Two operations in character recognition: Some evidence from reaction-time measurements[1]

SAUL STERNBERG

BELL TELEPHONE LABORATORIES

Theories of the recognition of a visual character may be divided into three sets, defined by the way in which the stimulus is encoded before being compared to a memorized target character. A character-classification experiment was performed in which the test stimuli were characters that were either intact or degraded by a superimposed pattern. Analyses of reaction-times in the experiment lead to the rejection of two of the three sets of theories. There appear to be at least two separate operations in the recognition or classification of a character. The first encodes the visual stimulus as an abstracted representation of its physical properties. The second, which may occur more than once, compares such a stimulus representation to a memory representation, producing either a match or a mismatch. A theory of high-speed exhaustive scanning in memory underlies the experiment and is given new support. The method of reaction-time analysis that is introduced, an elaboration of the Helmholtz-Donders subtraction method, may be applicable to the general problem of the invariance of perceived form under certain transformations of the stimulus.

Consider this simple task of recognition: a person is presented with a stimulus and is asked to judge whether or not it is a particular character. At some point between stimulus and response, a representation of the stimulus encounters a memory representation of the monitored character, and the two are compared. What is the nature of the encoded stimulus? In the formation of this representation, how much analysis of the stimulus is carried out?

At least two lines of approach to the problem of character recognition lead one to expect that the stimulus is processed to a considerable extent as its representation is formed. First, in some character-recognizing machines the stimulus is normalized, or subjected to filtering operations such as "thinning" and "smoothing," before being tested (Doyle, 1960; Minsky, 1963; Selfridge & Neisser, 1963; Stevens, 1961; Unger, 1959). Other artificial recognizers carry the preprocessing further, incorporating a stage in which features are extracted from the stimulus; these features are tested in the subsequent decision process (Bomba, 1959; Fischler et al, 1962; Stevens, 1961). Indeed, it has been argued that if characters are subject to noise or distortion then such a high degree of stimulus preprocessing is necessary in a workable recognition scheme (Uhr, 1963). A second line of approach is found in neurophysiology. Recent animal studies show that the information available at the higher visual centers is not simply a mapping of the retinal image, but that abstracted features of the stimulus can be represented as well (Hubel, 1963).

Also relevant to the preprocessing of visual stimuli are recent developments in the study of "short-term" or "operating" memory. If characters were retained in memory as visual images, there would be no reason to expect a high degree of stimulus preprocessing before stimulus- and memory-representations were compared. A raw image ("direct copy") of the stimulus would be sufficient for comparison to the image of a monitored character. But Conrad (1964), Sperling (1960), and others have concluded that a visual character is often retained in memory in the form of a representation of its spoken name. The existence of such a representation would introduce the possibility that in the recognition task the stimulus is preprocessed to the point of naming and, a fortiori, identification.[2]

But whereas (a) there seems to be a theoretical need for a highly processed stimulus representation, (b) evidence of the neural machinery is becoming available, and (c) its existence is plausible in the light of our knowledge of memory, the behavioral evidence is sparse. This is not surprising. In a certain sense, the stimulus is lost until the response occurs. More specifically, the behavioral effects of most experimental manipulations can be attributed as easily to the operation of comparing stimulus- and memory-representations as to the encoding operation by which the stimulus representation is formed. What is needed are tools that allow the dissection of the behavioral effects into those attributable to the comparison operation, and those attributable to the encoding operation. Such analysis would provide information about the nature of the operations, and therefore about the stimulus representation.

Character Classification

A tool of the required kind is provided by a recently discovered phenomenon in what may be described as a *character-classification task*. This is a generalization of the simple recognition task first mentioned. On each of a sequence of trials a character is presented as a test stimulus. The subject makes a *positive response* if the character is a member of a small memorized set of characters, called the *positive set*, and makes a *negative response* otherwise. For example, the subject may be told to operate the right-hand lever if the test stimulus is a "3"
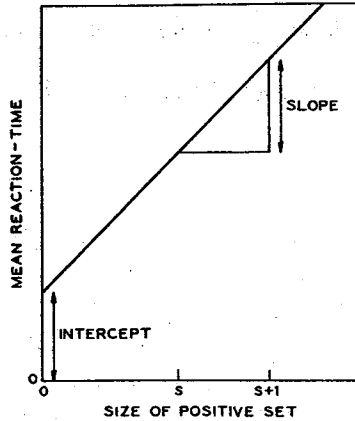
Fig. 1. Idealized data from a character-classification experiment. The slope is a measure of the mean time taken by the comparison of the stimulus representation to the memory representation of one character (Operation 2). The zero-intercept is a measure of the mean time taken by the events before and/or after the series of comparisons, which include the formation of the stimulus representation (Operation 1).

or a ·"7" and to operate the left-hand lever otherwise. In this example the positive set is of size two. The ensemble of possible test stimuli may consist of all the ten digits. The subject is encouraged to respond as rapidly as he can, while maintaining a low error-rate.

Results of several experiments on character-classification support a theory of high-speed scanning in memory (Sternberg, 1966). According to this theory, the time between stimulus and response is occupied, in part, by an exhaustive serial-comparison process. A representation of the test stimulus is compared successively to a sequence of memory representations, one for each member of the positive set, successive comparisons having the same mean duration. Each comparison results in either a match or a mismatch. After the search is completed, a positive response is initiated if there has been a match, and a negative response otherwise.

A few of the findings that support the exhaustive-scanning theory are as follows: first, mean latencies of both positive and negative responses increase linearly with the number of characters in the memorized set. This has been found for sets of up to six characters, with ensembles of digits or letters.[3] Second, the mean increase in latency per character is approximately the same for positive responses as for negative responses. This equality suggests that the search is exhaustive, rather than being terminated when a match occurs. The magnitude of the latency increase indicates an average scanning rate between 25 and 30 characters per second. Third, although the size of the positive set affects the reaction time, the size of the full ensemble does not.

Although these findings provide no direct evidence about the nature of the representations that are suc-

cessively compared, the rate of scanning, which is about four times the maximum rate of "internal speech" (Landauer, 1962), is suggestive. If the rate at which acoustic (or articulatory) memory representations can be scanned is limited by the rate of internal speech, then the use of such representations in the comparison operation is precluded.

Figure 1 is an idealization of the data obtained in a character-classification experiment. Mean reaction-time is plotted as a function of the size of the positive set. This *reaction-time function* is linear:

$$\overline{RT} = \alpha + \beta s. \qquad (1)$$

Assuming the validity of the scanning theory, one is led to the following interpretation of such data. The slope, β, of the function is a measure of the mean time taken by the comparison of the stimulus representation to the memory representation of one character (comparison time). This comparison process will be called *Operation 2*. The zero-intercept, α, is a measure of the mean time taken by events before and/or after the series of comparisons. These include the formation of the stimulus representation, which will be called *Operation 1*. Whereas Operation 1 is carried out only once, Operation 2 may occur several times, once for each character in the positive set.

Theories of Character Recognition

Theories of the character-recognition process fall into three broad classes, depending on the kind of stimulus representation postulated. Two or three examples from each class are described in Table 1. In theories of Class A, Operation 1 is inconsequential, and Operation 2 is carried out directly on a raw image of the stimulus. In theories of Classes B and C the stimulus representation is a processed version of the stimulus. In B1 and B2 it is a refined image; Operation 1 involves filtering and/or normalizing

Table 1. Some Alternative Theories of Character Recognition

Theory	Operation 1	Stimulus Representation	Memory Representation	Operation 2
A1	—	raw image	template	template-matching
A2	—	raw image	feature list	feature-testing
B1	image-refining	refined image	template	template-matching
B2	image-refining	refined image	feature list	feature-testing
B3	feature-extracting	feature list	feature list	feature-list matching
C1	identifying	imageless concept	imageless concept	concept-matching
C2	identifying and naming	image of spoken name	image of spoken name	acoustic or articulatory matching

processes. In B3 the representation is a list of relevant features of the stimulus; Operation 1 involves the extraction of features. In theories of Classes A and B, Operation 1 produces a representation based on physical properties of the stimulus that can be specified without identification of the stimulus; the memory representation then used in Operation 2 is a character prototype, of either the template or feature-list variety. According to theories of Class C, on the other hand, Operation 1 identifies the stimulus, producing a representation based on its meaning or name, which encounters a memory representation of similar kind in Operation 2. In C1 the representation produced by Operation 1 incorporates neither sensory nor motor components (see e.g., Humphrey, 1951). In C2 it is an acoustic or articulatory image of the spoken name of the character.

Two Possible Effects of Stimulus Degradation

It will now be evident how a character-classification experiment might provide a dissection tool. In the present study the experimental manipulation was the degrading of the test stimulus by superimposition of a checkerboard pattern. When the positive set contains just one member we have the simple recognition task in which the subject must decide whether the test stimulus is or is not a particular character. Suppose that his decision takes longer if the test stimulus is degraded. This fact alone does not allow us to determine whether the increase in reaction time is due to Operation 1, or Operation 2, or both. One can make this determination, however, by varying the size of the positive set and evaluating separately the effects of degradation on slope and intercept.

Let us consider how the reaction-time function (Equation 1) might change when the test stimuli are degraded. Two extreme possibilities are shown in Fig. 2. Suppose first that Operation 1 is inconsequential, and the stimulus representation is a replica of the stimulus, as in theories of Class A. The duration of Operation 1 would then be unaffected by degradation, and the increase in reaction time would have to arise
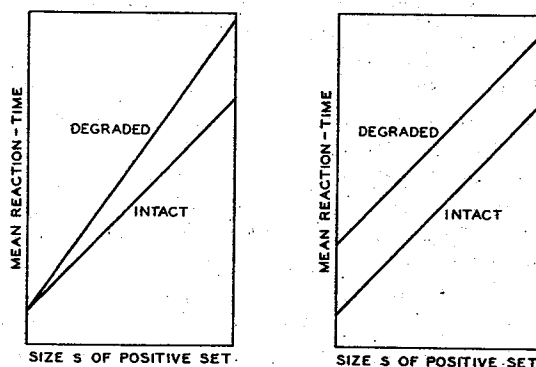

Fig. 2. Two possibilities for the effect of test-stimulus degradation on the reaction-time function. The left-hand panel, (a), corresponds to theories of Class A and Equation 2. The right-hand panel, (b), corresponds to theories of Class C and Equation 3.

from an increase, $\Delta\beta$, in the mean duration, β, of the comparison process. One would expect the slope of the function to increase, but not its zero-intercept (Fig. 2a):

$$\overline{RT} = \alpha + (\beta + \Delta\beta)s. \qquad (2)$$

At the other extreme, the stimulus representation might be a sufficiently processed version of the stimulus so that it incorporated none of the degradation. One example of such a representation is an image of the spoken name of the character, as in Theory C2. In this case there would be no reason for an increase in the duration, β, of the comparison process. The increase in reaction time would have to reflect an increase, $\Delta\alpha$, in the time to form the stimulus representation. One would expect an increase in the intercept, α, of the reaction-time function, but not in its slope (Fig. 2b):

$$\overline{RT} = (\alpha + \Delta\alpha) + \beta s. \qquad (3)$$

Analyses such as these depend, of course, on the validity of the exhaustive-scanning theory and, more generally, on the existence of additive reaction-time components and the validity of using the Helmholtz-Donders subtraction method for their measurement. On the other hand, the present study provides as byproducts several new tests of the theory, as well as additional information about the scanning process.

METHOD

Apparatus. Test stimuli were digits (about 0.6 in. high) produced by a Burroughs "Nixie" tube (Type 6844A). A beam-splitter caused a checkerboard pattern (about 7 cycles/in.) that was illuminated by an array of neon lamps (Type NE2H) to be superimposed on the digits. The degradation thus produced had previously been found to cause an increase in reaction time without substantially increasing the low error-rate. A warning signal and fixation aid was provided by an annulus that surrounded the digits and could be illuminated by a second set of neon lamps. On either side of the Nixie display was a translucent panel that could be illuminated to provide feedback signals. The subject viewed the display binocularly from a distance of about 29 in. while seated in a dimly lit booth, his head supported by a chin-rest. He rested his elbows on the table in front of him, positioning the fingers of each hand lightly on the table and immediately behind a lever which he could operate by flexing his fingers, thereby pulling it toward him. Near the beginning of its stroke the lever produced a contact-closure. The lever that was operated by the subject's dominant hand represented the positive response. The subject wore headphones through which white noise was steadily delivered at a comfortable level, in order to mask apparatus and other sounds.

Trial Events. A trial consisted of the following events: (a) intertrial interval of 2.0 sec., (b) warning signal (illumination of annulus) for 1.25 sec., (c) display of test stimulus for 44 msec., either intact or degraded, (d) subject's response (operation of one of the two levers), (e) feedback light displayed for 0.75 sec. from occurrence of response. (The panel on the side of the lever that had been operated was illuminated in green if the response was correct, in red if incorrect.) The time from onset of test stimulus to lever displacement was recorded with an accuracy of ±1 msec.

Test-stimulus Sequences. The purpose of the experiment was to determine the difference between reaction times to intact and degraded test stimuli for positive sets of size s = 1, 2, and 4. It was important to avoid confounding the variation in s with three other factors that might affect performance in this situation (Garner, 1962, Ch. 2), namely (a) response entropy, (b) stimulus entropy, and (c) systematic differences from digit to digit. This was accomplished by means of the arrangement shown in Fig. 3. For a given subject, each of the ten subscripted letters represents a particular digit, and the width of its cell represents the relative frequency with which it occurred in the population of test stimuli. The relative frequencies were x_1, 4/15; each y, 2/15; each z, 1/15; and each w, 1/15. For a given subject the stimulus population was the same in all three conditions; the important difference among conditions was in the assignment of stimuli to responses. In the row of cells associated with a condition, a hatched cell represents a digit assigned to the positive set of that condition. In each condition the test stimuli in the population were presented in a different random order.

With this arrangement, the size of the positive set could be varied without altering the relative frequency (4/15) of positive responses. Furthermore, for a given subject the relative frequency with which a particular digit occurred as a test stimulus and, *a fortiori*, the stimulus entropy and the sequential properties of the test-stimulus sequence, were the same
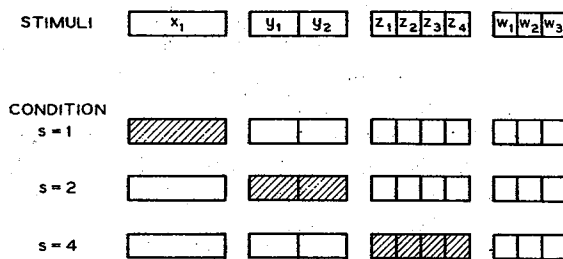

Fig. 3. Arrangement for varying size of positive set while stimulus population and response frequencies remain fixed. Letters represent test stimuli and cell widths represent their relative frequencies. Those cells in a row that are hatched represent stimuli in the positive set in that condition.

Table 2. Composition of Positive Sets of Size s = 1, 2, and 4

Subjects	Session 1			Session 2		
	s=1	s=2	s=4	s=1	s=2	s=4
1-4	5	4,9	0,1,3,7	7	3,8	1,5,6,9
5-8	8	2,7	3,5,6,9	9	0,6	2,4,5,8
9-12	7	0,9	2,3,5,6	2	1,5	0,3,4,7

from condition to condition. It should be noted that in order to exploit the invariance of the test-stimulus population so as to balance the effects of individual digits over conditions, data from the positive and negative responses in a condition had to be pooled rather than examined separately; this was done in the main analysis.

A fourth condition, with s = 3, was used for practice. For a given subject, the positive set was composed of the digits represented by w_1, w_2, and w_3. In order that the relative frequency of positive responses be 4/15, as in the other conditions, the test stimulus population had to be altered slightly.

Design. Each of twelve subjects was run for two sessions lasting about 1 hr. and separated by about a week. Each session had four parts, one for each value of s, and each part had two subparts, one with intact and one with degraded test stimuli. Part 1 was always the s = 3 condition. In the remaining parts a pair of subjects was assigned to each of the six possible orders of the s = 1, 2, and 4 conditions. One member of each pair had subparts in the order intact, degraded; the other member had the reverse order. For each subject the order of degraded and intact subparts and of conditions s = 1, 2, and 4 were reversed from Session 1 to Session 2. Trials were grouped in blocks of 18; there were three blocks per subpart except for condition s = 3, in which there were two.

In each session, three different identifications of digits with the letters of Fig. 3 were used, with two pairs of subjects assigned to each. The composition of the resulting positive sets is given in Table 2. For each subject the sets in the two sessions were "orthogonal": any two digits in the same set in Session 1 were in two different sets in Session 2.

Subjects. Subjects, students at the University of Pennsylvania, were paid for their services. The twelve subjects were selected from a group of twenty on the basis of the accuracy of their performance in Session 1. Any subject who made more than three errors in the last 36 trials in any of the last six subparts in that session was eliminated.

Payoffs and other aspects of procedure. For each block of 18 trials a subject received a score of one point per 0.01 sec. in his mean reaction-time and 10 points per error. In each session the lowest-scoring half of the subjects were each paid a $1.00 bonus. In both sessions a subject was told his mean reaction-
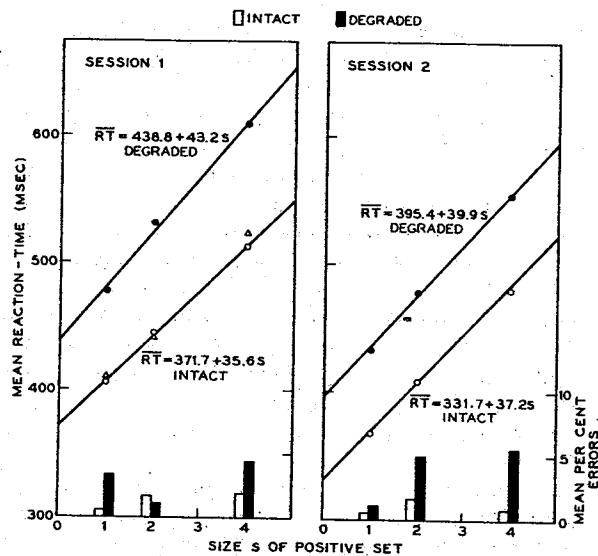
□ INTACT   ■ DEGRADED

Fig. 4. Mean reaction-time ($\overline{RT}$) and error percentage as functions of size of positive set for intact and degraded test stimuli. Left-hand and right-hand panels show data from Sessions 1 and 2, respectively. Each point (circles) represents about 29 observations from each of 12 subjects. Lines were fitted by least squares; their equations are displayed. Triangles represent data from a previous experiment.[4]

time, number of errors, and score, after each block in part 1. After this he was told only his overall errors, mean time, and score, for each part at the end of that part. He was able to rest after each block.

At the beginning of Session 1 the 10 digits were displayed serially several times, both intact and degraded. Before each part the subject was told the composition of the positive set. A right-handed subject, for example, would be told: "In the next part of the experiment, the digits for which the right-hand lever is correct are ... The other lever is correct for all other digits."

RESULTS

Excluded from the analyses were the data from part 1, the first block of each subpart, and the first three trials of each block. Also excluded were the occasional trials (2.2% in Session 1 and 2.5% in Session 2) on which the response was incorrect. There remained about 29 latencies per subpart for each subject, which included latencies of both positive and negative responses. To avoid possible effects on mean reaction-times of individual differences among digits (see section on test-stimulus sequences), the latencies of positive and negative responses were pooled and their arithmetic mean obtained. Means over subjects of the resulting values are displayed in Fig. 4. Also shown are four lines that were fitted by least squares, and their equations.

Basic Features of the Reaction-Time Function

Although the experiment was not designed as a test of the exhaustive-scanning theory, two features of the data are of particular relevance to its validity, and merit examination. The first is the linearity of the

reaction-time function. The data are well described by the linear functions shown in Fig. 4, which account for 99.6% of the variance within the four sets of means.

The second feature is the relation between the slopes of the latency functions for positive and negative responses. Lines fitted separately to the overall mean latencies of positive and of negative responses differ in slope by 5.4 ± 2.8 msec./character (SE [standard error of the mean] based on 11 df), with the negative function the steeper. This difference was not affected systematically by either test-stimulus degradation or session. The p-value derived from a t-test for slope equality is 0.08. A small slope difference in the opposite direction (-1.6 ± 3.0 msec./character) was observed in a previous experiment of similar design.[4] Taken together with the present results, that finding reduces, but does not eliminate, one's suspicion that there may be a small but systematic difference. On the other hand, the difference is small enough clearly to favor an exhaustive process (for which slopes are expected to be equal) over one that terminates when a match occurs (for which the positive slope is expected to be half the negative).

The zero-intercept of the estimated latency-function for positive responses, which were required on about 27% of the trials, exceeds the corresponding value for negative responses by 22.4 ± 9.5 msec. (SE based on 11 df); this intercept difference was not affected systematically by either test-stimulus degradation or session.

In the previous experiment,[4] which took place in a single session, test stimuli were intact, but instead of being flashed for 44 msec. they were exposed until the response occurred. Nevertheless the pooled means (triangular points in Fig. 4, least squares line being RT = 369.4 + 38.3s msec.) are almost identical to the values for intact stimuli in Session 1 of the present experiment (RT = 371.7 + 35.6s msec.).

Effects of Test-Stimulus Degradation on the Reaction-Time Function

The increases that degradation produced in the slope and zero-intercept of the reaction-time function are shown in Table 3 for the two sessions. Degradation affected the intercept markedly in both sessions. The difference between the effects in the two sessions is small, amounting to 3.4 ± 7.4 msec. (SE based on 5 df).[5] The effect of degradation on the slope is substantially smaller in absolute value than its effect on the zero-intercept and is significant in Session 1

Table 3. Effects of Test-Stimulus Degradation on Estimated Parameters of the Reaction-Time Function (SEs are based on 5 df.)

Effect of Degradation	Session 1	Session 2
Increase in mean zero-intercept (msec.)	67.1 ± 5.7	63.7 ± 10.4
Increase in mean slope (msec./character)	7.6 ± 2.4	2.7 ± 3.8

only. (In Session 1 the value of t for the slope effect is 3.22 with 5 df; the p-value for the positive tail is 0.013. The corresponding p-value for Session 2 is 0.25.) The difference between the slope effects in the two sessions is 4.9 ± 3.3 msec./character (SE based on 5 df; t is 1.49, and the p-value for one tail is 0.10).[6]

The reaction-time functions changed in two respects from session to session. First, as already indicated, the functions for intact and degraded test stimuli are more nearly parallel in Session 2 than in Session 1. Second, there was a general increase in speed, with the mean zero-intercept reduced by 41.7 ± 13.8 msec. from first to second session (SE based on 5 df). The increase in speed occurred gradually within both sessions; the overall mean reaction-time was 521.6, 490.0, and 480.1 msec. in parts 2-4 of Session 1, and 458.5, 451.6, and 450.5 msec. in parts 2-4 of Session 2. On the other hand, there appears to have been no general reduction in slope.

Comparison of Effects of Test-Stimulus Degradation on Latencies of Positive and Negative Responses

The magnitudes of the effects of degradation on latencies of positive and negative responses were determined from the combined data for conditions in which $s = 1$, 2, and 4. Combining these data produced considerable overlap of the populations of test stimuli that contributed to positive and negative means, although these populations were not identical (see Fig. 3). The results are displayed in Table 4. Evidently, the extent to which the response to a test stimulus is slowed by degradation does not depend on whether the stimulus is a member of the positive set.

DISCUSSION
Nature of the Stimulus Representation

On the basis of these findings and the exhaustive-scanning theory, what can be said about the nature of Operation 1 and the stimulus representation it produces? Although the reaction-time functions of Session 2 closely resemble the hypothetical functions shown in the right-hand panel of Fig. 2, the data from Session 1 depart significantly from both of the extreme possibilities shown in that figure. The representation is

Table 4. Comparison of Effects of Degradation on Latencies of Positive and Negative Responses

(Data from conditions with $s = 1$, 2, and 4 have been combined. SEs are based on 10 df.)

Effect of Degradation	Session 1	Session 2
Mean increase in positive-response latency (msec.)	84.8	68.6
Mean increase in negative-response latency (msec.)	84.2	70.6
Difference (msec.)	$+0.6 \pm 6.3$	-2.0 ± 8.6

apparently very far from a raw image of the stimulus, contrary to theories of Class A. It is sufficiently abstracted so that, in Session 2, the degradation of the test stimulus does not prolong the comparison operation. This conclusion follows from the virtual equality of the slopes of the two reaction-time functions in that session. The marked effect of degradation on reaction time is due, in both sessions, primarily to a change in some part of the process other than the comparison operation. It is plausible to assume that the affected part of the process occurs *before* the comparison stage and is, in fact, Operation 1. (The magnitude of the effect of degradation on the zero-intercept, about 65 msec., then provides a lower bound for the duration of Operation 1 with degraded test stimuli.)

The major effect of test-stimulus degradation then, is on Operation 1. But degradation can influence the comparison operation as well. This is shown by the 21% difference that was observed in Session 1 between the slopes of the reaction-time functions for intact and degraded stimuli. Visual degradation could influence Operation 2 only if there were residual degradation in the stimulus representation. It follows that, contrary to theories of Class C, what is represented are physical properties of the stimulus, rather than the identity or the name of the character. The observed reaction-time functions appear to be consistent only with theories of Class B.

According to these findings, then, identification need not be functionally prior to classification. This suggests (without implying) that it might not have temporal priority either, and that it could therefore be absent altogether in a classification task. Neisser (1963, 1964) has drawn this conclusion from subjects' reports of failure to "see" characters that are rejected in the course of visual search.

The existence of evidence showing that visual memory-representations are used in Operation 2 does not imply that they are also used in the retention of the positive set. Retained acoustic representations, for example, might be converted into the appropriate form when required. (By the same token, the occurrence of "acoustic confusions" in the recall of visually-presented characters does not per se imply acoustic storage, but only an acoustic representation at some stage.)

The effect of degradation on slope was less in the second session than in the first, suggesting that some sort of learning occurred. This would not be surprising, since the degradation was produced by superimposition of a fixed pattern, and the ensemble of test stimuli was the same from session to session. The improvement cannot be attributed to familiarity with particular positive sets, since their composition was changed from session to session (see section on design). Nor can it be attributed to a general increase in efficiency of the comparison operation, since the

slope of the reaction-time function for intact stimuli did not change from session to session. This invariance of slope, together with the equality of the intercept effects from session to session, also makes unlikely the possibility of the fundamental change, with practice, from a visual to a nonvisual stimulus representation. One cannot decide on the basis of these data, however, whether the improvement resulted from Operation 1 becoming more effective at eliminating degradation from the representation, or from Operation 2 becoming less sensitive to the degradation that remained.[7]

Alternative Mechanisms for the Slope Effect

Among the assumptions about the scanning process that were made in the previous section is that the number of comparisons on a trial equals the size, s, of the positive set. Given this assumption, any increase that stimulus degradation produces in the slope of the reaction-time function must reflect a corresponding increase in the mean comparison-time. An alternative is that there is an increase in the mean *number* of comparisons. This could result from at least two mechanisms that merit consideration.

Multiple memory-representations. The first is the generation of multiple memory-representations (variants) for some or all of the characters in the positive set. Such a technique is used in some artificial recognizers (e.g., Mermelstein & Eden, 1964). It would provide a means of counteracting potential defects in a visual stimulus-representation produced by Operation 1. For example, if degradation might alter or obliterate certain stimulus features then, for some of the characters in the positive set, Operation 2 would involve the comparison of the stimulus-representation to more than one feature list, thus increasing the effective size of the set. The number of variants per character, and hence the number of comparisons per trial, would be increased when the subject expected the test-stimulus to be degraded. This would lead to an increase in the mean comparison-time *per character*, and hence in the slope, with no change in the mean comparison-time *per variant*. Unless we make additional assumptions, the results of the present experiment do not permit us to discriminate between an increase in the number of memory representations and an increase in the mean comparison-time per representation.

Multiple stimulus-representations. The second possible mechanism in which degradation would influence the number of comparisons made rather than the comparison time, is the generation of multiple stimulus-representations, which might be proposed in an attempt to salvage theories of Class C. A degraded test stimulus might occasionally be ambiguous, producing representations of more than one character; a serial-comparison process would then be carried out for each. Suppose that two representations, rather

than one, were generated by Operation 1 on a proportion, p, of the trials. Then the mean number of comparisons would be (1 + p)s rather than s:

$$\overline{RT} = (a + \Delta a) + \beta (1 + p)s. \quad (4)$$

The result would be a slope increase of 100p% with no change in the mean time per comparison. This explanation would reconcile the observed slope effect with theories of Class C, thereby weakening our conclusions. Multiple stimulus-representations would also give rise to false-positive responses. The proposed mechanism must be rejected because it leads to several predictions about such errors that are inconsistent with the data, given the magnitude of the observed slope-effect. (For example, even if stimulus ambiguity were the only source of false positives, the mean proportion of positive responses on negative trials would be ps/9, or 0.095 for s = 4 in Session 1; the observed value is only 0.045 for the 20 subjects run in that session, and 0.023 for the subset of subjects who met the error-rate criterion.)

Implications for the Exhaustive-Scanning Theory and Some Competitors

It has been assumed that an increase in response latency caused by stimulus degradation might arise from one or both of two possible sources: (a) an effect on Operation 1. This operation precedes the classification decision; any change (Δa) in its duration should therefore affect positive and negative responses equally. (b) An effect on Operation 2. Since scanning is exhaustive, this operation occurs the same number of times (s) whichever decision is required; any change ($\Delta \beta$) in its duration should likewise affect positive and negative responses equally. The observed equality of the increases in positive- and negative-response latencies therefore supports the above assumption.

On the other hand, this equality of effects is hard to reconcile with a theory in which the amount of processing of the stimulus depends on whether or not it is a member of the positive set. Such a theory, involving the operation of a hierarchy of feature recognizers, has been put forward by Neisser (1963, 1964) for monitoring in visual search. He proposes that a stimulus is tested first for gross features, is processed further only if it passes the first test, and so on. More tests are therefore carried out for targets than for nontargets. If degradation increased the duration of such tests, its effect on response latency would increase with the number of tests. If this kind of theory were valid for the present experiment one might therefore expect stimulus degradation to have a greater effect on positive- than on negative-response latencies. (An alternative possibility is that by obliterating some features, degradation would increase the number of tests required

before a nontarget was rejected, thereby increasing the latency of negative responses. A positive response would entail the testing of all pertinent features, whether the stimulus was degraded or not; its latency would therefore be unaffected. Again, the effects of degradation would be unequal, but in the opposite direction.)

The exhaustive-scanning theory describes a process of comparison of stimulus- and memory-representations. Many of the previously available findings (Sternberg, 1966) were consistent with an alternative theory in which scanning consists, instead, of the search for a "marker" associated with one of the memory representations of the positive set. Each possible test stimulus is represented in memory before the trial begins. When a test stimulus is presented it is identified and the corresponding memory representation is "marked." The duration of this operation is independent of the size of the positive set. The representations of members of the positive set are then serially and exhaustively scanned for the presence of a marker. This theory may now be rejected because, like theories of Class C, it cannot be reconciled with an effect of stimulus degradation on the scanning rate.

The effects on reaction time of experimental variations have occasionally been explained in terms of changes in the amount of time occupied by the sampling of information from the stimulus (e.g., Fitts et al, 1963; Stone, 1960). That this kind of explanation may be of limited usefulness for simple-reaction time has been shown, for example, by Raab and Fehrer (1962), who found that differences in the luminance of a 2-msec. flash could produce differences of 120 msec. in reaction time. The present experiment provides similar evidence for choice-reaction time. First, the effects on reaction time of both degradation (65 msec.) and set size (as much as 130 msec.) were greater than the duration of the test stimulus (44 msec.). Second, the reaction-time function for intact stimuli appears to be unaffected by whether the test-stimulus exposure is brief or long (response-terminated).

Further Application of the Method

It may be possible to apply the experimental technique described here to the general problem of explaining "form constancy," the invariance of object identification with respect to certain transformations of the stimulus (Attneave, 1962). There are many ways of transforming a stimulus without necessarily altering a classification or identification response. The addition of a checkerboard is one example; others are rotation, reflection, magnification, blurring, and distortion. But whereas a response may be the same despite a stimulus transformation, its latency may not be. In such cases, investigating the locus of changes in latency may reveal something of the mechanism that underlies the response invariance.

References

Attneave, F. Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science.* New York: McGraw Hill, Vol. 4, 1962. Pp. 619-659.

Bomba, J. S. Alpha-numeric character recognition using local operations. *Proc. east. joint Comput. Conf.,* 1959, 16, 218-224.

Conrad, R. Acoustic confusions in immediate memory. *Brit. J. Psychol.,* 1964, 55, 75-84.

Doyle, W. Recognition of sloppy, hand-printed characters. *Proc. west. joint Comput. Conf.,* 1960, 17, 133-142.

Fischler, M., Mattson, R. L., Firschein, O., & Healy, L. D. An approach to general pattern recognition. *IRE Trans. inform. Theory,* 1962, 8, 64-73.

Fitts, P. M., Peterson, J. R., & Wolpe, G. Cognitive aspects of information processing: II Adjustments to stimulus redundancy. *J. exp. Psychol.,* 1963, 65, 423-432.

Garner, W. R. *Uncertainty and structure as psychological concepts.* New York: Wiley, 1962.

Hubel, D. H. Integrative processes in central visual pathways of the cat. *J. Opt. Soc. Amer.,* 1963, 53, 58-66.

Humphrey, G. *Thinking.* London: Methuen, 1951.

Landauer, T. K. Rate of implicit speech. *Percept. mot. Skills,* 1962, 15, 646.

Mermelstein, P., & Eden, M. Experiments on computer recognition of connected handwritten words. *Inf. Cont.,* 1964, 7, 255-270.

Minsky, M. Steps toward artificial intelligence. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought.* New York: McGraw Hill, 1963. Pp. 406-450.

Neisser, U. Decision-time without reaction-time: Experiments in visual scanning. *Amer. J. Psychol.,* 1963, 76, 376-385.

Neisser, U. Visual search. *Scient. American,* 1964, 210, 94-102.

Raab, D., & Fehrer, E. Supplementary report: The effect of stimulus duration and luminance on visual reaction time. *J. exp. Psychol.,* 1962, 64, 326-327.

Selfridge, O. G., & Neisser, U. Pattern recognition by machine. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought.* New York: McGraw Hill, 1963. Pp. 237-250.

Sperling, G. The information available in brief visual presentations. *Psychol. Monogr.,* 1960, 75, No. 11 (Whole No. 498).

Sternberg, S. High-speed scanning in human memory. *Science,* 1966, 153, 652-654.

Stevens, Mary E. *Automatic character recognition: A state-of-the-art report.* NBS Tech. Note 112. Washington, D. C.: U. S. Department of Commerce, 1961.

Stone, M. Models for choice-reaction time. *Psychometrika,* 1960, 25, 251-260.

Uhr, L. "Pattern recognition" computers as models for form perception. *Psychol. Bull.,* 1963, 60, 40-73.

Unger, S. H. Pattern detection and recognition. *Proc. IRE,* 1959, 47, 1737-1752.

digits defining the positive set is presented on each trial, and (a) the time interval between series and test stimulus is too short to allow the subject to find the complement and/or (b) the subject is required to recall the series after his binary response.

4. Experiment 2 in Sternberg, 1966.

5. The experimental unit in these analyses was the pair of subjects having the same order of conditions s = 1, 2, and 4 but opposite orders of intact and degraded subparts. The data for each pair were averaged and then corrected for the mean part-of-session effect. Lines were fitted by least squares to the corrected values. For each of the six pairs of subjects, and for each session, the increases produced by degradation in slope and zero-intercept were determined. SEs were based on these values for each of the two sessions, or on the appropriate differences between them.

6. For the six experimental units (pairs of subjects) the differences between estimated slopes for degraded and intact test stimuli in Session 1 were 14.7, 12.0, 9.8, 6.5, 3.9, and -1.2 msec./character. In Session 2 these differences became 16.2, 1.3, -2.0, -5.0, 11.8, and -6.3 msec./character respectively. The estimates of slope- and intercept-effects in Session 1 appear not to have been biased by the rejection of subjects who exceeded the error-rate criterion in that session (see section on design). For the rejected subjects both of the mean effects fell within one SE of the values for accepted subjects.

7. This argument is one of several in the present paper that depend on the assumption that any effect of degradation on an operation will be revealed by a change in the duration of that operation.