

# Analysis of tactile and visual confusion matrices

JACK M. LOOMIS

University of California, Santa Barbara, California 93106

Confusion matrices were compiled for uppercase letters and for braille characters presented to observers in two ways: as raised touch stimuli and as visual stimuli that had been optically filtered of their higher spatial frequencies. These and other existing matrices were subjected to a number of analyses, including the choice model and hierarchical clustering. The strong similarity of the visual and tactile matrices from this study lends additional support to the claim that visual recognition of low-pass filtered characters, to a first approximation, can be taken as a model of tactile recognition of small two-dimensional raised patterns. Besides this, the analysis questions the widely held assumption that response bias contributes significantly to the stimulus-response contingencies in a character-recognition task.

In recent work, Loomis (1981) demonstrated a close parallel between tactile recognition of raised letters and braille, using the finger and visual recognition of the same characters after they had been subjected to optical low-pass spatial filtering. In particular, the functions relating recognition performance to character size were quite similar for both modalities, with braille characters allowing better performance than letters of matched size throughout the range of stimulus size. One conclusion of that study was that visual recognition of low-pass filtered characters can be taken as an approximate model of tactile recognition of small raised characters. It follows from this that a more detailed look at the confusions between characters (either letters or braille) should also show the similarity between touch and vision. Because of the large number of character sets investigated in that study, insufficient data were gathered for any one character set to justify an analysis of the confusion errors. In order to carry out such an analysis, the present work involved the collection of much more extensive data on just two sets of characters, one comprising letters and the other, braille.

## METHOD

The method employed was essentially the same as that used in Experiment 3 of Loomis (1981), with the major difference being that the total set of stimuli in this study was much smaller. The method will be sketched here, but the interested reader should refer to Loomis (1981) for details.

This research was supported by Grant 1 R01 NS15129 from NINCDS. The author thanks Deborah Williams and Pauline McNulty for assistance in collecting and analyzing the data, Lawrence Hubert for help with the cluster analysis, Kenneth Johnson for preprints of his work, and John Foley and the reviewers for comments on an earlier draft of the manuscript.

## Stimuli

The letter and braille character sets used as stimuli are shown in Figure 1. They were first prepared using transfer lettering materials. The letters were taken from Letraset sheet No. 2665 (Helvetica Extra Light), while the braille characters were composed of dots from Chartpak sheet M451. Both stimulus sets were those employed in Experiment 3 of Loomis (1981). In the composition of the artwork, the two character sets were arranged in circles. The artwork was taken to a photoengraver, who created a circular plate out of zinc with the letters and braille characters that would serve as touch stimuli raised above the surface. The visual stimuli, prior to low-pass spatial filtering, were the transparent characters in the photographic negative used by the photoengraver to make the plate. In spatial extent, they were of the same physical dimensions as the touch stimuli. Prior to viewing, they were filtered of their higher spatial frequencies by means of blurring by diffusion, per the method of Carlson and Heyman (1979). The optical filter was adjusted so as to equate visual resolution to tactile resolution; as in Experiment 3 of Loomis (1981), the point spread function in the output plane was closely approximated by a Gaussian of half-power width equal to 5.8 mm.

In the previous work, different sizes of the letter and braille character sets were studied in order to determine how tactile and visual recognition varied with character size. Here it was desired to choose for each of letters and braille, the character set size that would yield close to 50% recognition accuracy. Based on the results of the previous work (Loomis, 1981, Figure 5), braille charac-

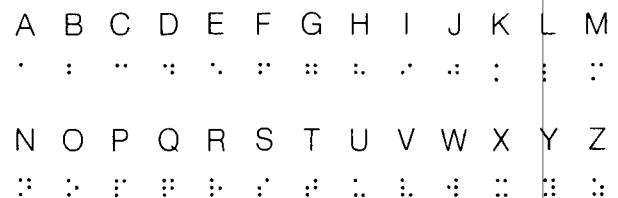


Figure 1. The letters and braille characters used in this experiment. The touch stimuli were raised facsimiles of the characters in this figure. The visual stimuli were back-illuminated transparencies of the above subjected to optical low-pass filtering prior to viewing. The letters were 5.7 mm high, while the braille characters were based on a cell height of 4.3 mm. Each braille character corresponds alphabetically with the character above it.

ters based on a cell height of 4.3 mm and uppercase letters 5.7 mm high were selected. Figure 1 represents the actual relative sizes of the two character sets.

The visual stimuli were viewed with the right eye in a darkened room from a distance of 104 cm. Each millimeter in the plane of the visual stimuli thus subtended 3.3 min of arc. Stimulus duration was 2 sec as controlled by a shutter. During the presentation, subjects were free to scan the stimulus. All other details were as before.

The touch stimuli were sensed by the distal pad of the right index finger using a force of roughly 5-10 g. The same mode of touch used in Experiment 3 of Loomis (1981) was employed here. The subjects made very slight circular motions of the finger as it rested on the character, taking care that there be no slippage between the corneum of the epidermis and the character. The subjects were permitted to feel each stimulus for approximately 2 sec.

### Subjects

Six females, ranging in age from 18 to 24 years, participated as paid subjects. Three of these, run in 1980, had had at least 30 h of experience as subjects in experiments on touch, either with raised characters or with the Optacon, a vibrotactile reading aid for the blind. The other three, run in 1981, were naive psychophysical observers. All six subjects had normal or corrected-to-normal vision and ostensibly normal touch. None of the subjects was aware of the purpose of the experiment.

### Procedure

Before the experiment proper, the subjects were taught the alphabetic characters of Standard English braille using flashcards. Training continued until no identification errors were made in 10 successive runs through the alphabet. Each session of the experiment proper involved the presentation of the full complement of alphabet characters for each of the four conditions. A change in conditions occurred every 13 characters with the orders of conditions and characters within conditions being random. The subject was given feedback after each trial. Four subjects participated for 36 sessions (each lasting nearly 1 h), while the remaining two participated for 51 sessions. Each character, therefore, was presented a total of 246 times within each of the four conditions.

## RESULTS

Averaged over subjects and characters within each set, recognition performance proved to be somewhat better than the targeted level of 50% correct. The average recognition accuracies were 62%, 60%, 61%, and 64% for letters/touch, letters/vision, braille/touch, and braille/vision, respectively. Because the braille characters were based on a characterspace size having an area less than 60% that of the letterspace size, these results reconfirm the finding of Loomis (1981) that braille characters are more identifiable than uppercase letters of the *same* dimensions for both senses.

The tactile and visual confusion matrices for letters are presented in combined form in Table 1 and those for braille, in Table 2. In each cell of the tables, the tactile and visual data are given as the upper and lower entries, respectively. Each entry is the proportion of trials on which the row stimulus resulted in the column response and is referred to as the stimulus-response contingency; each proportion is based on 246 presentations of the stimulus. Anticipating later

discussion, the diagonal cell values will also be referred to as hit rates.

Some of the analyses to follow were based not on the confusion matrices of Tables 1 and 2, but on symmetrized transformations of them. If  $P_{ij}$  is the original contingency between stimulus  $i$  and response  $j$ , the symmetrized contingency  $P'_{ij}$  is simply:

$$P'_{ij} = (P_{ij} + P_{ji})/2,$$

Although the confusion error (off-diagonal) values between two stimuli exhibit significant asymmetry in some cases, as Kikuchi, Yamashita, Sagawa, and Wake (1979) have noted, symmetrizing the matrix values has the effect of reducing the uncorrelated variation between symmetrically corresponding off-diagonal values relative to the correlation confusion tendencies.<sup>1</sup> One indication of the noise reduction is the increase in the reliability coefficients<sup>2</sup> computed for all the off-diagonal cell values in each matrix as a result of symmetrizing; the reliabilities went from .88 (letters/touch), .89 (braille/touch), .86 (letters/vision), and .87 (braille/vision) for the original matrices up to .94, .92, .93, and .93, respectively, for the symmetrized matrices. Naturally, the diagonal cell values (the hit rates) remain unchanged by symmetrizing. The reliability coefficients for these values were .96 (letters/touch), .96 (braille/touch), .90 (letters/vision), and .84 (braille, vision).

If, as claimed, visual recognition of low-pass filtered characters can be taken as a model of tactile recognition of small raised characters, then one would expect the symmetrized confusion matrices to be very similar for the two modalities for both types of characters. One way of assessing the similarity of two matrices is to compute the product-moment correlation coefficient separately for the hit rates and for the confusion error values. (Computing the correlation for all  $26 \times 26$  cells would be misleading, for the correlation coefficient approaches 1.0 as the two recognition tasks being compared are made very easy, even when the tasks and stimuli are not at all alike.)

For uppercase letters, the correlations obtained between vision and touch were .84 for the hit rates and .83 for the (symmetrized) off-diagonal values; for braille characters, the values were .82 for the hit rates and .79 for the confusion error values. To get a comparative sense of how these correlations fare, it is instructive to examine the same measures computed between a number of other tactile and visual confusion matrices computed by other investigators. Each of the studies compared is briefly described below. Except where otherwise stated, the dependent variable of ultimate interest was the stimulus-response contingency.

*K. O. Johnson and Phillips (1981).* The conditions were very similar to those of the present study involving touch. Subjects





attempted to identify raised uppercase letters 4.5 mm in height presented for 4 sec to the distal pad of the right index finger. No lateral motion of the finger over the letter was permitted. Contact force was constant at 60 g.

*Loomis (1974)*. Uppercase letters, 204 mm in height, were presented statically for 1.5 sec to the backs of observers, using a 20 × 20 array of solenoid-driven stimulators. Vibration frequency was 40 Hz.

*Kikuchi, Yamashita, Sagawa, and Wake (1979)*. Uppercase letters, 208 mm in height, were presented to the backs of observers using a 17 × 17 array of air-driven stimulators. Vibration frequency was 2.5 Hz. Observers were permitted to move relative to the letters and were given unlimited time for sensing them.

*Craig (1979)*. Uppercase letters, 20 mm in height, were presented statically to the ventral surface of the finger, using the 24 row × 6 column vibrotactile array of the Optacon, a commercially available reading aid for the blind. Its piezoelectric bimorph stimulators vibrate at 230 Hz when activated. The letters presented in this study were much larger than those used by Johnson and Phillips (1981) and those used in the present study, but recognition accuracy was held to 52% by using short stimulus durations ranging from 4 to 100 msec. For most of the trials, the left index finger was used.

*Gilmore, Hersh, Caramazza, and Griffin (1979)*. Uppercase dot matrix letters, 20 min high in visual angle, were presented tachistoscopically by computer to observers who viewed them foveally. Stimulus duration ranged from 10 to 70 msec, with a mean of 33 msec. The letters were green in color against a dark background and were preceded and followed by a dark field of at least 1 sec duration.

*Townsend (1971)*. Uppercase letters, 30 min high in visual angle, were presented tachistoscopically to observers who maintained fixation 10 min of arc beneath the position of the letter. The letters were black against a uniform white background with the pre- and poststimulus fields being white of the same luminance (not including the fixation point).

*Nolan and Kederis (1969)*. All 63 braille characters based on the 2 × 3 braille cell were presented for very short durations to the fingertip using a device called the tachistotactometer. The braille characters were of essentially standard dimensions (4.6 mm high and 2.3 mm measuring between the centers of the appropriate cell positions). The confusion matrices were not reported. Instead, the authors gave for each character the threshold duration required for accurate identification. For purposes of comparison with the other studies here, the recognition thresholds for the 26 alphabetic characters were taken as the legibility measures after sign inversion. (Since a low recognition threshold presumably signifies high legibility, the thresholds were inverted in sign so that positive correlations with other data sets would represent positive relationship).

*Bürklen (1932)*. In this study, the relative legibility of 39 of the braille characters was assessed. Experienced braille readers attempted to identify the characters in a variety of touching modes including static touching and scanning from left to right. Because characters more than twice normal height and width were em-

ployed, Bürklen found it necessary to cover the finger with a rubber cap to bring recognition performance down to a level at which confusions were frequent. His legibility values for the 26 alphabetic characters were used in the comparison with the measures from other studies. Confusion matrices were not reported.

For each confusion matrix, the character-specific hit rates were tentatively taken as measures of legibility to permit comparison with the legibility values of other studies. The product moment correlation coefficients computed between the legibility values are given in Table 3 (for letter recognition) and Table 4 (for braille recognition). The correlations calculated between the symmetrized confusion error (off-diagonal) values of experiments involving letter recognition are given in Table 5. (The only corresponding value for braille is the .79 obtained between the two conditions of this study, since the other studies of braille did not report confusion matrices.) It is clear that most of the matrices, whether they were obtained using touch or vision, are quite similar in terms of both legibility and confusion-error values. The one obvious exception is the matrix of Townsend (1971), the legibility values of which correlate only poorly with those of other experiments on letter recognition. It ought to be mentioned that the same correlations, when computed between one study involving letters and the other involving braille, are uniformly low in absolute value. As examples, the diagonal value and off-diagonal value correlations between letters/touch and braille/touch are only .11 and .08. Similarly, the same measures for letter/vision and braille/vision are .21 and .03. This fact indicates that the pattern of confusion errors for letters is quite different from that for the alphabetically corresponding braille characters, as one would expect, and that the hit rates for the different letters show little resemblance to those of the corresponding braille characters.

The reader might well object, as did a reviewer, that the diagonal cell values should be referred to as legibility values, for the term suggests that the relative frequency with which a character is correctly identified is solely a function of its discriminability from other characters in the set. The reviewer raised the possibility that the diagonal cell values might have

Table 3  
Matrix of Correlations Between Legibility Values for Letters

	Touch				Vision		
	Johnson & Phillips	Loomis (1974)	Kikuchi et al.	Craig	This Study	Gilmore et al.	Townsend
This Study, Touch	.81	.69	.74	.77	.84	.58	.46
Johnson & Phillips, Touch		.83	.81	.74	.85	.56	.22
Loomis (1974), Touch			.72	.85	.72	.72	.11
Kikuchi et al., Touch				.83	.76	.63	.48
Craig, Touch					.78	.88	.30
This Study, Vision						.68	.39
Gilmore et al., Vision							.21

**Table 4**  
Matrix of Correlations Between Legibility Values for Braille

	Nolan & Kederis, Touch	Bürklen, Touch	This Study, Vision
This Study, Touch	.72	.66	.82
Nolan & Kederis, Touch		.57	.68
Bürklen, Touch			.50

been contaminated significantly by the variation in tendency to respond with each of the 26 alphabetic names, so-called response bias. In the extreme, one might argue that all variation in the diagonal cell values reflects response bias; if this were so, the positive correlations between the diagonal cell values of different experiments might be ascribable to constancy of response bias across subjects in the different situations.

One plausible view of how response bias enters into the stimulus-response contingencies in an identification task is provided by the choice model (Luce, 1963), a model that others (Gilmore et al., 1979; Townsend, 1971) have since brought to bear on the analysis of letter confusion matrices. If  $P_{ij}$  represents the proportion of trials on which stimulus  $i$  results in response  $j$  and  $E(P_{ij})$  its expectation, the choice model assumes the following:

$$E(P_{ij}) = \frac{\eta_{ij} \beta_j}{\sum_{k=1}^{26} \eta_{ik} \beta_k}$$

where  $\eta_{ij}$  is a ratio scale measure of similarity and  $\beta_k$  is a ratio scale measure of response bias. (The model assumes that the bias to respond with a particular stimulus name is independent of the stimulus presented.) The intent of the choice model, when its assumptions are satisfied, is to allow decomposition of an empirical  $n \times n$  confusion matrix into its estimated underlying constituents, the matrix of similarity values  $[\eta_{ij}]$  for all possible pairs of stimuli and the response bias vector  $(\beta_1, \beta_2, \dots, \beta_n)$  representing the tendencies for responding with each of the  $n$  stimulus labels. The expressions for estimating  $\eta_{ij}$  and  $\beta_j$  are given in the Appendix of Townsend (1971).

In order to begin assessing the validity of this notion of response bias, each of the original confusion matrices discussed so far was analyzed by the choice model. (After conversion to proportions when necessary, the cell values of zero were converted to .001, the value Gilmore et al., 1979, found most satisfactory in fitting the choice model.) The choice model analysis resulted in a similarity matrix and response bias vector for each of the 10 matrices. As in the case with the 325 symmetrized off-diagonal values of the original matrix, the 325 similarity values of the different studies were correlated, taking two studies at a time. The intercorrelation matrix is given in Table 6. There is relatively little difference between the correlations in this table and those in Table 5 computed for the symmetrized off-diagonal values of the original matrix. The value comparing braille/touch and braille/vision was .70, which is close to the value of .79 for the off-diagonal values. Of greater interest is the matrix of correlations of response bias given in

**Table 5**  
Matrix of Correlations Between Confusion Error Values for Letters

	Touch				Vision		
	Johnson & Phillips	Loomis (1974)	Kikuchi et al.	Craig	This Study	Gilmore et al.	Townsend
This Study, Touch	.67	.47	.55	.67	.83	.70	.52
Johnson & Phillips, Touch		.49	.58	.47	.67	.55	.43
Loomis (1974), Touch			.52	.49	.57	.51	.38
Kikuchi et al., Touch				.54	.57	.52	.41
Craig, Touch					.65	.67	.56
This Study, Vision						.74	.61
Gilmore et al., Vision							.56

**Table 6**  
Matrix of Correlations Between Similarity Values Derived From the Choice Model

	Touch				Vision		
	Johnson & Phillips	Loomis (1974)	Kikuchi et al.	Craig	This Study	Gilmore et al.	Townsend
This Study, Touch	.75	.47	.56	.65	.84	.66	.49
Johnson & Phillips, Touch		.47	.62	.48	.71	.58	.38
Loomis (1974), Touch			.56	.54	.52	.52	.40
Kikuchi et al., Touch				.56	.54	.43	.39
Craig, Touch					.66	.71	.51
This Study, Vision						.69	.61
Gilmore et al., Vision							.50

Table 7  
Matrix of Correlations Between Bias Parameters Derived From the Choice Model

	Touch				Vision		
	Johnson & Phillips	Loomis (1974)	Kikuchi et al.	Craig	This Study	Gilmore et al.	Townsend
This Study, Touch	.11	.23	.06	.23	.19	.19	.58
Johnson & Phillips, Touch		.68	.27	.62	.73	.59	.26
Loomis (1974), Touch			.40	.74	.51	.67	.23
Kikuchi et al., Touch				.36	.21	.22	.16
Craig, Touch					.57	.85	.51
This Study, Vision						.63	.38
Gilmore et al., Vision							.53

Table 7. It is noted here that the correlations are generally quite low, averaging only .45. (The average does not include those correlations involving the study by Kikuchi et al., 1979, for one would not expect the letter response biases of Japanese subjects to be like those of subjects more accustomed to Roman letters.) By themselves, the low correlations between the derived response biases from studies employing uppercase Roman letters raise some question about the assumptions of the choice model, the response bias assumption being just one. More will be said about this in the Discussion.

As a means of reducing the confusion-error data in order to facilitate the task of comparing the tactile and visual results of the present study, each of the four symmetrized confusion matrices was analyzed by hierarchical clustering (Hubert & Baker, 1976; S. C. Johnson, 1967), following the lead of Gilmore et al. (1979) and Kikuchi et al. (1979). Gilmore et al. (1979) used the similarity matrix resulting from the choice model as the input to the cluster analysis, but we chose to use the original symmetrized matrices instead, since the evidence was not compelling that the similarity matrix represented an improvement. The analysis was carried out using program P1M from the BMDP statistical package (Dixon & Brown, 1977); the complete link (or maximum) method was employed. Because of the large number of cell values of zero in the off-diagonal cells of each matrix, the formation of clusters at the levels of least similarity

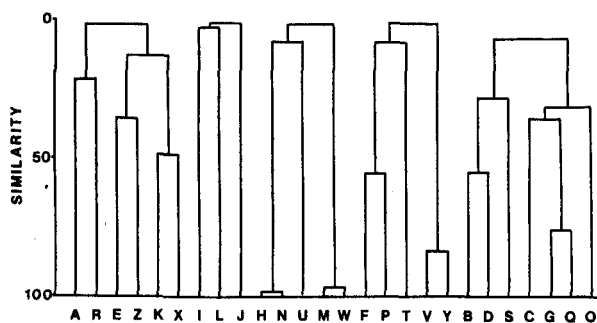


Figure 2. The results of the cluster analysis for letters sensed by touch.

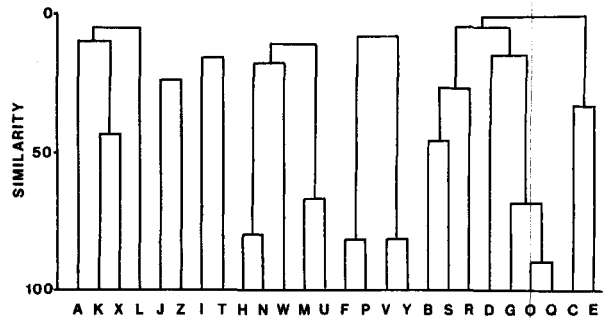


Figure 3. The results of the cluster analysis for letters sensed by vision.

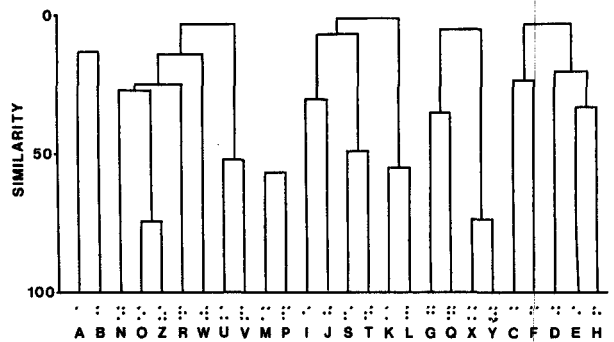


Figure 4. The results of the cluster analysis for braille characters sensed by touch.

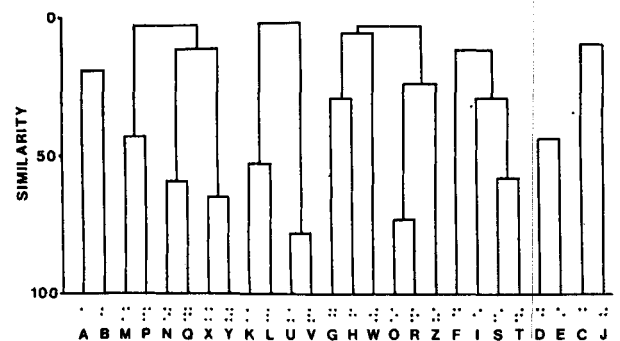


Figure 5. The results of the cluster analysis for braille characters sensed by vision.

would have been arbitrary and therefore was not done. The results of the four analyses are shown in Figures 2-5.

## DISCUSSION

The discussion is divided into two sections. The first considers the extent to which response bias actually enters into the stimulus-response contingencies of a character-recognition task. The second section deals with the central issue of this paper—the degree of similarity between visual and tactile character recognition.

### Response Bias

The notion of response bias we are considering is that a subject's tendency to use a response category varies from one category to the next and that these varying response tendencies are independent of the sensory state induced by the stimulus on any given trial. Which response the subject actually chooses, of course, depends upon both response bias and the sensory state induced, as is assumed by the choice model. Since all of the confusion matrices considered in this paper were based on more than one subject, our interest is in the response bias shared by the subjects in each experiment. Uncorrelated response biases of different subjects within each experiment would presumably show up largely as random noise within the matrix and ought not to play an important role in the interpretation of the confusion matrix.

One possible way of evaluating whether response bias is a significant factor to contend with is to consider the results of applying the choice model to each confusion matrix. The first point of interest is the finding that the correlations between the response bias vectors of different letter-recognition experiments (Table 7) average only .45. Since all the experiments on which this average is based made use of uppercase Roman letters and English-speaking subjects, one would be hard put to argue from the choice model results for a stable set of letter response biases within this population. Another could counter, however, that the response biases are specific to the particular font of uppercase characters used in each experiment.

A more serious question is raised about the response bias notion, as embodied in the choice model, when one compares the response bias vectors derived from the model with measures from the confusion matrices. As the first row of Table 8 shows, the hit rates (legibility values) and response bias parameters are highly correlated. By itself, this fact could be interpreted to mean that the hit rates primarily reflect variation in response bias and only slightly, the discriminability of the characters. However, with such strong response bias operating, one would expect it also to show up as variation in the false alarm rate,  $P(j | \bar{j})$  for each response category  $j$  where:

$$P(j | \bar{j}) = \frac{\sum_{i=1}^{26} P(i, j)}{25} \text{ for } i \neq j.$$

However, one sees in the second row of Table 8 that the false-alarm rates in general exhibit rather low correlations with the response bias parameters; only for the Townsend (1971) study is the correlation positive and significantly different from 0 ( $p < .001$ ). Thus, taking these two facts together, the high correlations between the bias parameters and hit rates and the low correlations between the bias parameters and false-alarm rate, we are forced to wonder whether the bias parameters actually measure response bias.

A most enlightening result is given in the third row of Table 8. These values are the correlations between the hit rates and the false-alarm rates. Except for Townsend's experiment, the correlations are all negative, some significantly. This means that the variation in discriminability of the different characters is not overestimated by the hit rates, as the earlier view of response bias contamination would have it, but is, in fact, underestimated. The fact that the false-alarm rates tend to be negatively correlated with the hit rates can be viewed as a consequence of the symmetry of similarity, when response bias is minimal. To see this, consider that when a character of high legibility is presented, it will most likely be identified, meaning that other response categories are unlikely to be chosen. By the symmetry of similarity, a subject is unlikely to

Table 8  
Correlations Between Different Measures for 10 Experiments

	Touch, This Study		Johnson & Phillips	Loomis (1974)	Kikuchi et al.	Craig	Vision, This Study		Gilmore et al.	Townsend	Average
	Letters	Braille					Letters	Braille			
Bias vs. HR	.71	.89	.85	.89	.72	.90	.81	.82	.93	.58	.81
Bias vs. FAR	.06	-.01	.19	.05	-.32	-.33	.04	-.21	-.08	.77	.02
HR vs. FAR	-.50	-.31	-.19	-.16	-.83	-.52	-.26	-.49	-.16	.17	-.33
d' vs. -z <sub>HR</sub>	.98	.98	.99	.98	.99	.99	.98	.98	.97	.66	.95
HR vs. β	-.13	-.17	-.24	.09	-.04	.07	-.42	-.17	-.04	-.12	.12



**Table 9**  
Estimates of  $d'$  and  $\beta$  for the Four Conditions of This Study

	Letters				Braille			
	Touch		Vision		Touch		Vision	
	$d'$	$\beta$	$d'$	$\beta$	$d'$	$\beta$	$d'$	$\beta$
A	3.73	7.65	3.45	7.31	5.37	1.22	4.11	8.06
B	1.64	8.95	1.58	8.59	3.56	11.58	2.82	11.26
C	2.34	10.73	2.64	7.21	3.70	4.39	3.03	14.12
D	2.17	7.18	2.20	9.32	2.24	11.31	2.50	10.23
E	2.35	9.89	2.13	12.79	3.28	7.60	3.63	4.99
F	2.42	13.25	2.46	6.64	2.04	8.42	2.43	10.35
G	1.60	6.83	1.99	5.15	2.69	6.66	2.35	7.21
H	1.99	7.26	2.00	7.62	2.48	13.85	2.52	10.50
I	5.19	6.03	3.98	6.66	2.92	7.93	2.97	9.42
J	4.68	12.93	4.05	6.00	2.21	7.79	2.91	13.50
K	2.08	6.29	1.92	11.28	3.44	3.37	3.15	9.25
L	4.63	11.24	4.43	1.40	3.22	42.41	2.83	7.74
M	2.00	9.82	1.90	9.95	2.90	5.07	2.85	17.37
N	1.60	5.81	1.94	7.05	1.35	5.80	1.74	9.17
O	2.61	6.00	2.12	5.10	1.79	6.34	2.40	9.41
P	2.68	25.00	2.72	7.32	1.78	11.39	2.02	9.17
Q	1.85	7.88	1.82	8.62	2.09	12.44	1.87	6.73
R	1.97	6.82	1.80	9.28	1.83	10.38	2.04	7.87
S	1.71	16.78	1.77	9.13	2.05	7.92	2.76	6.67
T	3.36	10.87	2.98	12.45	1.74	8.03	2.24	10.06
U	3.43	6.03	2.69	7.06	2.81	6.09	2.63	11.14
V	3.02	11.12	2.57	10.09	2.57	15.41	2.70	8.70
W	2.88	4.83	2.45	12.35	2.02	9.53	2.72	14.48
X	2.18	8.89	1.98	12.42	2.87	6.40	2.49	9.61
Y	2.42	10.47	2.38	13.67	1.93	12.22	2.15	13.92
Z	2.66	10.99	2.83	17.13	1.90	7.71	2.27	11.91

use the label of this character when any other is presented. For characters of low legibility, the converse argument can be given. Consequently, if response bias is minimal, one expects hit rates and false-alarm rates to vary inversely, each, accordingly, being a measure of legibility. If response bias is present as well, it reduces the symmetry of the empirical confusion matrix and lessens the negative correlation between the two measures, perhaps even making it positive.

In light of the questions raised about the choice model, how else might we partial out whatever response bias might be represented in the hit rates? A suggestion by Massaro and Schmuller (1975, p. 227) seems to provide a solution, one that has been hinted at by the "hit-rate" and "false-alarm-rate" labels

used throughout the analysis. Subject to the usual assumptions of signal detection theory,  $d'$  ought to be a measure of bias-free character legibility and  $\beta$ , a measure of bias. Accordingly,  $d'$  and  $\beta$  were computed for each character within each experiment, using the expressions  $d' = z_{FAR} - z_{HR}$  and  $\beta = f_g(z_{HR})/f_g(z_{FAR})$ , where  $z_{FAR}$  and  $z_{HR}$  are the normal standard scores corresponding to the false-alarm rate and hit rate, respectively, and  $f_g$  is the unit normal density function. (Where the false alarm rate was equal to 0 and the hit rate, equal to 1, these values were changed to .0001 and .9999.) The  $d'$  and  $\beta$  values for the braille characters and letters of the four conditions of this study are presented in Table 9. The  $d'$  values for letters/touch and letters/vision exhibit a correlation of .95 and those for braille/touch and braille/vision, a correlation of .86. The correlations between the  $d'$  values of the eight letter-recognition studies are given in Table 10. Excluding the correlations involving the Townsend (1971) study, which showed decreases, the  $d'$  correlations uniformly increased from an average of .75 to one of .85. This average increase can be attributed in part to taking false-alarm rate into account; however, the change in scale from one of probability to one of standard scores also accounts for part of the increase, the average of the correlations between the  $z_{HR}$  values being .79 for the same subset as above.

Granted that  $d'$  is a better measure of character legibility than either the hit rate or the corresponding sign-inverted standard score,<sup>3</sup>  $-z_{HR}$ , the question is *how much* better. Variance calculations for the 26 characters showed that for all of the 10 letter and braille studies listed in Table 8, except that of Townsend (1971), the variance of the  $z_{FAR}$  values never exceeded 15% of the  $z_{HR}$  variance and 9% of the  $d'$  variance, the averages being 8% and 6%. (In contrast, these proportions were 91% and 59% for the Townsend, 1971, study.) Since response bias accounts for only part of the variation in the  $z_{FAR}$  values, the effect of response bias must be minimal. Another indication of this is seen in the fourth row of Table 8, which gives, for each study, the correlation between the  $d'$  values and the corresponding  $-z_{HR}$  values for the different characters. Except for the Townsend

**Table 10**  
Matrix of Correlations Between  $d'$  Values

	Touch				Vision		
	Johnson & Phillips	Loomis (1974)	Kikuchi et al.	Craig	This Study	Gilmore et al.	Townsend
This Study, Touch	.92	.85	.88	.88	.95	.73	.11
Johnson & Phillips, Touch		.85	.89	.80	.87	.71	.03
Loomis (1974), Touch			.89	.92	.80	.79	.02
Kikuchi et al., Touch				.88	.87	.74	.18
Craig, Touch					.88	.91	.04
This Study, Vision						.77	.17
Gilmore et al., Vision							.00

(1971) study, these correlations are very high. If we are to take the  $d'$  values as fairly pure measures of legibility, then the  $-z_{HR}$  values or the hit rates themselves must also be good measures of legibility. This means that the hit rates can be only minimally contaminated by response bias. Assuming the  $\beta$  values to be measures of response bias, this latter point is corroborated by the last row of Table 8—the correlations between the hit rates and  $\beta$  values are all very low.

Having demonstrated that response bias is, at best, a very small component of the stimulus-response contingencies in most recognition tasks, we still can ask whether it manifests itself as a stable set of response tendencies across different studies. If the  $\beta$  values are purified measures of response bias, then constancy of response bias across different experiments ought to show up as high correlations between the  $\beta$  values. Such is not the case. The correlation between the  $\beta$  values of letters/touch and letters/vision (Table 9) was only  $-.02$ ; the value for the two corresponding braille conditions was  $-.18$ . The average of all 28 correlation coefficients between the eight letter-recognition studies was only  $.29$ . We conclude that response bias is neither a potent factor within nor a stable factor across recognition tasks.

#### Similarity of Vision and Touch

At the outset, it needs to be recognized that a demonstration of similarity between the tactile and visual confusion matrices of this study is only a necessary, but not sufficient, condition for claiming that tactile recognition is approximated by visual recognition of optically filtered characters. As will be argued later, tactile and visual confusion matrices might be quite similar even if the tactile and visual senses differ considerably. The main intent of this work is to assess whether the confusion matrices are similar, as required by the implication of necessity above. That tactile and visual character recognition under these conditions are indeed similar has already been established in the previous paper (Loomis, 1981). The evidence for this conclusion from that and other work will be briefly reviewed.

The essence of the paper by Loomis (1981) is that when the effective spatial bandwidth of vision is reduced by optical low-pass filtering to that of touch on the basis of independently taken measures of spatial resolution, visual character recognition closely mimics tactile character recognition. The similarity between the senses was demonstrated in several different circumstances involving the variation of character size and character shape. Although there may be situations in which distinctly different forms of degraded vision also result in similarity of the two senses, the critical point that supports the present way of thinking is that a priori we know that small tactile characters are necessarily filtered of their higher spatial frequency content because of the poor spatial

resolution of touch (relative to vision). In a recent and important paper, K. O. Johnson and Lamb (1981) have established this point very clearly. Recording from mechanoreceptive afferents in the finger of the macaque monkey, they were able to reconstruct the spatial responses of three populations of mechanoreceptors to braille-like stimuli. A sample of their results is given in Figure 6; shown are the reconstructed responses of slowly adapting (SA), rapidly adapting (RA), and pacinian (PC) afferents to stimuli of roughly the dimensions of standard braille characters. The responses of the three types of units represent gradations of low-pass spatial filtering. For the size of characters presented, even the responses of the SA fibers, which are the most spatially acute of the three, exhibit a considerable loss in fidelity. Since the physiology of mechanoreception is quite similar for man and monkey (Darian-Smith & Kenins, 1980; Johansson & Vallbo, 1979), these results are in accord with Loomis's (1981) model and, in particular, with the explanation for the superior tangibility of braille (re letters). However, this and work by K. O. Johnson and Phillips (1981; Phillips & K. O. Johnson, 1981) have also revealed some interesting properties of touch not anticipated by the model: (1) There is consider-

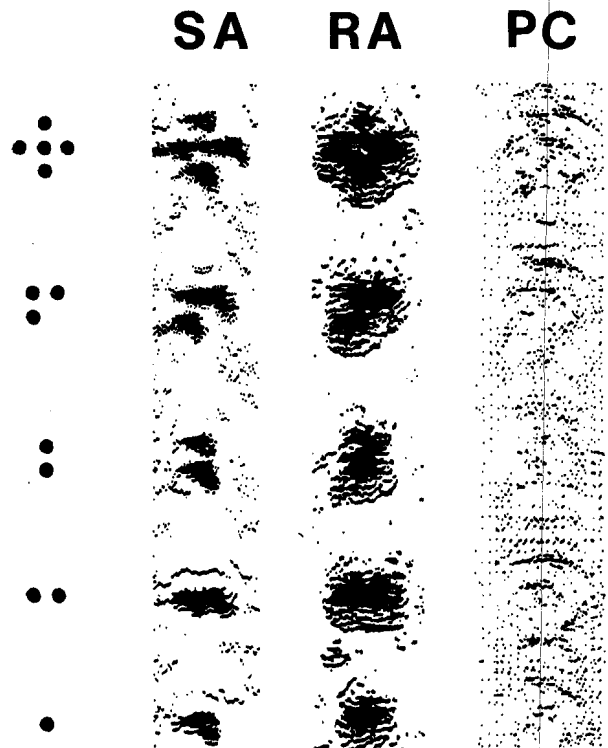


Figure 6. Reconstructed responses of slowly adapting (SA), rapidly adapting (RA), and PC (pacinian) mechanoreceptive units in the monkey finger to the braille-like stimuli shown at the left. Each stimulus dot was 1.2 mm in diameter. (This figure is part of Figure 2B from Johnson and Lamb, 1981, and was reproduced with permission of the first author.)

able anisotropy in the receptor responses reflecting the elongation of the receptive fields in the longitudinal direction of the finger, and (2) the responses to stimulus elements surrounded by neighboring elements are depressed relative to those of elements in isolation, presumably because the effective stimulation varies by virtue of the mechanical properties of the skin (Phillips & K. O. Johnson, 1981).

It ought to be mentioned that using band-limited vision as a model for touch may not require that the stimuli be optically filtered prior to the eye. Since vision is limited in spatial bandwidth by the optics of the eye, the limited sampling rate of the retinal mosaic, and, in some cases, neural organization, one would expect direct foveal viewing of appropriately sized patterns to give the same results as the method used in the present study. In the pilot work using foveal viewing prior to the published experiments (Loomis, 1981), we did not find, for whatever reason, that foveal vision led to as close a parallel with touch as did blurred vision. On the other hand, K. O. Johnson and Phillips (1981) found that the effect of letter size in a tactile recognition experiment was very much like that from an experiment using foveal vision and letters of appropriate size.

Having presented the evidence that visual recognition of low-pass filtered characters is an approximate model of tactile recognition, we can proceed to ask whether the visual and tactile confusion matrices of this study closely resemble one another, as they ought to. The facts are clear—the correlations between letters/touch and letters/vision were .95 for the  $d'$  values and .83 for the confusion error values, while the corresponding correlations between braille/touch and braille/vision were .86 and .79.

What is perhaps more interesting at this point is whether there are systematic differences between the corresponding confusion matrices. There is little to be found in the  $d'$  values, but the confusion errors do exhibit some reliable discrepancies. These are best considered by referring to the results of the hierarchical cluster analyses (Figures 2-5). With regard to letters, one striking difference is that M and W are highly confusable for touch, but not for vision. Another difference is that E is found in one cluster (E, K, X, and Z) for touch and in another (C, E) for vision. With regard to braille confusions, we also see some reliable differences. N and Q are confusable for vision, but not for touch, while O and Z are much more confusable for touch. Although these few appreciable discrepancies need to be reckoned with, there is no clear generalization one can make about them. Nor does it appear that these discrepancies can be made intelligible by modifying the band-limited model of touch to include the two additional properties of the primary cutaneous afferents found by K. O. Johnson and Lamb (1981): their anisotropy

and their reduced responsiveness in the presence of neighboring stimuli.

As required by the model we have been considering, the tactile and visual confusion matrices of this study were shown to be quite similar. However, what are we to make of the similarity between these and other matrices obtained under conditions in which there was no ostensible low-pass spatial filtering? There are at least two possibilities to entertain.

The first is that two matrices could be similar even if quite different information about the characters were utilized in the two cases; this could come about if the "shape" of a character is an invariant across a wide range of degrading transformations. For example, if the spectrum of each character of a set were divided into lower and higher halves, the resulting nonoverlapping spectra might still share the same "shape." (For some nice examples of band-pass filtered patterns, see Ginsburg, 1978.) Thus, it is possible that confusion matrices compiled for low-pass and high-pass filtered characters might be similar even though different spectral information is used by the observers in each case. This argument, extended to other sorts of transformations, might account for the similarity of two or more of the matrices discussed in this paper.

The second possibility is that most, or even all, of the confusion matrices considered here do reflect low-pass spatial filtering, including those from experiments in which the researcher did not intend it. For example, the visual confusion matrix of Gilmore et al. (1979) and the tactile confusion matrix of Craig (1979) are similar to each other and to the matrices of the present study, yet both involved very short presentations of characters that were not so small as to be limited by the sensory spatial bandwidth at long durations. Nonetheless, at short exposure durations, the characters may have been effectively filtered of their critical higher spatial frequencies. How this might be so is presented in the following reiteration of the argument from Note 6 of Loomis (1981).

In the case of vision, reducing stimulus exposure duration below the critical duration of Bloch's law amounts to a lowering of stimulus contrast (provided that there are no preceding or following pattern masking fields). Because the overall contrast sensitivity function of photopic vision (Campbell & Robson, 1968; Ginsburg, 1978) signifies that stimuli are band-pass filtered in the course of visual processing, the bandwidth of the above-threshold spectral components diminishes with decreasing contrast. If the stimuli are sufficiently small, lowering their contrast (e.g., by reducing exposure duration) means that their critical information will be effectively low-pass filtered. If, in addition, reducing exposure duration should have the effect of shifting the band-pass filtering characteristics of vision to lower frequencies

as might occur with a transition from sustained to transient channels (Breitmeyer & Ganz, 1976), the low-pass filtering would be all the more pronounced. An analogous argument might be proposed for touch, with the SA system taking the place of the sustained channels and the RA system taking the place of the transient channels.

#### REFERENCES

- BREITMEYER, B. G., & GANZ, L. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, 1976, **83**, 1-36.
- BÜRKLEN, K. *Touch reading of the blind* (F. K. Merry, trans.). New York: American Foundation for the Blind, 1932.
- CAMPBELL, F. W., & ROBSON, J. G. Application of Fourier analysis to the visibility of gratings. *Journal of Physiology* (London), 1968, **197**, 551-566.
- CARLSON, C. R., & HEYMAN, P. M. A large format optical display for the generation of generalized visual stimuli. *Vision Research*, 1979, **19**, 99-103.
- CRAIG, J. C. A confusion matrix for tactually presented letters. *Perception & Psychophysics*, 1979, **26**, 409-411.
- DARIAN-SMITH, I., & KENINS, P. Innervation density of mechanoreceptive fibers supplying glabrous skin of the monkey's index finger. *Journal of Physiology* (London), 1980, **309**, 147-155.
- DIXON, W. J., & BROWN, M. B. (Eds.). *Biomedical computer programs: P-series*. Los Angeles: University of California Press, 1977.
- GILMORE, G. C., HERSH, H., CARAMAZZA, A., & GRIFFIN, J. Multidimensional letter similarity derived from recognition errors. *Perception & Psychophysics*, 1979, **25**, 425-431.
- GINSBURG, A. P. *Visual information processing based on spatial filters constrained by biological data*. Doctoral dissertation, University of Cambridge, England, 1978.
- HUBERT, L., & BAKER, F. B. Data analysis by single-link and complete-link hierarchical clustering. *Journal of Educational Statistics*, 1976, **1**, 87-111.
- JOHANSSON, R. S., & VALLBO, A. B. Tactile sensibility in the human hand: Relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *Journal of Physiology* (London), 1979, **286**, 283-300.
- JOHNSON, K. O., & LAMB, G. D. Neural mechanisms of spatial tactile discrimination: Neural patterns evoked by Braille-like dot patterns in the monkey. *Journal of Physiology* (London), 1981, **310**, 117-144.
- JOHNSON, K. O., & PHILLIPS, J. R. Tactile spatial resolution: I. Two-point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology*, 1981, **46**, 1177-1191.
- JOHNSON, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, **32**, 241-254.
- KIKUCHI, T., YAMASHITA, Y., SAGAWA, K., & WAKE, T. An analysis of tactile letter confusions. *Perception & Psychophysics*, 1979, **26**, 295-301.
- LOOMIS, J. M. Tactile letter recognition under different modes of stimulus presentation. *Perception & Psychophysics*, 1974, **16**, 401-408.
- LOOMIS, J. M. On the tangibility of letters and braille. *Perception & Psychophysics*, 1981, **29**, 37-46.
- LUCE, R. D. Detection and recognition. In R. D. Luce, B. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1). New York: Wiley, 1963.
- MASSARO, D. W., & SCHMULLER, J. Visual features, preperceptual storage, and processing time in reading. In D. W. Massaro (Ed.), *Understanding language*. New York: Academic Press, 1975.
- MCNEMAR, Q. *Psychological statistics*. New York: Wiley, 1962.
- NOLAN, C. Y., & KEDERIS, C. J. *Perceptual factors in braille word recognition*. New York: American Foundation for the Blind, 1969.
- PHILLIPS, J. R., & JOHNSON, K. O. Tactile spatial resolution: III. A continuum mechanics model of skin predicting mechanoreceptor responses to bars, edges, and gratings. *Journal of Neurophysiology*, 1981, **46**, 1204-1225.
- TOWNSEND, J. T. Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, 1971, **9**, 40-50.

#### NOTES

1. The product-moment correlation coefficients computed for symmetrically corresponding off-diagonal values were .65 (letters/touch), .44 (braille/touch), .63 (letters/vision), and .66 (braille/vision), indicating a tendency for confusion errors to be symmetrical in direction between any two letters. These values compare with the value of .7 found by Kikuchi, Yamashita, Sagawa, and Wake (1979) for uppercase letters sensed by touch.

2. During the compilation of the confusion matrices in Tables 1 and 2, the data for the three subjects run in 1980 were kept distinct from the data of the three run in 1981. Split-half reliability coefficients (McNemar, 1962) were calculated between the two sets of data for corresponding subsets of each confusion matrix. In each case, the whole test reliability coefficient  $r_{xx}$  was calculated on the basis of the split-half reliability  $r_{12}$  using Equation 10.17 from McNemar (1962, p. 150):

$$r_{xx} = \frac{2r_{12}}{1 + r_{12}}$$

It is whole test reliability that is referred to in the text by the term "reliability."

3. High hit rates ( $p > .5$ ) are represented by z scores of negative value. To avoid the situation in which negative correlations represent positive relationships, the sign-inverted z-score,  $-z_{HR}$ , is used to represent legibility.

(Manuscript received December 1, 1980;  
revision accepted for publication October 20, 1981.)