183	183	150	183	142	125	242	392	308	242	158	150	183	125	133	142	142	167	P
158	150	208	192	158	117	167	142	117	158	158	175	117	158	316	167	216	175	150	167	233	133	133	125	167	158	Q
167	208	208	200	167	192	175	158	133	167	242	142	158	150	167	275	158	233	158	150	158	158	117	175	133	167	R
216	175	225	150	183	158	216	175	133	133	133	158	133	150	167	158	142	192	433	150	142	142	150	125	167	242	S
142	150	133	175	142	208	117	175	208	142	142	292	167	175	150	142	142	175	142	300	125	158	158	117	150	142	T
133	183	258	183	133	150	133	216	133	225	192	216	200	258	250	142	142	167	142	133	417	342	258	142	275	142	U
125	142	192	158	150	142	125	192	142	158	192	158	150	183	167	133	133	150	167	158	350	408	250	233	250	150	V
117	167	133	142	142	133	133	142	150	142	233	192	233	175	133	108	142	167	158	150	292	266	425	192	175	150	W
108	133	167	150	125	133	117	142	133	108	275	125	150	175	133	125	133	142	142	150	208	250	200	400	200	175	X
125	125	216	142	142	167	167	208	142	175	167	216	150	167	142	150	167	183	125	200	258	283	200	158	316	125	Y
167	125	158	133	142	142	125	117	142	142	142	150	150	158	133	117	125	167	242	142	150	133	167	167	133	442	Z
500	266	233	233	292	200	242	183	200	158	133	158	250	225	300	225	242	225	258	192	192	200	200	167	150	242	a
225	500	292	350	225	183	266	342	175	167	250	225	167	225	283	350	375	225	233	200	250	192	150	150	216	150	b
283	242	483	242	300	142	258	258	158	216	158	133	200	308	342	250	242	292	275	200	325	292	208	167	258	233	c
216	383	225	500	208	208	375	333	183	208	225	216	175	167	266	333	350	233	183	242	233	200	158	158	216	150	d
375	308	325	250	483	208	300	183	158	175	175	142	192	233	325	283	242	308	283	200	266	192	183	183	158	250	e
200	175	208	192	216	500	250	208	192	250	167	225	167	183	142	183	175	225	175	300	167	216	175	150	183	175	f
325	325	266	275	233	200	500	216	175	266	208	175	225	208	258	316	417	175	233	200	208	183	167	175	383	158	g
183	316	233	283	216	192	167	500	192	225	333	192	242	325	183	266	183	342	150	208	308	250	142	150	266	175	h
167	142	167	183	150	192	200	175	467	358	175	392	200	208	183	183	142	192	183	266	216	183	158	158	192	175	i
175	183	225	208	167	275	283	200	350	500	175	266	167	208	192	250	192	258	208	308	233	250	167	183	225	200	j
158	266	200	216	150	158	167	325	175	150	500	192	183	150	142	208	150	208	175	183	175	233	200	366	242	183	k
158	208	125	225	158	283	158	208	366	208	225	500	142	167	142	158	158	167	175	300	158	142	142	167	150	158	l
200	175	225	192	208	200	216	216	192	133	167	167	500	375	225	200	167	300	192	183	233	208	316	233	183	250	m
242	216	350	225	216	192	183	358	192	216	133	133	375	500	242	216	216	308	175	175	333	225	233	175	233	208	n
250	308	316	316	325	150	233	208	200	216	158	133	225	250	500	333	258	192	233	158	333	283	167	150	200	158	o
258	350	258	300	250	183	325	242	200	216	183	150	175	216	300	500	408	266	233	216	200	183	233	158	258	216	p
216	342	266	383	208	192	467	258	158	167	175	200	167	225	242	408	500	208	183	183	216	183	183	158	258	175	q
250	242	283	283	275	275	242	333	208	233	167	125	266	300	266	316	225	500	216	266	258	200	183	242	216	200	r
258	167	292	225	292	208	216	158	192	175	133	125	225	233	200	258	192	208	483	200	208	192	242	242	192	308	s
158	258	208	225	150	342	208	200	200	233	233	300	200	200	192	192	158	192	183	467	192	233	167	183	192	216	t
283	183	316	225	266	183	225	316	192	266	200	183	225	333	292	208	216	275	200	192	442	400	366	208	400	225	u
150	183	292	175	192	167	192	225	175	225	200	167	175	266	250	183	200	258	208	183	400	500	308	283	325	242	v
175	150	216	150	167	175	183	208	192	175	150	167	350	275	150	167	175	183	242	183	333	325	500	192	242	258	w
208	167	225	158	208	142	175	183	192	158	392	142	200	175	167	158	200	200	175	183	225	258	183	500	175	258	x
233	250	216	250	216	175	366	250	175	266	216	158	216	225	208	233	275	225	216	192	358	350	258	233	500	192	y
250	183	283	133	292	158	192	183	192	158	200	142	192	233	175	192	167	258	333	183	233	175	266	242	233	500	z

APPENDIX B
Mean Generation and Veridical Similarity Values with Order Collapsed

	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	484	240	248	232	272	232	261	227	179	187	191	179	233	240	253	259	222	259	227	194	252	212	195	214	210	210
	237	152	154	148	152	166	133	171	123	148	126	141	152	172	163	159	152	163	159	144	150	159	142	123	140	145
B		500	261	355	286	225	275	278	198	190	244	221	183	208	290	342	290	278	233	232	221	174	163	174	221	183
		263	195	210	171	154	190	183	152	156	150	180	148	154	207	230	171	170	167	161	165	148	145	127	148	133
C			475	259	300	182	290	214	185	233	193	188	202	257	353	237	292	265	275	208	323	259	190	193	221	225
			413	210	210	162	198	171	133	183	143	173	143	158	244	181	179	204	181	147	244	194	145	145	181	152
D				492	259	216	302	263	198	214	206	235	195	210	307	302	330	240	192	228	240	195	171	166	210	170
				242	159	148	165	166	148	169	143	185	153	162	192	212	182	163	144	148	193	143	139	133	152	133
E					488	290	261	230	195	203	203	218	225	223	271	280	208	280	277	208	255	190	208	206	175	273
					259	167	192	141	140	147	159	174	159	133	171	173	150	157	210	144	152	154	175	152	144	163
F						500	227	216	229	225	194	257	198	195	170	263	194	271	193	296	190	182	179	158	188	208
						359	145	170	179	184	165	165	154	154	141	233	129	181	157	195	145	134	134	125	142	160
G							500	207	194	240	188	184	206	191	303	284	388	216	233	195	225	188	188	173	271	200
							221	154	129	160	125	139	145	145	180	185	169	175	173	135	165	137	134	129	159	131
H								500	228	208	286	242	233	298	191	257	205	286	177	227	282	233	216	218	249	195
								321	173	165	193	208	181	223	158	161	154	160	145	175	207	165	163	157	193	158
I									484	321	208	326	212	205	202	223	163	216	173	270	218	191	183	182	192	192
									313	175	163	294	139	157	133	154	139	143	139	201	140	135	145	134	137	135
J										500	185	249	169	214	223	240	207	232	191	271	273	230	177	169	250	179
										379	163	207	135	148	160	163	141	173	145	195	202	192	145	126	198	143
K											488	228	198	206	175	199	182	237	168	223	195	214	198	359	226	197
											434	221	166	177	139	135	150	201	131	152	171	180	169	273	175	161
L												500	192	190	160	198	175	179	163	277	229	183	190	175	175	190
												255	156	168	137	173	153	161	137	242	190	171	167	139	177	167
M													488	361	202	191	178	253	195	200	232	216	342	232	203	218
													334	205	135	137	117	165	143	145	165	167	248	152	150	161
N														500	230	204	216	261	195	214	294	240	263	242	225	273
														304	145	134	152	156	145	147	210	179	194	163	165	214
O															500	284	309	231	223	179	336	244	171	165	195	195
															434	223	271	157	165	137	230	169	139	139	159	134
P																500	318	338	261	216	212	187	202	177	248	208
																413	235	237	154	157	154	139	131	127	150	142
Q																	500	216	214	187	250	183	191	175	223	188
																	225	162	143	150	169	137	127	125	161	133
R																		496	233	227	244	198	198	218	206	225
																		233	163	163	152	156	142	150	154	161
S																			492	183	208	183	210	216	197	307
																			417	144	139	148	150	135	142	237
T																				484	198	200	183	183	221	208
																				309	141	159	150	139	167	152
U																					471	384	325	214	329	233
																					409	338	255	163	266	155
V																						500	302	280	349	205
																						429	253	227	261	147
W																							496	222	257	232
																							409	182	180	154
X																								500	248	233
																								417	185	167
Y																									500	194
																									309	137
Z																										500
																										442

Note—upper entry = generation similarity. Lower entry = veridical similarity.

(Manuscript received December 22, 1988;
revision accepted for publication October 27, 1989.)