A Model of Character Recognition and Legibility

Jack M. Loomis
University of California, Santa Barbara

This article presents a model of character recognition and the experiments used to develop and test it. The model applies to foveal viewing of blurred or unblurred characters and to tactile sensing of raised characters using the fingerpad. The primary goal of the model is to account for variations in legibility across character sets; a secondary goal is to account for variations in the cell entries of the confusion matrix for a given character set. The model consists of two distinct processing stages. The first involves transformation of each stimulus into an internal representation; this transformation consists of linear low-pass spatial filtering followed by nonlinear compression of stimulus intensity. The second stage involves both template matching of the transformed test stimulus with each of the stored internal representations of the characters within the set and response selection, which is assumed to conform to the unbiased choice model of Luce (1963). Though purely stimulus driven, the model accounts quite well for differences in the legibility of character sets differing in character type, size of character, and number of characters within the set; it is somewhat less successful in accounting for the details of each confusion matrix.

Over the years much effort has been devoted to the study of character recognition and legibility in perceptual psychology. Many of the experimental studies in vision have been motivated by applied concerns, such as determining optimal character types for print and computer displays (e.g., Berger, 1944; Maddox, Burnette, & Gutmann, 1977; S. L. Smith, 1978, 1979; Tinker, 1963) and assessing how various factors—such as achromatic and chromatic contrast, character size, and display resolution—influence character legibility (e.g., Arps, Erdmann, Neal, & Schlaepfer, 1969; Pastoor, Schwarz, & Beldie, 1983; S. L. Smith, 1979; Snyder & Taylor, 1979; Tinker, 1963; Van Nes & Bouma, 1980). Similarly, many tactile studies, especially in earlier years, have been concerned with optimizing raised print and graphics symbols (Bürklen, 1932; Gill & James, 1973; Nolan & Kederis, 1969; Nolan & Morris, 1971; Uniform Type Committee of the American Association of Workers for the Blind, 1913, 1915). In recent years there has been a shift, in both vision and touch research, toward more basic research concerned with understanding

the sensory, perceptual, and decision processes underlying character recognition and legibility. This research includes a wide range of theoretical and experimental studies (Appelman & Mayzner, 1981, 1982; Ashby & Perrin, 1988; Blommaert & Timmers, 1987; Bouma, 1971; Coffin, 1978; Craig, 1977, 1979, 1980; Dawson & Harshman, 1986; Engel, Dougherty, & Jones, 1973; Farrell, 1988; Fujita, Yamada, & Tanaka, 1984; Gervais, Harvey, & Roberts, 1984; Geyer & DeWald, 1973; Gilmore, Hersh, Caramazza, & Griffin, 1979; Ginsburg, 1986; Grossberg, 1987; Hall & Newman, 1987; Holbrook, 1975; Johnson & Phillips, 1981; Keren & Baggen, 1981; Legge, Pelli, Rubin, & Schleske, 1985; Loomis, 1981a, 1982, 1985; Lupker, 1979; Millar, 1987; Oden, 1979; Parish & Sperling, 1987; Phillips, Johnson, & Browne, 1983; Podgorny & Garner, 1979; Santee & Egeth, 1982; J. E. K. Smith, 1980; Townsend, 1971; Townsend & Ashby, 1982; Townsend, Hu, & Evans, 1984; Townsend & Landon, 1982, 1983). The most common strategy in these studies has been to collect extensive confusion data on one or two different character sets under conditions that produce imperfect recognition performance and then to attempt to account for the details of the resulting confusion matrices using a descriptive or process model. This study adopted quite a different approach to studying character legibility and the recognition process: A wide variety of character sets that differ greatly in legibility was used, and an attempt was then made to account for this variation in legibility, with a subsidiary goal being to account for the details of each confusion matrix. Much of the experimental work was published in two articles (Loomis, 1981a, 1982). This article reports additional experiments and a process model of character recognition and legibility. *Legibility* refers to the overall recognition performance (measured in percentage correct) of a given character set under some set of conditions; *recognition* refers to the process by which a given character within a set is selected in response to presentation of a stimulus from the set.

The original goal of this research was to arrive at a better understanding of tactile pattern perception and the factors that limit it.[1] The primary experimental approach was to simulate aspects of tactile sensing of raised characters by using blurred visual sensing of the corresponding optical stimuli; this approach is taken because tactile sensing can share many functional properties with vision when vision is effectively reduced to the spatial bandwidth of tactile sensing by an external low-pass spatial filter. The first use of this experimental method was reported by Apkarian-Stielau and Loomis (1975); subsequent refinements of the technique (Loomis, 1981a, 1982) have led to results that suggest a close parallel between tactile character recognition and visual character recognition when the spatial information available to vision is that available to touch. The ordering of different character sets in terms of tactile legibility is nearly identical to their ordering in terms of visual legibility; moreover, the tactile confusion matrices that have been studied have proven quite similar to the corresponding visual confusion matrices. Besides demonstrating that cutaneous spatial resolution is one of the important factors limiting tactile legibility (see also Johnson & Lamb, 1981; Johnson & Phillips, 1981; Phillips et al., 1983), the research has established that a model of character legibility that applies to either of the sensory modalities can be fruitfully applied to the other. Thus, although the original purpose of this work was to understand tactile legibility, it has broadened to encompass visual legibility as well.

Experiment 1 provides further evidence beyond that just mentioned for the near-functional equivalence of tactile and blurred visual recognition. Experiment 1 is similar to an experiment reported earlier (Loomis, 1981a, Experiment 1); whereas the latter dealt with only letters and braille, the present experiment made use of a wider variety of characters.

Experiment 1

Method

Stimuli. The seven sets of 26 characters used in this experiment are those numbered 17 to 23 in Figure 1. These characters were produced using utility programs in the DOS Toolkit running on an Apple II computer. Sets 17 and 18 were fonts provided with the software. The characters in Set 17 are the standard uppercase characters of the Apple II; those of Set 18 are a modified subset of katakana, the simplest of three character types used in written Japanese. Sets 19–23 were created using a font editor in the Toolkit. Set 19 is an arbitrary set of graphics characters. Set 20 is the alphabetic character set of standard English braille. Set 21 is the same as Set 20 except that dots in adjacent vertical or horizontal positions within the braille cell are connected by line segments. Sets 22 and 23 consist of the characters of Set 20 with added surrounds. The seven character sets were printed out and arranged as concentric circles in the artwork; this artwork was subsequently used to create a photographic negative and then, by photoengraving, a disk with raised characters. The scale of the characters in the artwork, negative, and zinc disk was such that the characters of Set 17 were 5.8 mm in height.

The touch stimuli on the zinc plate were raised 0.75 mm with respect to the surrounding etched region. By moving the disk, any desired character could be positioned beneath the finger. The visual stimuli, prior to optical low-pass spatial filtering, were the transparent characters in the photographic negative. The negative was affixed to a Plexiglas disk and mounted vertically. On one side of the disk, a tungsten source, condensing lens, and shutter provided timed and collimated illumination of a restricted portion of the disk, with rotation and lateral displacement of the disk permitting back-illumination of any desired character. Just on the other side of the disk were two parallel diffusing plates mounted vertically. With back-parallel projection of any character onto the first of these two diffusers, the desired low-pass spatial filtering could be accomplished by adjusting the separation between the two diffusers. In earlier work (Loomis, 1981a, Experiment 3), this optical filter was adjusted so that visual resolution performance matched tactile resolution performance; the setting of the filter was such that its Gaussian point spread function (measured in the plane of the second diffuser) had a full width at half amplitude equal to 5.8 mm.

Photometric measurement using a Spectra Spotmeter indicated that within the plane of the second diffuser (the output plane of the filter), the spread function corresponding to a single point input (braille character A in Set 20) had a luminous intensity of 5.0×10^{-6} candelas (cd). With the viewing distance of 81 cm, such a stimulus was well within the photopic range (approximately 1.5 log units above foveal cone threshold). Each millimeter in the plane of the second diffuser subtended 4.24 arc min; thus, the 5.8-mm-high characters of Set 17 were 24.6 arc min high in angular subtense.

Subjects. One male and 5 female subjects, all in their twenties, participated as paid observers. On the basis of an assessment made with a Keystone Orthoscope, all subjects had 20/20 vision in both eyes, except one who had 20/30 vision in the left eye. Five of the 6 had participated in an earlier experiment involving letter recognition under conditions of blurring. All were ignorant of the purpose of the experiment.

Procedure. Each subject participated in all 14 conditions obtained by crossing the seven character sets with the two sensory modalities. Counterbalancing across the seven sets was carried out to minimize sequence effects. Prior to data collection on a given character set, each subject was taught labels for the 26 characters of the set being tested; the labels used were always those of the roman alphabetic

---

[1] The goal of this research is to gain a better understanding of cutaneous (tactile) pattern perception. This is a far less ambitious goal than attempting to understand all of tactual perception, which involves varying combinations of kinesthetic and cutaneous input (Gibson, 1962; Loomis & Lederman, 1986). In pursuing this more limited goal, I acknowledge the multiplicity of known or hypothesized factors limiting tactile perception (Loomis, 1981a, 1981b, 1982; Loomis & Lederman, 1986); in addition to cutaneous spatial and temporal bandwidth limitations, there are any number of central processing limitations involving memory, attention, integration, and interpretation that are manifest in performance variations brought about by instruction, practice, and attentional manipulations (e.g., Craig, 1977, 1985; Epstein, Hughes, Schneider, & Bach-y-Rita, 1989; Hall & Newman, 1987; Loomis, 1981b; Millar, 1987). While recognizing the many factors involved in tactile pattern perception, I maintain that spatial bandwidth is the primary limiting factor in experiments involving the recognition of small raised characters sensed by the finger, provided that subjects are practiced and that they are allowed sufficient time for sensing the characters and for responding. I make no claim that these experimental results or the model have immediate implications for the perception of rapid sequences of spatial patterns (e.g., reading or tactile speech), for recognition tasks in which reaction time is the dependent measure of interest, or for any other task that departs from the conditions stated here.
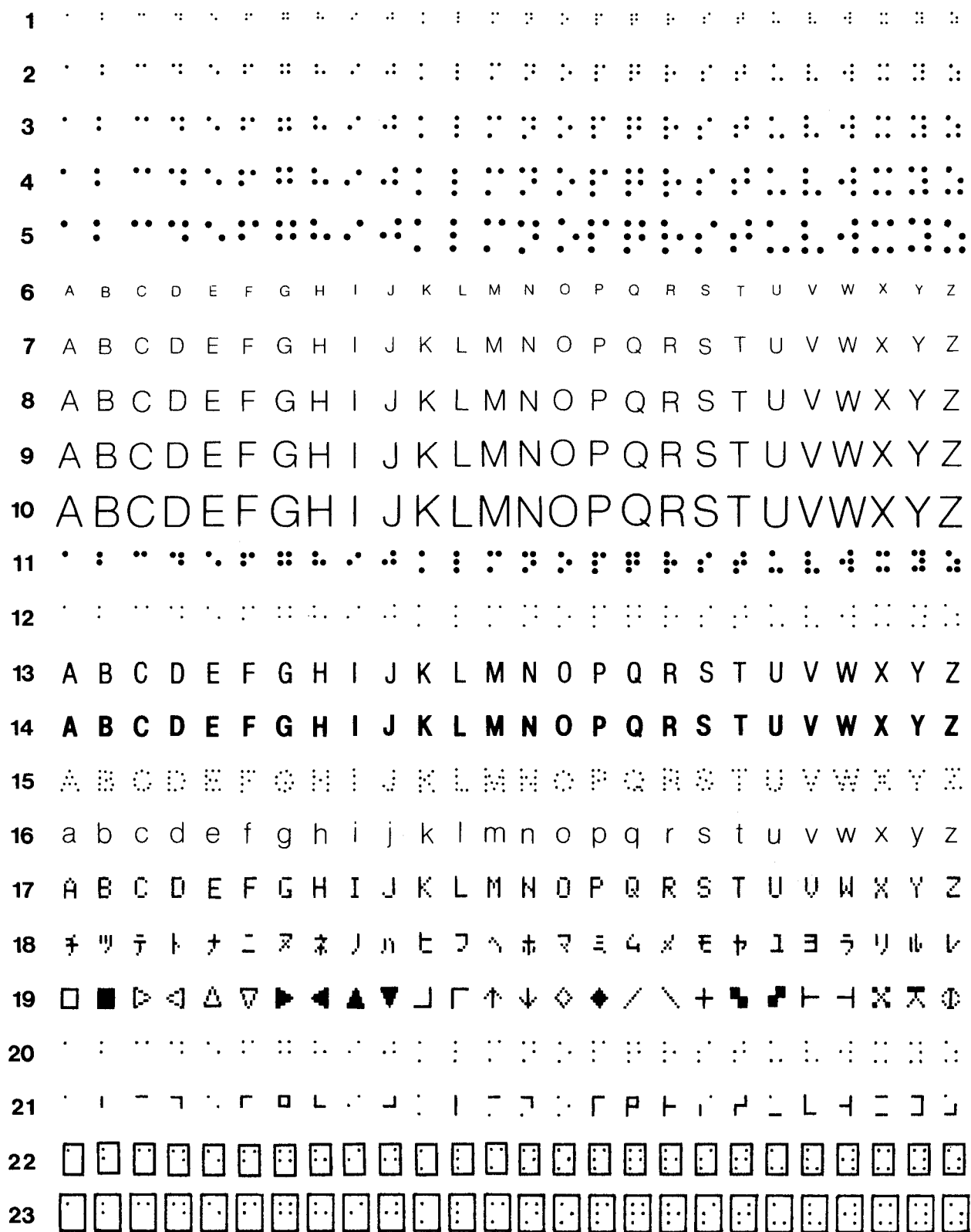
Figure 1. Character sets used in a variety of experiments. (Sets 17–23 were used in Experiment 1, and the others were used in two previous studies [Loomis, 1981a, 1982]. The scale is such that, in both the visual and tactile experiments, the characters of Set 17 were 5.8 mm high.)

characters appearing in the corresponding positions of Figure 1. Subjects were taught the labels by use of flashcards that were viewed normally at close distance. Training on a set continued until the subject achieved five errorless runs of the 26 characters. To minimize memorial interference between the labels used for the different character sets, each subject worked with only one character set per week. After training, each subject then participated in three 1-hr sessions of testing for each character set. In each session the subject received three presentations of each of the 26 characters for touch and three of each for vision, with the order of sense modality randomly determined for each session. Subjects received feedback as to the identity of the stimulus after each response.

In the tactile task, all stimuli were presented to the approximate center of the distal pad of the right index finger, which was held roughly in position by a vertically mounted finger guide. The sides of the finger guide were spaced apart far enough so that the finger could make slight circular motions over the surface of each character, with some slippage between corneum and finger permitted. Contact with the character was limited to 2 s. Subjects were instructed to make sure that the fingerpad always made even contact with the entire character and that the character remained roughly in the center of the pad. The purpose in so restricting movement was to prevent scanning of the edges of the character, a mode of touch that is more complex (Loomis, 1985). At the time that the experiment was done, measurement of the contact forces during exploration was not possible, but subsequent work (see Loomis, 1985) has shown that subjects choose to exert forces averaging about 50 g. These measurements reveal that the reported values of 5–10 g reported in two previous studies (Loomis, 1981a, 1982) were incorrect.

In the vision task, the subject viewed each stimulus with the left eye, the right eye being occluded by an eye patch. Stimulus duration was 2 s. The room was otherwise dark, and no light appeared before or after the stimulus presentation.

Results and Discussion

Recognition performance averaged over subjects is presented in Table 1 as a function of character set and sensory modality. The character set variable produced large variations in performance for both touch and vision; the mean performance values for each set were in close agreement for the two modalities, with the product–moment correlation over the seven pairs being .95. A more detailed analysis involves comparing the hit rates of the two modalities for each character set; the hit rates are simply the diagonal cell entries in the confusion matrix. In order from Set 17 to Set 23, the product–moment correlations were .88, .86, .87, .63, .71, −.09, and .40. The low values for Sets 22 and 23 are not

surprising given that performance was so close to chance. The confusion probabilities (off-diagonal cell entries in the confusion matrix) were not analyzed because each character within a given set was presented only 21 times.

These results confirm and extend the conclusion of earlier work (Loomis, 1981a, 1982) that when the effective spatial bandwidth of vision is reduced to that of touch by low-pass filtering the visual stimuli, visual character recognition bears a strong similarity to tactile character recognition. One hypothesis for this near congruence of the two senses is that the legibility of a character set is primarily determined by its criterion spatial bandwidth. This concept can be defined with reference to some low-pass spatial communication channel over which the character set is being transmitted for the purpose of recognition; an example would be a very low resolution closed-circuit television system in which the field of view is matched to the size of the character space. The criterion spatial bandwidth of a character set can be defined as the spatial bandwidth of that channel that permits recognition of that set at some criterion level of performance (Ginsburg, 1986; Legge et al., 1985; Loomis & Lederman, 1986; Parish & Sperling, 1987). Under this hypothesis, braille characters constitute a set of low-criterion bandwidth, because they require a low bandwidth communication channel. Conversely, the braille characters with small surround (Set 22) constitute a set of high-criterion bandwidth, because they require a relatively high resolution communication channel for recognition at some specified level.

Besides confirming that vision and touch have much in common in terms of legibility, Experiment 1 also shows that the effective spatial bandwidth of vision (with external filter) was well matched to that of touch, thus adding to the evidence of previous experiments, especially those involving the discrimination of resolution targets (Loomis, 1981a; Loomis & Lederman, 1986). Because visual and tactile performance levels were similar and because the visual stimuli were of the same spatial dimensions (in millimeters) as those of touch, one can conclude that the two senses were well matched in terms of effective spatial bandwidth when measured in cycles per millimeter (at the fingerpad or the output plane of the optical filter). The effective spatial bandwidth of vision might be thought to be a function of both the external optical filter and the intrinsic filtering properties of the visual system. However, a subsidiary experiment[2] (Loomis, 1983) showed

Table 1
Percentage Correct Values for Each of the Seven Character Sets Presented to Two Sensory Modalities

Character set	Modality	
	Touch	Vision
17	39.9	39.2
18	22.6	16.5
19	44.9	50.6
20	55.8	81.1
21	66.0	70.4
22	15.5	9.5
23	24.1	7.4

[2] In an experiment that is subsidiary to one dealing with character recognition (Loomis, 1983), visual spatial discrimination performance was assessed as a function of viewing distance. The visual stimuli were two-bar resolution targets (Loomis, 1981a) varying in size and orientation that were filtered by the same optical filter used in Experiment 1. The subject's task was to discriminate between targets oriented vertically and horizontally. Over a wide range of viewing distances that encompassed the viewing distance used in Experiment 1 and previous studies (Loomis, 1981a, 1982), discrimination performance varied with the size of the target in the output plane of the optical filter but was independent of threefold changes in angular size occasioned by changing viewing distance. Only for a much larger viewing distance outside this range did discrimination performance show a decline, indicating that for this distance the angular sizes of

that the external optical filter alone controls visual performance. Thus, because the optical patterns in all of these experiments have the same lateral dimensions as the corresponding raised tactile patterns, one can conclude that the point spread function of the external optical filter measured in millimeters is a fair approximation of the point spread function representing the intrinsic mechanical/neural filter of cutaneous processing for the pad of the index finger.

Before proceeding to the model, we summarize the results of three experiments that compare tactile and visual legibility by using the 23 character sets of Figure 1. Sets 17 to 23 were used in Experiment 1. Sets 1 to 10 are the roman and braille sets used to investigate the effect of varying character size (Loomis, 1981a, Experiment 3), and Sets 11 to 16 are six of the eight sets used in Experiment 1 of that same study (Loomis, 1981a). The other two sets used in that same experiment are Sets 3 and 8. Figure 2 summarizes the percentage of correct recognition data for all 23 sets, plotting visual performance of each numbered set against its tactile performance. The similarity of tactile and visual legibility is readily apparent. Figure 2 also shows clearly that the slope relating visual legibility to tactile legibility is greater than unity, a result that is at odds with the hypothesis that spatial bandwidth is the sole limiting factor for both senses. Loomis (1981a) speculated that the reason tactile legibility fails to increase at the same rate as visual legibility might be that tactile perception is subject to some additional limiting factor. One possibility is that the integration and interpretation of the available sensory information is less competently carried out in the case of touch for the simple reason that observers have much less experience with tactile pattern sensing.

The Model

The primary goal of the model presented here is to predict the legibility of an arbitrary character set presented either visually or tactually under some range of conditions; a secondary goal is to account for variations in the cell entries of the confusion matrix for a given character set. The model deals with what are thought to be the essential functions involved in human character recognition, starting with stimulus encoding and ending with response selection; the detailed physiological mechanisms underlying these functions are not addressed. Most of the ideas about the different processing stages have been considered by others in earlier work. Although this model is developed independently, it is similar in some respects to the model presented by Gervais et al. (1984).

Description

The model consists of two distinct processing stages. The first involves transformation of the stimulus into an internal

the targets were sufficiently small to cause the intrinsic filtering of visual processing to exert some influence. Thus, for the viewing distances used in the character recognition experiments, one can conclude that visual performance is controlled solely by the external optical filter. Legge et al. (1985) and Parish and Sperling (1987) followed similar lines of reasoning and reached similar conclusions.
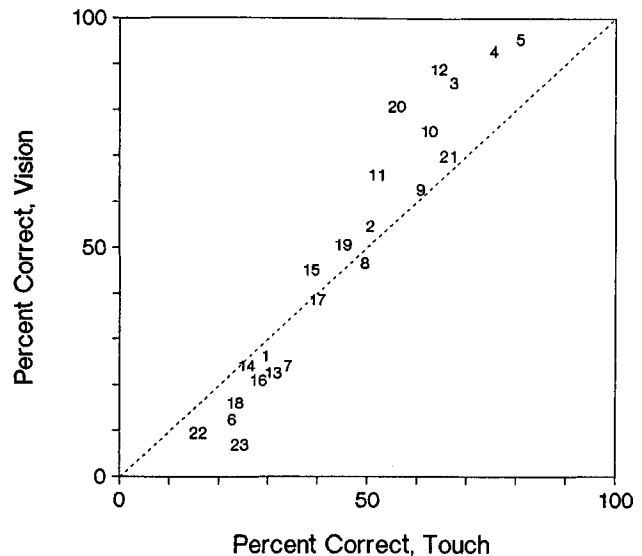

Figure 2. Summary results of three experiments. (Visual recognition performance is plotted against tactile recognition performance for each of the 23 character sets of Figure 1. The data for Sets 17–23 are from Experiment 1, and those for Sets 11–16 and Sets 1–10 are from Experiments 1 and 3, respectively, of Loomis [1981a].)

representation. The second involves the comparison of the transformed stimulus with templates of all the transformed characters in the set, and then the selection of a response. In actuality, the model does not simulate the determination of a response to a stimulus on a trial-by-trial basis. Rather, the input to the model is a set of characters and output is the theoretical confusion matrix for that set.

Each stimulus and each successive transformation of it is represented by a matrix of numbers, with each number indicating an intensity at a discrete spatial location, expressed in pixel coordinates. Thus, a stimulus, identified by the index i, is symbolized by I_{i0} (x, y) where x and y are the horizontal and vertical coordinates, respectively, and the 0 subscript refers to the unprocessed stimulus level; higher values of the second subscript will refer to the patterns after processing. Matrix entries for each stimulus were either 0 or 1, because the stimuli were all of high contrast, whereas the entries for all transformed stimuli varied continuously between 0 and 1.

$$I_{i1}(x, y) = \sum_u \sum_v I_{i0}(u, v)\exp(-\pi\theta) \qquad (1)$$

where $\theta = [(x - u)s_x c]^2 + [(y - v)s_y c]^2$.

Equation 1 represents the first transformation of the stimulus: linear low-pass spatial filtering of the stimulus. It is expressed as discrete spatial convolution of the stimulus pattern with the point spread function of the spatial filter, a discrete two-dimensional Gaussian. In the case of vision, the filtering is that produced by the external optical filter, at least for this initial analysis. The constant c with value equal to 0.163 mm^{-1}, corresponding to a full width at half amplitude of 5.8 mm, was derived from measurements of the optical point spread function of the filter in the plane of the stimulus. The horizontal and vertical scale factors, s_x and s_y, are the

pixel spacings that link the integral pixel coordinates of the stimulus with the horizontal and vertical coordinates expressed in millimeters. The filtered output for vision is the actual blurred stimulus. No further loss of spatial information due to blurring is assumed to occur in the visual transfer, because a subsidiary experiment (see Footnote 2) demonstrated no measurable contribution to filtering by the visual system under the conditions of these experiments.

In the case of touch, the spatial filtering represents the loss of high spatial frequency information caused by the overall mechanical/neural transfer from skin to cortex (Johnson & Lamb, 1981; Johnson & Phillips, 1981; Loomis, 1981a, 1982; Loomis & Lederman, 1986; Phillips & Johnson, 1981a, 1981b; Phillips et al., 1983). The best estimate to date of the overall point spread function of the cutaneous transfer for the index fingerpad of the adult human is given by the external optical filter used in the vision task, according to the argument presented in the discussion of Experiment 1.

$$I_{i2}(x, y) = [I_{i1}(x, y)]^{0.5}. \qquad (2)$$

The second step of the stimulus transformation stage, which results in the internal perceptual representation, involves a power function with an exponent of 0.5. The preceding filtering step is a linear transformation of the stimulus, but the square-root transformation is nonlinear, causing a greater fractional attenuation of the higher intensity values in the filtered output than of the lower intensities. The decision to include this nonlinear step was made because spatial masking of targets by uniform and structured backgrounds is a common finding in both target detection and target recognition paradigms, both for touch and vision (Breitmeyer & Ganz, 1976; Craig, 1974, 1982; Eriksen, 1966; Legge & Foley, 1980; Turvey, 1973). Although there may be multiple causes of masking for recognition of targets on structured backgrounds, masking of a target on a uniform background in a detection task virtually requires some sort of compressive nonlinearity (see Legge & Foley, 1980). The choice of a power function with an exponent of 0.5 was based in part on the need for a compressive nonlinearity and in part on a variety of results in the psychophysical literature (touch: Knibestol & Vallbo, 1976; LaMotte, 1977; Sherrick & Cholewiak, 1986; vision: Ginsburg, 1986; Legge & Foley, 1980). The output of this second step, I_{i2}, can be considered as a partial representation of what the subject perceives in the presence of stimulus I_{i0}. It captures the spatial information available to the subject but not those aspects incidental to the recognition process, such as color (for vision) and texture (for touch).

The second stage of the model deals with the decision process by which the subject selects a character label in response to presentation of the stimulus. Recall that in the experiments, subjects were taught labels for each of the stimuli prior to data collection and that during training they viewed each stimulus normally, so that they experienced neither external optical filtering nor cutaneous filtering. When subjects began the experiment proper, they had to match up the learned high-fidelity characters with the filtered patterns experienced during the experimental trials. In the template matching step of the model, it is assumed that the transformed stimulus is matched not with the high-fidelity templates but

with filtered templates. It is as if the subject imaginally degrades each high-fidelity template with the filtering operation in order to compare it with the filtered stimulus. This imaginal filtering of the templates is assumed to be perfect in the model. Under this assumption, the transformed stimuli and the transformed templates are equivalent in their internal representations and thus written in the same symbolic form:

$$I_{i3}(x, y) = I_{i2}(x, y) \Big/ \sum_x \sum_y I_{i2}(x, y). \qquad (3)$$

A more critical assumption of the model, expressed by Equation 3, is that the subject ignores the overall intensity of each stimulus or template, making use of only shape information. Equation 3 states that each stimulus is rescaled by its total intensity. This is a reasonable assumption for large character sets, because intensity variation is perceptually less salient than shape variation and because subjects are probably less able to store this one-dimensional aspect of the pattern in memory. However, it is an incorrect assumption for identification tasks involving very small character sets or, more so, discrimination tasks involving just two characters, because subjects would be expected to use the intensity cue when shape is highly degraded. The model, however, is intended to apply only to recognition tasks involving sets of five or more characters.

$$D(i, j) = \min_{\substack{\Delta x = -3 \text{ to } 3 \\ \Delta y = -3 \text{ to } 3}} \sqrt{d} \qquad (4)$$

where $d = \sum_x \sum_y [I_{i3}(x,y) - I_{j3}(x + \Delta x, y + \Delta y)]^2$.

Equation 4 states that the transformed stimulus I_{i3} is compared with each of the transformed templates I_{j3}, resulting in a psychological distance measure $D(i, j)$ for each comparison. The measure is the euclidean distance with the component dimensions being the intensity values associated with the various image coordinates. Because subjects in neither the visual nor tactile task had precise information about the absolute position of the stimulus, it was necessary to allow some positional variation of the template in relation to the stimulus. Hence, each template was varied in position in relation to the matrix center from 3 pixels left to 3 pixels right while also varying from 3 pixels above to 3 pixels below. $D(i, j)$ is defined as the minimum of all the euclidean distances calculated for the 49 different relative position locations of stimulus i and template j. This means, for example, that uppercase J more closely matches uppercase T when T is displaced laterally with respect to the center of J. The model could be improved by allowing for still other variations in the template, such as size rescaling, tilting, and stretching in one dimension; this was not done because of the intolerable increases in computation time that would have resulted.

Although the template-matching process results in the psychological distance between stimulus i and template j, what is needed is a measure of similarity that is reciprocally related. Such a measure—suggested by Shepard (1958, 1987) and used in a similar application by Getty, Swets, Swets, and Green (1979)—is given by

$$S(i, j) = \exp[-\tau D(i, j)]. \qquad (5)$$

One way in which to think of similarity is in terms of the activation of nodes in a network. Given stimulus i, each response node j is activated in proportion to the similarity $S(i, j)$ between the associated template and the stimulus.

Given the various activations of each of the response alternatives, how does the subject go about selecting a response? A good descriptive model of response selection is the unbiased choice model of Luce (1963; Townsend, 1971):

$$P(i, j) = S(i, j) \bigg/ \sum_k S(i, k) \qquad (6)$$

The probability, $P(i, j)$, of responding with label j given stimulus i is given by the fraction of total response activation accounted for by template j. The totality of probability values $P(i, j)$ is referred to as a theoretical confusion matrix.

As was described earlier, the model is purely stimulus driven. If response bias were known to be a significant factor in the recognition process, then a model of response selection that correctly incorporates response bias would be desirable. Unfortunately, there is as yet little consensus on how it enters into response selection and on how to measure it (Appelman & Mayzner, 1981, 1982; Keren & Baggen, 1981; Loomis, 1982; Townsend, 1971; Townsend & Ashby, 1982; Townsend, Hu, & Kadlec, 1988). Accordingly, I have not attempted to include it in the model.

Implementation

The model was first developed in BASIC on IBM microcomputers and subsequently translated into C. All critical computations have been carried out using double precision numbers. The 7×7 matrices representing the characters of Sets 17 to 23 were transferred from an Apple computer to an IBM and centered within 21×21 matrices that were otherwise filled with zeros. In the convolution step of Equation 1, the horizontal and vertical pixel spacings, s_x and s_y, were 0.713 and 0.85 mm, respectively. The filtered and compressed patterns were also represented by 21×21 matrices. The stimuli of Sets 1 to 16 were entered into 55×35 matrices by digitization. Their transformed representations, following the filtering and compression steps, were also in the 21×21 format. For a given relative placement of stimulus and template, the rescaling of each pattern by its total intensity and the computation of euclidean distance were computed only over the central 15×15 matrices of the two aligned patterns; limiting the computation to the central regions was done to allow displacement of the template up to three pixels in each of the four directions. In early stages of the work, a city block metric was also computed for distance and showed little difference from the euclidean distance.

The model has three parameters. The space constant of the point-spread function in Equation 1 was measured experimentally. (The two scale factors in the same equation are part of the stimulus specification.) The exponent of the power function in Equation 2 was selected in part because of the need to incorporate some compressive nonlinearity in the response and in part from an assessment of the psychophysical literature. The only parameter in the model obtained by

outright data fitting is the scaling constant, τ, of the exponential function in Equation 5. It was fit by maximizing the agreement between the theoretical and empirical legibility values (overall percentage of correct values) for Sets 17 to 23 in the vision task. The resulting value was used to generate the confusion matrixes for all of the remaining character sets (Sets 1–16). Precisely the same model is used for vision and touch, although as will become evident, the fit is much better for vision.

Evaluation

If a model of character recognition and legibility is to be useful, at the very least it should predict with some accuracy the relative legibility of different character sets. To assess the legibility predictions of the model, the average of the 26 hit rates of each confusion matrix was computed. Figure 3 plots the observed legibility values of the 23 character sets of Figure 1 against the predicted values from the model; the visual results are shown in the left panel and the tactile results in the right panel. Recall that the only parameter fitting was the selection of the value of the constant in Equation 5 that maximized the product–moment correlation between the empirical and theoretical values for Sets 17 to 23 in the vision task. Using that value, the product–moment correlation for all 23 sets in the vision task was .94, and the root-mean-square (rms) error between predicted and observed values was 0.111. The success with which the model predicts legibility is impressive, given that the data in Figures 2 and 3 come from three experiments in which there were different subjects and slight procedural differences. The results for touch are almost as good; the product–moment correlation and rms error were .92 and 0.166, respectively. Although the tactile data deviate more from the model predictions as indicated by the higher rms error, the high correlation indicates that with ad hoc correction thrown in, tactile legibility can be predicted nearly as well as visual legibility. The success of the model with vision and its relative failure with touch are, of course, the consequence of the decision to fit the same model to both. The lack of perfect congruence between tactile and visual legibility in Figure 2 means that no single model can account well for both. I believe that tactile performance is limited by some other factor in addition to spatial bandwidth and hypothesize this factor to be incomplete perceptual integration. Because I do not know how to model such a factor, I am content to let the model fail in the way that it does for touch.

Although the model accounts for much of the variation in legibility within a modality, certain character sets have empirical legibility values far from predicted. Only in the case of katakana, Set 18, do I have some possible explanation for the discrepancy. When all of the character sets from Sets 17 to 23 are simultaneously displayed on a video monitor and blurred an appropriate amount, one is able to visually assess the degree of shape variation within a single set with that of any other set; for the most part, this evaluation of shape variation agrees with the legibility values in Figure 2. When this was done prior to development of the model, the katakana characters seemed approximately as variable in shape as the roman characters. This was surprising at the time, because perform-
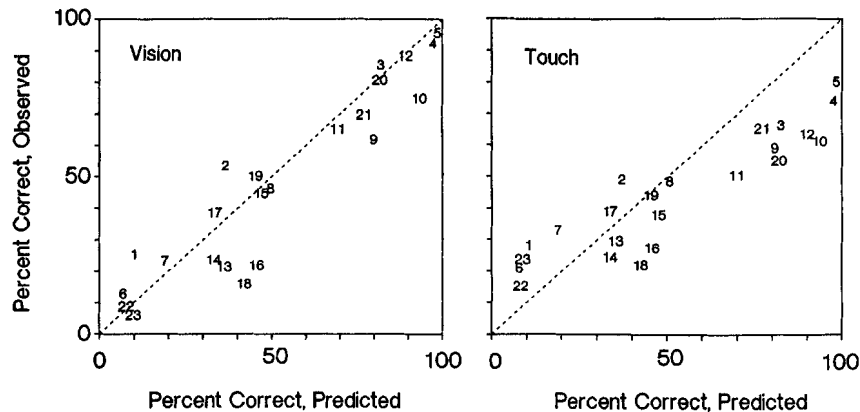
Figure 3. Visual and tactile recognition performance plotted against performance predicted by the model for each of the 23 character sets of Figure 1.

ance in the recognition task was so much poorer with the katakana characters. Subsequently, development of the model led to the prediction that katakana ought to be slightly more legible than roman. A possible explanation of why performance in the recognition task falls short of what seems possible when all blurred characters are simultaneously available for comparison involves the putative template matching step. It may be that even though subjects were able to learn the labels of the unblurred characters during training, they found it difficult to imaginally degrade the learned katakana patterns in order to match them with the filtered stimuli, as was assumed by the model. This hypothesized difficulty in performing the mental transformation could stem from the greater apparent stylistic uniformity of katakana characters. Clearly, this interpretation involves a central limiting factor that is not part of the model.

Going beyond the prediction of the relative legibility of different character sets, one might hope that a model would also be able to predict the individual cell entries of each confusion matrix. For this part of the evaluation, the data of Loomis (1982) were used; in this study, extensive recognition data were collected on Sets 2 and 8, in both the tactile and visual tasks of Experiment 1. Because subjects developed much greater proficiency than the subjects in the earlier experiment (Loomis, 1981a), performance levels were higher than those given in Figure 2. The legibility values, based on 246 presentations of each character in each condition, were 62%, 60%, 61%, and 64% for letters/touch, letters/vision, braille/touch, and braille/vision. This effect of practice was compensated for by choosing a slightly different value of τ in Equation 5; otherwise, the model was unchanged. Table 2 gives summary statistics that indicate how well the model accounts for details of the confusion matrix. Also included are measures of the congruence between the visual and tactile confusion matrices. Finally, for purposes of comparison, theoretical confusion matrices were generated by the biased choice model (Luce, 1963; Townsend, 1971): first, the empirical confusion matrices were decomposed into bias and similarity parameters using the estimation equations given by Townsend (1971), and from these, the theoretical confusion matrices were reconstructed. As before, correlation and rms

measures indicate how well this model performs. In all of these analyses, the diagonal cell entries (hit rates) are analyzed separately from the off-diagonal cell entries (confusion probabilities), because treating them together gives a misleading picture of the goodness of fit (LeBlanc & Muise, 1985; Loomis, 1982). Because the hit rates have a greater range of variation and are more reliable (Loomis, 1982), it is not surprising that they correlate more highly with the predictions of the two models. Of greater interest is the comparison between the present model and the biased choice model. The latter is more successful in terms of both rms error and correlation, but it achieves its high goodness of fit for a given character set through a large number of parameters: 351 parameters are needed to fit each 26 × 26 confusion matrix (676 cells). In contrast, the present model achieves a reasonable goodness of fit for a wide range of character sets with just three parameters, albeit with multiple processing stages. The purpose of this comparison is to show how well the present model fares on objective measures in relation to the choice model, which serves as a good standard in the analysis of confusion matrices. The comparison should not be taken as

Table 2

Product–Moment Correlations (r) and Root-Mean-Square (rms) Values for Diagonal and Off-Diagonal Cells of Confusion Matrices

Character type	Diagonal cells		Off-diagonal cells	
	r	rms	r	rms
Letters				
Touch/vision	.84	0.119	.73	0.024
Touch/process	.58	0.172	.59	0.028
Vision/process	.68	0.144	.54	0.027
Touch/choice	.98	0.047	.94	0.012
Vision/choice	.98	0.041	.93	0.011
Braille				
Touch/vision	.82	0.128	.60	0.024
Touch/process	.80	0.140	.36	0.028
Vision/process	.70	0.104	.44	0.024
Touch/choice	.98	0.050	.88	0.014
Vision/choice	.98	0.036	.91	0.010

an indication of the superiority of one model over the other, for the two are quite different in nature. Whereas the biased-choice model deals only with the response selection stage and starts with an empirical confusion matrix as input, the present model is a model of the entire process from stimulus encoding to response selection. Despite the small number of parameters, it achieves its success by having the physical stimuli as inputs and by having multiple processing stages.

To further evaluate the success of the model in predicting the details of confusion matrices, I conducted the following experiment. Extensive confusion data in a visual letter recognition task were collected for each of four levels of performance.

Experiment 2

Method

Stimuli. The characters of Set 17 used in Experiment 1 were presented in this experiment using an Apple II computer. These characters are the normal uppercase letters used by this computer. They were displayed on a Video 100 black-and-white monitor and viewed by the observer through a glass diffuser. By positioning the diffuser at different distances from the monitor, different degrees of blurring (low-pass spatial filtering) could be accomplished. The diffuser was positioned at the following distances, increasing in geometric progression: 60, 80, 106, and 142 mm. The distance between the observer's left eye and the video monitor screen was maintained at 610 mm. The characters were based on a matrix of 5 pixels wide and 7 pixels high on 0.70-mm horizontal centers and 0.76-mm vertical centers. Thus, character heights were 25.7 arc min measured between the centers of the top and bottom pixels. A single illuminated pixel was measured to have a luminous intensity of 3.5×10^{-6} cd.

The variation of the point spread function with distance of the glass diffuser was assessed in the following way. At six different distances ranging from 40 mm to 188 mm, the empirical point spread function was determined by moving a pinhole aperture vertically and horizontally within a plane just in front of the diffuser and measuring the amount of light at each position with a Spectra Spotmeter. In each case, the measured spread function was well fit by a two-dimensional Gaussian. Because the measured function depended on the distance separating the screen and the photometer lens, the measured function was then referred back to the video monitor by similar triangles (with the center of the photometer lens as the projection center). The full widths at half amplitude characterizing the different Gaussians were well fit by a linear function passing through zero at zero distance. This function was used to give best estimates of the degree of optical spread when referred to the plane of the stimuli. The corresponding full widths at half amplitude were 2.26, 3.02, 4.00, and 5.36 mm at the diffuser distances of 60, 80, 106, and 142 mm; these width values should be compared with the letters which were 4.56 mm in height (equivalent to 25.7 arc min).

Subjects. Five female and 2 male subjects, all close to 20 years old, participated as paid subjects for the 2-hr sessions each. None of the subjects was aware of the purpose of the experiment.

Procedure. Subjects were presented each letter an average of four times within each condition for each of the 10 sessions. The stimulus was presented for 1 s in an otherwise dark room; subjects viewed with the left eye. After responding with the label of one of the 26 letters, subjects were given feedback about the identity of the stimulus.

Results and Discussion

Confusion matrices were compiled for each of the four conditions pooling over all subjects and all sessions, excluding the first, which was considered practice. For each condition, the confusion matrix entries were based on an average of 252 presentations of each letter. As before, legibility for the set of characters was defined as the average of the 26 hit rates. Table 3 gives the legibility values for the empirical confusion matrices of Experiment 2 and for the theoretical confusion matrices predicted by the model. The model predictions were generated using the appropriate values of the Gaussian spread constant and the scale factors used in specifying character size. These changes in the model parameters alone ought to have predicted the observed performance levels in Table 3. However, for reasons that are not obvious, predicted performance proved somewhat better than was observed for all four conditions. The parameter of the exponential function of Equation 5 was decreased slightly (by a factor of 0.75) to obtain a better fit. The values in Table 3 indicate that with this modified parameter the model predicts the slope of the psychometric function quite well. As a further evaluation of the model, product–moment correlation and rms error values were computed in the comparison of predicted and observed confusion matrix entries; these were computed separately for diagonal and off-diagonal cells. Both measures indicate a fairly close fit of the model.

In all of the character recognition experiments reported so far, there have been 26 characters in each set. Experiment 3 was conducted to determine whether the model, without modification, can account for the results obtained with character sets having fewer than 26 characters.

Experiment 3

Method

Stimuli. The experiment was a variant of Experiment 2; the same computer setup, uppercase roman characters, viewing conditions, and procedure were used. Five different sets of characters were used: the full set of 26, two subsets of 13 (A through L and M through Z), and two subsets of 7 (A through G and M through S).

Table 3
Observed and Predicted Percentage Correct Values for the Four Conditions of Experiment 2 and Goodness-of-Fit Measures for Diagonal and Off-Diagonal Cells of the Four Confusion Matrices

| | % correct | | Goodness of fit | | | |
| | | | Diagonal | | Off diagonal | |
Condition	Observed	Predicted	r	rms	r	rms
1	89.3	93.4	.87	0.074	.83	0.013
2	73.7	75.1	.82	0.119	.67	0.024
3	44.7	45.0	.83	0.165	.52	0.033
4	23.5	19.1	.79	0.174	.34	0.034

Subjects. The 20 paid subjects were divided evenly among the four subsets. Subjects ranged from 20 to 33 years of age, with a roughly equal number of men and women. None of the subjects was aware of the purpose of the experiment.

Procedure. Each subject participated in four sessions. In the 7-character set conditions, subjects received eight presentations of each letter for each level of blur in each session. The number of presentations per stimulus was reduced to four for the 13-character set conditions and to two for the full set condition.

Results and Discussion

Confusion matrices were compiled for each of the conditions in the experiment; data from different subjects were pooled in the compilation, with the data from the first session excluded as practice. The left panel in Figure 4 depicts observed recognition performance (percentage correct) as a function of character set and stimulus bandwidth; bandwidth is expressed in terms of the filter cutoff spatial frequency (corresponding to a transmittance of 0.5) divided by the character height (see Ginsburg, 1986). The model was used to generate theoretical confusion matrices using precisely the same parameters from the analysis of Experiment 2. The right panel in Figure 4 shows the predicted performance values as a function of character set and stimulus bandwidth. As was true in Experiment 2, the model captures quite well the slope of the function relating performance to increasing stimulus bandwidth. Two other results are also important. First, the model accounts for the higher performance levels of the smaller character sets in relation to those of the full set of 26. Second, the model predicts rather well the ordering of performance between the two members of each pair of sets with the same number of characters and the same degree of blur; all 8 such order comparisons (four for the 7-character set and four for the 13-character set) were correctly predicted.

Experiment 3 shows that, without modification, the model accounts quite well for the results obtained with different character sets varying in number of characters, at least over the range studied. The model would probably fail in the case of a set of just two or three characters, for the subject would

undoubtedly make more use of intensity variation in the stimulus that is assumed to be ignored in the model. It is possible that the model could be extended to these cases by eliminating the step represented by Equation 3.

Given the success of the model in accounting for visual recognition of blurred characters, one would hope that the model could be extended to visual recognition using direct foveal viewing. At the outset of this study, the feasibility of using direct foveal viewing as a model of tactile character recognition was investigated. The pilot work preceding the work reported by Loomis (1981a) involved viewing the characters of that study foveally without any intervening optical filter. By viewing from a sufficiently far distance, visual recognition performance was lowered to that of touch. Nevertheless, that pilot work suggested that direct foveal viewing of targets of small angular size did not constitute an apt model of tactile pattern perception (Loomis, 1982, p. 51). In the meantime, Phillips et al. (1983) obtained evidence that visual character recognition is in fact quite similar to tactile character perception when the angular sizes of the visual characters bear the same ratio to the spacing between foveal cones as the sizes of the tactile characters do to the spacing of the mechanoreceptors in the fingerpad. The purpose of Experiment 4 is to examine more closely how well character recognition using foveal viewing compares with visual recognition of visual characters optically filtered prior to viewing.

Experiment 4

Method

Stimuli. The seven character sets of Experiment 1 were again used but this time were displayed on a USI Pi-3 display driven by an Apple II computer. In order to vary angular size (or retinal image size), I used five viewing distances in geometric progression: 267, 336, 423, 531, and 668 cm. Characters 5 pixels wide and 7 pixels high measured 5.2 mm high and 3.2 wide on the monitor screen. The corresponding angular heights at the five viewing distances were 6.7, 5.3, 4.2, 3.4, and 2.7 arc min. The luminous intensity of a single illuminated pixel on the monitor (e.g., braille character *A*) was meas-
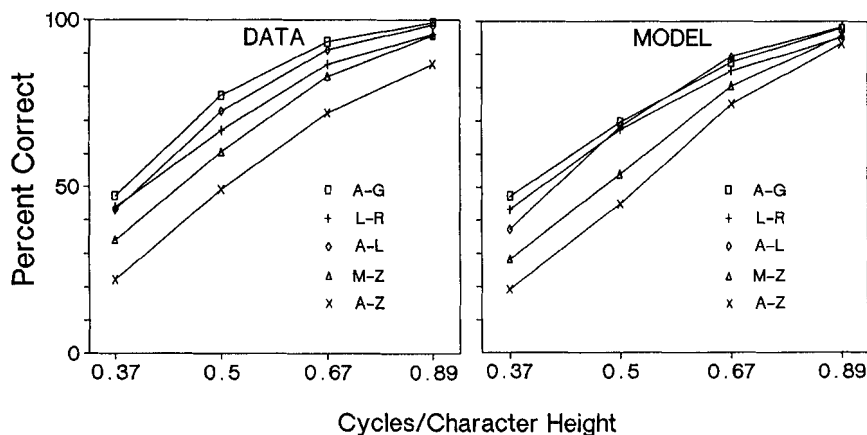

Figure 4. Observed and predicted recognition performance in Experiment 3 as a function of character set and stimulus bandwidth (cycles per character height).

ured at 2.9 × 10⁻⁶ cd using a Spectra Spotmeter. Stimuli were viewed binocularly in a completely darkened room for 2 s.

Subjects. Three female and 2 male subjects, all close to 20 years old, participated as paid observers. All had visual acuity in each eye of at least 20/20 as measured by a Keystone Orthoscope. They had no previous experience in psychophysical experiments and were unaware of the purpose of the experiment.

Procedure. As in Experiment 1, it was necessary to teach subjects the labels for the unfamiliar character sets. In the first hour of a 2-hr session dealing with a novel character set, subjects learned to associate the 26 alphabetic labels of the roman alphabet with the 26 characters in the novel set. This training hour was conducted using the computer-driven display with a viewing distance of 152 cm. Training proceeded until subjects were able to identify all 26 characters in five errorless runs. During the 1-hr experimental session, a random sample of 45 characters drawn from the character set was displayed in sequence for each of the five viewing distances. The order of character sets was counterbalanced across subjects, and the order of distances within character sets for a given subject was determined by random sampling.

During each intertrial interval, subjects were maintained in a state of light adaptation. A room light under computer control provided an illuminance of 5.3 lux in the monitor plane between trials. In addition, the entire video screen was turned on to provide a homogeneous field of 9.4 cd/m². Subjects fixated the center of the field throughout the period between stimuli. The stimulus sequence on a trial went as follows. The room lamp was extinguished, and the uniformly illuminated video screen was replaced by a fixation field. This field was dark except for four illuminated pixels centered on the location where the target character would appear. This fixation pattern remained on for 4.0 s. At its termination, the target character appeared in isolation for 2.0 s, then the trial terminated. Subjects received feedback after each response.

Results and Discussion

The left panel in Figure 5 gives recognition performance averaged over subjects as a function of character set and angular height of the 5 × 7 characters. As was expected, recognition performance increased with angular size for all character sets. The more important result is that for the medium-sized characters for which performance spanned the full range, the relative legibility of the different character sets

agreed well with their legibility in the vision task of Experiment 1. Fortuitously, the 3.4 arc min characters gave rise to performance values closely matching those of the vision task in Experiment 1. The product–moment correlation (.98) confirms the similarity of the two conditions. In agreement with the results of Phillips et al. (1983) obtained with roman letters, these data obtained with a wider range of characters indicate that visual character recognition with direct foveal viewing closely matches character recognition using the fingerpad. I speculate that the reason the pilot experiment reported earlier (Loomis, 1982) using foveal viewing provided results different from these two experiments is that the stimulus intensity was not sufficiently above detection threshold.

The results of Experiment 4 suggest that the model presented earlier can be extended to character recognition using direct foveal viewing with little modification. The expected modification would be to use the optical–neural transfer function of foveal viewing in place of the point-spread function of the external optical filter used in the blurring experiments. Although the shape of the visual transfer function of the average normal observer is known (Ginsburg, 1986; Olzak & Thomas, 1986), I opt for a different but related modification. The idea is to continue using the two-dimensional Gaussian spread function that is part of the model but to choose a parameter that is appropriate for the optical–neural transfer associated with foveal character recognition. I begin by assuming that recognition performance will be the same for two character sets with characters of unequal size when each is filtered by a point-spread function having a width that is a constant proportion of the character height; in spatial frequency terms, the cutoff spatial frequency of the filter should be a constant when expressed in cycles per character height (Ginsburg, 1986; Legge et al., 1985). The rationale behind this assumption is that recognition performance depends only on the spatial information in the stimulus actually available to the observer and not on the absolute size of the stimulus per se; in other words, it is the spatial frequency content, expressed in cycles per character height, that determines recognition performance (Ginsburg, 1986; Legge et al., 1985). In Experiment 1 the external optical filter (with a 5.8-
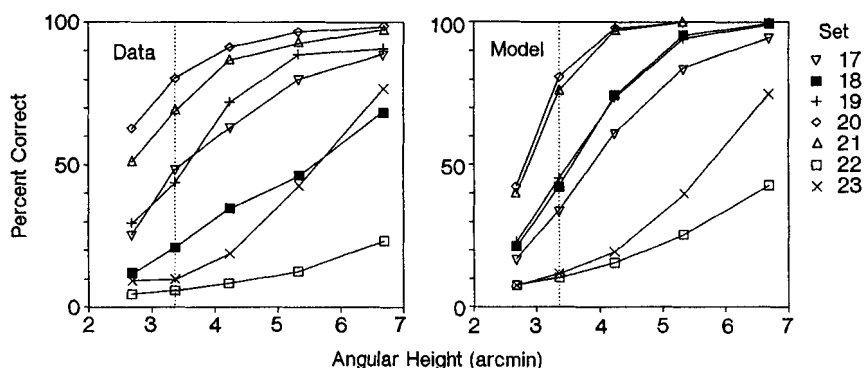

Figure 5. Observed recognition performance in a task involving direct foveal viewing, Experiment 4 (*left*), and recognition performance predicted by the model (*right*). (Five viewing distances were used resulting in five character heights, expressed in arc min. Character sets 17–23 of Figure 1 were used. The performance levels for the 3.4 arc min characters, indicated by the vertical dotted line, agree closely with those obtained with blurred viewing in Experiment 1).

mm full-width spread function) was found to limit recognition performance to 39% correct when used with characters 5.8 mm high (in the plane of the diffuser.) In Experiment 3, the intrinsic optical and neural filter of visual processing limited recognition to virtually the same performance level when characters 3.4 arc min were tested. Thus, using the constant proportion assumption from above, one would expect that the spread function representing the optical–neural transfer of foveal processing for the present conditions would be approximated by a two-dimensional Gaussian function with full width at half amplitude of 3.4 arc min. However, because the aspect (height/width) ratio of the characters in the foveal vision condition in Experiment 4 was somewhat larger than in Experiment 1, which used an external optical filter, a larger value of the full width is appropriate. In the modeling process, it was found that a full-width value for 3.95 arc min led to the closest congruence in performance for the two experiments. The corresponding modulation transfer function for this point spread function is also a two-dimensional Gaussian function (of orthogonal spatial frequency axes). This calculated low-pass filter function has modulation transfer of 1.0, 0.94, 0.78, 0.57, 0.37, and 0.21 at spatial frequencies of 0, 2, 4, 6, 8, and 10 cycles/deg. The fall-off in values above 2 cycles/deg is in fair agreement with sine wave contrast sensitivity measured psychophysically at the low photopic level used here (Olzak & Thomas, 1986). However, the filter function does not exhibit the falloff in visual sensitivity to spatial frequencies less than 2 cycles/deg (Olzak & Thomas, 1986.) In view of the similarity of results for direct foveal viewing and for lowpass filtering by an external filter, it would seem that the reduced sensitivity to spatial frequencies less than 2 cycles/deg is inconsequential for character recognition, at least for the conditions of these experiments.

The right panel in Figure 5 gives the results of the modeling analysis. For Sets 17 through 21, the model was applied directly using the filter value appropriate for the angular dimensions of the characters, as indicated earlier. The results for Sets 22 and 23 (with constant surrounds) are those obtained with a slight modification to the model; without modification the model predicted recognition performance considerably worse than that observed for the larger angular sizes of these sets. The failure of the unmodified model in dealing with the masking effects of a constant surround arises in the following way. The low-pass filtering presumed in the model means that except for characters that are very large in relation to the point spread function, the surround stimulation encroaches on the central region of the character that varies (the target). This added stimulation of the constant surround reduces the template matching differences (Equation 4) of the targets in the following two ways: (a) by the compressive nonlinearity in the stage (Equation 1) after linear filtering, and (b) by the rescaling of the filtered patterns by the overall intensity within the target region (Equation 3). In the unmodified model the area over which the template match is performed is fixed, yet it is clear that in the presence of a constant surround, the template-matching differences could be increased by simply narrowing the matching area, especially when the spacing between target and surround is larger (e.g., Set 23). Adaptive adjustment of the template-matching area

can be thought of as the subjects searching for discriminating features rather than matching globally over the entire pattern presented as stimulus. The difficulty in incorporating such adaptive adjustment into the model is that there is little experimental evidence to guide this aspect of model development and to justify the additional degrees of freedom. In the present case, the region over which the template match was performed was narrowed from the central 15 × 15 matrix of the pattern to the central 11 × 11 matrix. This ad hoc modification of the model accounts reasonably well for the results of both Sets 22 and 23.

In the comparison of performance values of model and data for all conditions, the rms value was 0.145, and the correlation coefficient was .92. Obviously, the model correctly predicts the increase in performance with character size and is reasonably successful in ordering the character sets in legibility. At the same time, however, the model fails in several significant ways beyond the need to adjust it for Sets 22 and 23. First, the slopes of the predicted psychometric functions for Sets 20 and 21 (the two versions of braille) are appreciably too large for the smaller characters, a result also apparent in Figure 3. Second, the model predicts much better performance for Set 18 over the entire range of sizes than was actually observed, a result noted earlier in the discussion of Figure 3.

General Discussion

I have presented a three-parameter stimulus-driven model of visual and tactile character recognition that is reasonably successful in accounting for the variation in legibility of character sets varying in type and number of characters; in addition, the model exhibits some success in accounting for detailed properties of individual confusion matrices, such as the variation in the diagonal cell entries. In the case of vision, the model accounts for foveal character recognition whether the characters are optically low-pass filtered prior to viewing or are filtered by the optical and neural transfer of the eye and visual pathway. In its application to tactile character recognition, this model is less successful, for there seems to be some factor limiting tactile performance in addition to the spatial low-pass filtering of the cutaneous transfer from skin to cortex. Even so, given the systematic way in which the tactile results deviate from the model predictions, one can make ad hoc adjustments of the model in order to predict the legibility of an arbitrary set of tactile characters.

There are several ways in which the model can be improved, but these will increase structural complexity of the model and entail substantial increases in computational load. The first, mentioned earlier, is to permit modest amounts of size scaling and stretching of the templates in the template-matching stage. The second, suggested in the discussion of Experiment 4, is to allow template matching within subareas of the present character space. This change amounts to a shift away from global template matching to one of adaptive feature extraction and feature matching. The third is to incorporate response bias into the model. In each case, it is far better to make the modification under the guidance of independent analysis than to simply incorporate additional free parameters to improve the fit of the model.

Even though the model has already been shown to apply quite broadly, it probably can be extended in further directions. One concerns tactile character recognition with body loci other than the fingerpad. Although there is scant evidence so far (Loomis & Lederman, 1986, pp. 10–11), it is fair to suppose that character recognition is functionally equivalent at different body loci when the stimuli are scaled in size at each site in relation to the spatial resolution at that site; the latter assumes that spatial resolution is measured under certain conditions (degree of skin indentation, frequency of vibration, etc.) comparable to those used in character presentation. If the supposition is correct, the model can be applied directly by setting the spread constant and stimulus scaling constants in Equation 1 to their appropriate values. If the subjects, however, are to explore the stimulus in ways other than the passive mode studied here, such as free scanning with the fingertip, then the model will not apply without modification, because there will be variations in spatial resolution and only parts of the stimulus will be in contact with the skin at any one moment, a circumstance that produces quite different recognition performance (Loomis & Lederman, 1986, pp. 12–15).

With respect to vision, it might be expected that the model can be extended to character recognition using nonfoveal (eccentric) viewing in analogy with the preceding supposition about character recognition with different body sites. This extension seems plausible, given evidence in the vision literature of a strong degree of invariance in spatial contrast sensitivity with variation in retinal eccentricity when the sensitivity functions are scaled by the so-called cortical magnification factor (Rovamo & Virsu, 1979; Virsu & Rovamo, 1979). However, when it comes to visual character recognition, there are systematic differences between foveal and eccentric viewing that go beyond differences in spatial resolution (Bouma, 1971; Loomis, 1978). Thus, it appears that extension of the model to eccentric vision will be only partially successful.

Two effects that a modified version of the model might explain without too much difficulty are the declines in visual legibility with decreasing contrast and with decreasing exposure duration. When a stimulus that is being passed through a linear filter is lowered in contrast, its component spatial frequencies are lowered proportionally. If one assumes that there is some thresholding of the spectral components after filtering, then the output pattern corresponding to some stimulus pattern will gradually lose its higher spatial frequency components if the filter exhibits a gradual falloff in its high-frequency transmission characteristics. If a stimulus is so small that its details are barely resolvable at high contrast, then its details will be lost with a sufficient reduction in contrast. Perhaps a modification of the model in which absolute stimulus level is taken into account might permit the model to deal with the effect of contrast. If so, then it might be possible to go a bit further and account for the decline in recognition accuracy of characters presented tachistoscopically without forward and backward maskers because, as Loomis (1982) has argued, reducing exposure duration below the critical duration of Bloch's law ought to be equivalent to a reduction in contrast (see Eriksen, 1966).

Given these directions in which the model might be extended, what are some of the obvious limits on applicability of the model? First, the model is probably inappropriate for tasks in which reaction time, rather than recognition accuracy, is the dependent measure of interest, because variation in performance often reflects limitations of central processing (Santee & Egeth, 1982) not built into the model. Second, the model obviously cannot deal with any variations in recognition accuracy that are due to practice, instruction, motivation, and attention (see Footnote 1) other than through ad hoc adjustment of the constant in Equation 5. Finally, a spatial-filtering/template-matching model, even when generalized to spatiotemporal patterns, is unlikely to account for the effects of a backward pattern masker on tachistoscopically presented stimuli in view of the conclusion that backward pattern masking involves more than an attenuation of the effective target pattern by the masker (Breitmeyer & Ganz, 1976; Craig, 1982; Turvey, 1973).

References

Apkarian-Stielau, P., & Loomis, J. M. (1975). A comparison of tactile and blurred visual form perception. *Perception & Psychophysics, 18,* 362–368.

Appelman, I. B., & Mayzner, M. S. (1981). The letter-frequency effect and the generality of familiarity effects on perception. *Perception & Psychophysics, 30,* 436–446.

Appelman, I. B., & Mayzner, M. S. (1982). Application of geometric models to letter recognition: Distance and density. *Journal of Experimental Psychology: General, 111,* 60–100.

Arps, R. B., Erdmann, R. L. Neal, A. S., & Schlaepfer, C. E. (1969). Character legibility versus resolution in image processing of printed matter. *IEEE Transactions on Man-Machine Systems, MMS-10,* 66–71.

Ashby, F. G., & Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review, 95,* 124–150.

Berger, C. (1944). Stroke-width, form, and horizontal spacing of numerals as determinants of the threshold of recognition. *Journal of Applied Psychology, 28,* 208–230.

Blommaert, F. J. J., & Timmers, H. (1987). Letter recognition at low contrast levels: Effects of letter size. *Perception, 16,* 421–432.

Bouma, H. (1971). Visual recognition of isolated lower-case letters. *Vision Research, 11,* 459–474.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppresion, and information processing. *Psychological Review, 83,* 1–36.

Bürklen, K. (1932). *Touch reading of the blind* (F. K. Merry, Trans.). New York: American Foundation for the Blind.

Coffin, S. (1978). Spatial frequency analysis of block letters does not predict experimental confusions. *Perception & Psychophysics, 23,* 69–74.

Craig, J. C. (1974). Vibrotactile difference thresholds for intensity and the effect of a masking stimulus. *Perception & Psychophysics, 15,* 123–127.

Craig, J. C. (1977). Vibrotactile pattern perception: Extraordinary observers. *Science, 196,* 450–452.

Craig, J. C. (1979). A confusion matrix for tactually presented letters. *Perception & Psychophysics, 26,* 409–411.

Craig, J. C. (1980). Modes of vibrotactile pattern generation. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 151–166.

Craig, J. C. (1982). Vibrotactile masking: A comparison of energy and patterns. *Perception & Psychophysics, 31,* 523–529.

Craig, J. C. (1985). Attending to two fingers: Two hands are better than one. *Perception & Psychophysics, 38,* 496–511.

Dawson, M. R. W., & Harshman, R. A. (1986). The multidimensional analysis of asymmetries in alphabetic confusion matrices: Evidence for global-to-local and local-to-global processing. *Perception & Psychophysics, 40,* 370–383.

Engel, G. R., Dougherty, W. G., & Jones, G. B. (1973). Correlation and letter recognition. *Canadian Journal of Psychology, 27,* 317–326.

Epstein, W., Hughes, B., Schneider, S. L., & Bach-y-Rita, P. (1989). Perceptual learning of spatiotemporal events: Evidence from an unfamiliar modality. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 28–44.

Eriksen, C. W. (1966). Temporal luminance summation effects in backward and forward masking. *Perception & Psychophysics, 1,* 87–92.

Farrell, J. E. (1988). Image quality of digital characters. *Investigative Ophthalmology and Visual Science, 29* (Suppl.), 139.

Fujita, N., Yamada, W., & Tanaka, O. (1984). [Perception of braille: A linear translation model from stimulus patterns to perceived patterns]. *Bulletin of the Faculty of Education, Kochi University* (in Japanese), Series 2, No. 36.

Gervais, M. J., Harvey, L. O., Jr., & Roberts, J. O. (1984). Identification confusions among letters of the alphabet. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 655–666.

Getty, D. J., Swets, J. A., Swets, J. B., & Green, D. M. (1979). On the prediction of confusion matrices from similarity judgments. *Perception & Psychophysics, 26,* 1–19.

Geyer, L. H., & DeWald, C. G. (1973). Feature lists and confusion matrices. *Perception & Psychophysics, 14,* 471–482.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review, 69,* 477–490.

Gill, J. M., & James, G. A. (1973). A study on the discriminability of tactual point stimuli. *American Foundation for Blind Research Bulletin, 26,* 19–34.

Gilmore, G. C., Hersh, H., Caramazza, A., & Griffin, J. (1979). Multidimensional letter similarity derived from recognition errors. *Perception & Psychophysics, 25,* 425–431.

Ginsburg, A. P. (1986). Spatial filtering and visual form perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance* (pp. 34/1–34/41). New York: Wiley.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science, 11,* 23–63.

Hall, A. D., & Newman, S. E. (1987). Braille learning: Relative importance of seven variables. *Applied Cognitive Psychology, 1,* 133–141.

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

Johnson, K. O., & Lamb, G. D. (1981). Neural mechanisms of spatial tactile discrimination: Neural patterns evoked by Braille-like dot patterns in the monkey. *Journal of Physiology* (London), *310,* 117–144.

Johnson, K. O., & Phillips, J. R. (1981). Tactile spatial resolution: 1. Two-point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology, 46,* 1177–1191.

Keren, G., & Baggen, S. (1981). Recognition models of alphanumeric characters. *Perception & Psychophysics, 29,* 234–246.

Knibestol, M., & Vallbo, A. B. (1976). Stimulus-response functions of primary afferents and psychophysical intensity estimation on mechanical skin stimulation in the human hand. In Y. Zotterman

(Ed.), *Sensory functions of the skin in primates* (pp. 201–213). New York: Pergamon Press.

LaMotte, R. H. (1977). Psychophysical and neurophysiological studies of tactile sensibility. In N. R. Hollies & R. F. Goldman (Eds.), *Clothing comfort: Interaction of thermal, ventilation. construction, and assessment factors* (pp. 83–105). Ann Arbor, MI: Ann Arbor Science.

LeBlanc, R. S., & Muise, J. G. (1985). Alphabetic confusion: A clarification. *Perception & Psychophysics, 37,* 588–591.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70,* 1458–1471.

Legge, G. E., Pelli, D. G., Rubin, G. S., & Schleske, M. M. (1985). Psychophysics of reading: 1. Normal vision. *Vision Research, 25,* 239–252.

Loomis, J. M. (1978). Lateral masking in foveal and eccentric vision. *Vision Research, 18,* 335–338.

Loomis, J. M. (1981a). On the tangibility of letters and braille. *Perception & Psychophysics, 29,* 37–46.

Loomis, J. M. (1981b). Tactile pattern perception. *Perception, 10,* 5–27.

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

Loomis, J. M. (1983, November). *Tactile and visual legibility of seven character sets.* Paper presented at the annual meeting of the Psychonomic Society, San Diego, CA.

Loomis, J. M. (1985). Tactile recognition of raised characters: A parametric study. *Bulletin of the Psychonomic Society, 23,* 18–20.

Loomis, J. M., & Lederman, S. J. (1986). Tactual perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance* (pp. 31/1–31/41). New York: Wiley.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, B. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 103–189). New York: Wiley.

Lupker, S. J. (1979). On the nature of perceptual information during letter perception. *Perception & Psychophysics, 25,* 303–312.

Maddox, M. E., Burnette, J. T., & Gutmann, J. C. (1977). Font comparisons for 5 × 7 dot matrix letters. *Human Factors, 19,* 89–93.

Millar, S. (1987). Perceptual and task factors in fluent braille. *Perception, 16,* 521–536.

Nolan, C. Y., & Kederis, C. J. (1969). *Perceptual factors in braille work recognition.* New York: American Foundation for the Blind.

Nolan, C. Y., & Morris, J. E. (1971). *Improvement of tactual symbols for blind children.* Louisville, KY: American Printing House for the Blind.

Oden, G. C. (1979). A fuzzy logical model of letter identification. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 336–352.

Olzak, L. A., & Thomas, J. P. (1986). Seeing spatial patterns. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 7/1–7/56). New York: Wiley.

Parish, D. H., & Sperling, G. (1987). Object spatial frequency, not retinal spatial frequency, determines identification efficiency. *Investigative Ophthalmology and Visual Science, 28* (Suppl), 359.

Pastoor, S., Schwarz, E., & Beldie, I. P. (1983). The relative suitability of four dot-matrix sizes for text presentation on color television screens. *Human Factors, 25,* 265–272.

Phillips, J. R., & Johnson, K. O. (1981a). Tactile spatial resolution: 2. Neural representation of bars, edges, and gratings in monkey primary afferents. *Journal of Neurophysiology, 46,* 1192–1203.

Phillips, J. R., & Johnson, K. O. (1981b). Tactile spatial resolution: 3. A continuum mechanics model of skin predicting mechanoreceptor responses of bars, edges, and gratings. *Journal of Neuro-*

physiology, 46, 1204–1225.

Phillips, J. R., Johnson, K. O., & Browne, H. M. (1983). A comparison of visual and two modes of tactual letter resolution. *Perception & Psychophysics, 34,* 243–249.

Podgorny, P., & Garner, W. R. (1979). Reaction time as a measure of inter- and intraobject visual similarity: Letters of the alphabet. *Perception & Psychophysics, 26,* 37–52.

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research, 37,* 495–510.

Santee, J. L., & Egeth, H. E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance, 8,* 489–501.

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

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237,* 1317–1323.

Sherrick, C. E., & Cholewiak, R. W. (1986). Cutaneous sensitivity. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 11/1–11/30). New York: Wiley.

Smith, J. E. K. (1980). Models of identification. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 129–158). Hillsdale, NJ: Erlbaum.

Smith, S. L. (1978). The limited readability of Lansdell numerals. *Human Factors, 20,* 57–64.

Smith, S. L. (1979). Letter size and legibility. *Human Factors, 21,* 661–670.

Snyder, H. L., & Taylor, G. B. (1979). The sensitivity of response measures of alphanumeric legibility to variations in dot matrix display parameters. *Human Factors, 21,* 457–471.

Tinker, M. A. (1963). *Legibility of print.* Ames: Iowa University Press.

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

Townsend, J. T., & Ashby, F. G. (1982). Experimental test of contemporary mathematical models of visual letter recognition. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 834–863.

Townsend, J. T., Hu, G. G., & Evans, R. J. (1984). Modeling feature perception in brief displays with evidence for positive interdependencies. *Perception & Psychophysics, 36,* 35–49.

Townsend, J. T., Hu, G. G., & Kadlec, H. (1988). Feature sensitivity, bias, and interdependencies as a function of energy and payoffs. *Perception & Psychophysics, 43,* 575–591.

Townsend, J. T., & Landon, D. E. (1982). An experimental and theoretical investigation of the constant-ratio rule and other models of visual letter confusion. *Journal of Mathematical Psychology, 25,* 119–162.

Townsend, J. T., & Landon, D. E. (1983). Mathematical models of recognition and confusion in psychology. *Mathematical Social Sciences, 4,* 25–71.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 80,* 1–52.

Uniform Type Committee of the American Association of Workers for the Blind. (1913). Fourth Biennial Report. *Outlook for the Blind, 7,* 1–48.

Uniform Type Committee of the American Association of Workers for the Blind. (1915). Fifth Biennial Report. *Outlook for the Blind, 9,* 1–92.

Van Nes, F. L., & Bouma, H. (1980). On the legibility of segmented numerals. *Human Factors, 22,* 463–474.

Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research, 37,* 475–494.