Journal of Experimental Psychology:
Human Perception and Performance

VOL. 4, NO. 1 FEBRUARY 1978

Size Scaling in Visual Pattern Recognition

Axel Larsen and Claus Bundesen
Copenhagen University, Denmark

Human visual recognition on the basis of shape but regardless of size was investigated by reaction time methods. For successive matching of random figures, reaction time increased linearly with the linear size ratio of stimulus pairs. For single-character classification, reaction time increased with divergence between cued size format and stimulus format such that for character nonrepetitions, the increment in latency was approximately proportional to the logarithm of the linear size ratio of the two formats. However, when reactions to character repetitions were faster than those to nonrepetitions, the repetition reaction time function was similar to that for successive matching of random figures. The results suggested two processes of size scaling: mental-image transformation and perceptual-scale transformation. Image transformation accounted for matching performance based on visual short-term memory, whereas scale transformation accounted for size invariance in recognition based on comparison against visual representations in long-term memory.

Our visual capacity to classify objects on the basis of shape but regardless of size constitutes a fundamental problem of visual perception. The capacity is expressed when two objects of different sizes are perceived as identically shaped, or equally, when a single object of a specific shape is perceived as a member of a given category regardless of the specific size of the object. The theoretical possibilities in accounting for these facts of size invariance in visual recognition are critically dependent on the basic assumptions concerning the pattern recognition process.

Visual pattern recognition is presumably achieved by comparing stimulus patterns with memory representations. In one type of interpretations, the mode of comparison is essentially *position-wise:* The memory representation specifies a spatial arrangement of pattern elements (points or subpatterns), and the comparison is made with respect to particular positions in the field of view. For example, in template theory (see Neisser, 1967), recently revived in the context of Fourier analysis (e.g., Pribram, Nuwer, & Baron, 1974), the memory representation specifies a canonical spatial distribution of points, and recognition is attempted by a process of position-wise com-

---

This research was supported in part by a grant from the Danish Research Council for the Humanities. The authors are indebted to Jørgen Rathje for important technical assistance and to Sven Kreiner Møller for valuable suggestions concerning the statistical analysis.

Requests for reprints should be sent to Claus Bundesen, Psychological Laboratory, Copenhagen University, Njalsgade 90, DK-2300 Copenhagen S., Denmark.

parison through point-by-point correlation. Similarly, in some structural schema theories (e.g., Hochberg, 1970; Noton & Stark, 1971), a schema defines a canonical spatial arrangement of subpatterns by means of the attention shifts required to pass from one subpattern to another; in evaluating a stimulus against a schema, then, a position-wise comparison is performed.

If visual pattern recognition is based on position-wise comparison of stimulus patterns against memory representations, the problem of size invariance may be approached as follows: A set of long-term memory representations, each of which specifies a canonical spatial arrangement of pattern elements in relation to a standard reference system, is postulated. To use these memory representations for recognition, a correspondence must be assumed between positions in the memory reference system and positions in the current field of view; that is, the standard reference system must be assigned an interpretation in the field of view. A very simple assumption is that the memory reference system has a fixed interpretation in terms of retinal coordinates. Alternatively, the correspondence between positions in the memory reference system and positions in a given field of view could be variable. The latter assumption can be stated by postulating the positional correspondence to be established by imposition of a variable perceptual reference system on the visual field such that this perceptual reference system constitutes the effective interpretation of the memory standard reference system.

Consider a standard stimulus forming a spatial arrangement of pattern elements in relation to a given perceptual reference system such that this arrangement conforms directly to the specification of a given long-term memory representation. If a size transform is substituted for the standard pattern, the new stimulus will not conform directly to the memory specification under the given interpretation of the memory standard reference system. However, size-invariant recognition may be achieved by two types of simple processes: image transformations and scale transformations.

Processes of *image transformation* may serve for size-invariant recognition with a fixed perceptual reference system. Three possibilities may be considered in relation to the particular example above. First, suppose that the process of comparing a long-term memory representation against a stimulus is mediated by a comparison between that memory representation and a transformable visual image of the stimulus. In this case, size invariance may be obtained by normalizing the new stimulus to fit the given perceptual reference system (cf. Minsky, 1961), that is, by transforming the visual image of this stimulus so that the represented size changes to that of the standard stimulus. Second, assume that the process of comparing a long-term memory representation against a stimulus may be mediated by matching the stimulus against a transformable (and possibly schematic) visual image which is generated from the long-term representation (cf. Posner, Boies, Eichelman, & Taylor, 1969). If so, size invariance may be achieved by generating an image that represents the standard pattern and transforming this image so that the represented size changes to that of the new stimulus. Third, if a visual image of the standard persists from the first stimulus presentation, this image may be transformed and used for position-wise comparison against the new stimulus, thus bypassing the long-term representation.

If the perceptual reference system is allowed to vary, a process of *scale transformation* may serve for size-invariant recognition. In the above example, size-invariant recognition can be obtained by transforming the perceptual scale (i.e., the unit of the perceptual reference system) in proportion to the changing stimulus size. Following the appropriate scale transformation, the new stimulus pattern conforms position-wise to the specification of the given long-term memory representation.

In general, if recognition is based on position-wise comparison of stimulus patterns with memory representations, size-invariant recognition may be explained by means of image transformations, scale transformations, or both types of processes. Discounting the implausible possibility that visual patterns are represented in memory at all possible magnifications,[1] other types of explanations are difficult to envisage.

As image and scale transformations should take time, the suggested account of size invariance may be evaluated by reaction time methods. In previously reported experiments (Bundesen & Larsen, 1975), the time necessary to decide whether two simultaneously presented random figures had the same shape was found to be a linearly increasing function of the linear size ratio of the figures. Absolute sizes and size differences apparently had no effect per se. The results suggested that the task was performed by encoding one of the figures as a visual image, by transforming this image to the size format of the other figure, and then testing for a match. Further experiments are reported in the present article. Experiment 1 extended the previous findings on image transformations. The possible role of scale transformations in visual recognition was investigated in Experiments 2 and 3.

Experiment 1

The interpretation of the findings on simultaneous matching reported in Bundesen and Larsen (1975) suggested that a similar pattern of results could be obtained by using a comparable successive-matching task. Experiment 1 tested this conjecture.

Method

Subjects. Seven subjects participated, including the authors. The subjects were students or members of the staff at Copenhagen University. Five subjects had previous experience with same-different reaction time tasks. All subjects were between 20 and 40 years old and had normal or corrected vision.

Stimuli. The stimulus material consisted of 400 pairs of slides. Each slide showed a black solid shape on a white background. Within each of the 200 positive pairs, the solid shapes were identical except for a geometric multiplication. The linear size ratio of a pair was either 1, 2, 3, 4, or 5. These five values were equally frequent, as were magnification and demagnification within pairs. The negative pairs were constructed in the same way as the positive ones, except that one shape in each pair was rotated π rad in the picture plane (see Figure 1).

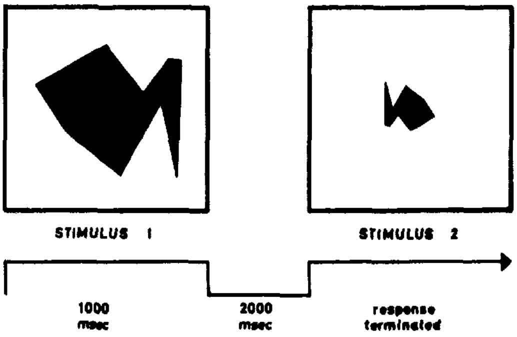

Figure 1. Example of a negative stimulus display in Experiment 1.

The slides were prepared from closed outline drawings filled in with india ink. The outline drawings were constructed randomly by essentially the same method as used in Experiment 2 of Bundesen and Larsen (1975). A total of 200 different pairs of drawings were employed such that two identical pairs of slides were made from each pair of drawings. The 400 pairs of slides were arranged in a standard sequence which was generated at random with the constraint that identical pairs of slides were separated by at least 50 other stimulus pairs.

Procedure. Each subject served individually in two experimental sessions. In the first experimental session, the 800 slides were presented once in the standard sequence. In the second session, the stimulus sequence was repeated backwards so that the order of presentation was reversed both within and between stimulus pairs. The experimental sessions were preceded by practice sessions of about 30 min, in which similar stimulus material was employed to familiarize the subject with the apparatus and procedure.

The subject was seated about 3 m in front of a screen on which the projections of the slides spanned approximately .33 rad horizontally and .22 rad vertically. Each random shape was presented such that its center of gravity (defined as the first moment of area) was positioned at the midpoint of the screen. The largest projected

---

[1] It might be objected that Fourier techniques have shown the feasibility of the related hypothesis that templates may effectively be stored for all possible translations of a given pattern (see, e.g., Duda & Hart, 1973). Anyway, the experiments to be described speak strongly against the hypothesis re size.

shapes spanned about .20 rad and the smallest shapes about .03 rad. During projection of a slide, the pupils of the subject received an illuminance of approximately 10 lx from the stimulus field and 3 lx from the surroundings.

A trial began when the subject depressed a starting key which immediately released the exposure of the first shape in a stimulus pair. The exposure lasted for 1 sec. After a blank interstimulus interval of 2 sec, the second shape was projected. The subject was instructed to decide "as quickly as possible" whether the two stimulus shapes were identical except for a change of size. If they were, he pressed a button on his right; otherwise he pressed a button on his left. The exposure of the second stimulus shape terminated when one of these buttons was pressed. The experiment was run by a laboratory computer, which measured reaction time (msec) from the onset of the second stimulus exposure. After each experimental session, the subject was asked to report upon his strategies for performing the task.

Results

All reaction times longer than 1,500 msec were excluded from the analysis, which eliminated 10 out of 5,600 trials. The individual error rates ranged between 2% and 8%, which seemed reasonably low. Only correct reactions were analyzed with respect to latency.

As shown in Figure 2, mean reaction time across subjects, sessions, and response types (positive vs. negative) increased rather precisely as a linear function of the size ratio of stimulus pairs. The slope constant was about 14 msec. Mean reaction time for positive responses was consistently shorter than mean reaction time for negative ones, but both were approximately linear functions of size ratio, and the rates of increase were very nearly the same. While false alarms were more frequent than misses, the error rates were roughly constant over values of size ratio.

The latency effect of magnification versus demagnification within stimulus pairs is also illustrated in Figure 2. For each type of pairs, mean reaction time increased approximately linearly as a function of size ratio, but the magnification reaction time function was less steep than that for demagnification: The slopes were about 8 and 21 msec, respectively. Error rates were less informative.

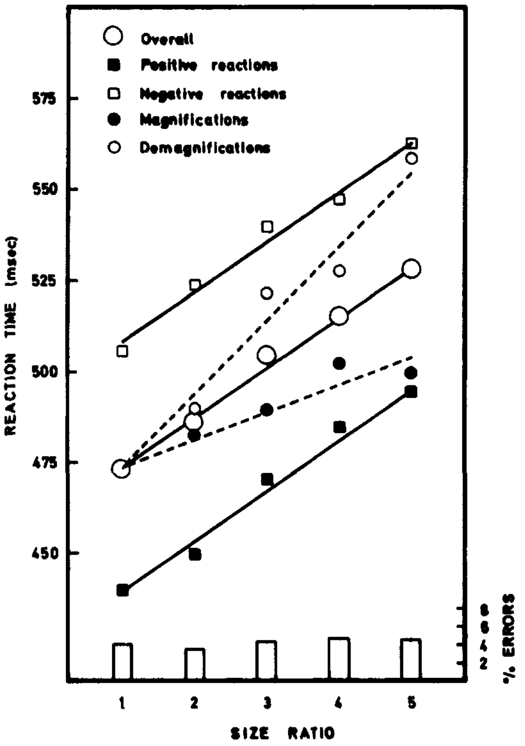

Figure 2. Mean reaction times for correct positive and negative responses, their mean, and the means for correct responses to magnification and demagnification stimulus-pairs as functions of linear size ratio in Experiment 1. (Bottom panel shows mean rate of errors.)

Individual data were subjected to a median-based statistical analysis. For each subject, session, and response type, a pair of straight lines intersecting at size ratio equal to 1 was fitted to median reaction time as a function of size ratio: one for responses to magnification pairs, one for demagnification pairs. Slopes were the same for positive and negative responses. The line fitting was done by an iterative method of minimum chi-square, where goodness of fit was evaluated by testing the hypothesis that for each value of size ratio, the probability that a reaction time fell above the fitted line was .5. Overall, the minimum chi-square fits were acceptable, $\chi^2(196) = 217.8$, $p = .14$.

Subjective reports. All subjects reported that they retained the first stimulus of a pair in a visual form during the interstimulus interval. Several subjects claimed that the image retained was highly schematic, whereas one subject stressed the importance of physiognomic characteristics. The introspective data were less clear concerning the basis of the yes–no decision once the second stimulus had appeared.

Discussion

The pattern of reaction times obtained in Experiment 1 agrees well with the findings reported in Bundesen and Larsen (1975), and the interpretation that was previously advanced for the case of simultaneous matching may accordingly be extended to the successive-matching task. On this interpretation, the subjects used a strategy of encoding the first stimulus in a pair as a more or less schematic visual image which was retained during the interstimulus interval. When the second stimulus appeared, the visual image was gradually transformed to fit the size format of that stimulus. Following the transformation, the image was used for position-wise comparison against the stimulus. If they matched, a positive response was made, and otherwise a negative response. The interpretation explains readily that reaction time increased with size ratio and that the rate of increase was the same for negative as for positive reactions. It is also consistent with, and partially supported by, the subjective reports.

Concerning the generality of the findings in Experiment 1, it should be noted that in unpublished work we have obtained linear successive-matching reaction time functions with slopes comparable to those reported for the present experiment except that the function for magnification was as steep as, or sometimes steeper than, that for demagnification. Whereas linearity appears to be a general characteristic, the effect of magnification versus demagnification is apparently dependent on minor variations in procedure.

Experiment 2

The purpose of Experiment 2 was to investigate the possible role of scale transformations in visual pattern recognition.

Suppose that size-invariant pattern recognition is obtained by scale transformations when recognition is based on the comparison of stimulus patterns against visual representations in long-term memory. A scale transformation was defined as a transformation of the unit of a perceptual reference system which should constitute the current interpretation of a standard reference system for long-term representations of visual patterns. Let the *assumed format* of a stimulus pattern be that size format for which the perceptual reference system is currently set. Provided that a scale transformation is a gradual process, recognition time would then be expected to increase systematically with size divergence between the assumed format and the actual stimulus format. Experiment 2 was an attempt to test this prediction in a serial character-recognition task which used the transitional probabilities in the sequence of stimulus formats for controlling the assumed formats of stimulus patterns.

Method

Subjects. Eight subjects with normal or corrected vision participated. The subjects were students or members of the Copenhagen University staff, with ages ranging between 25 and 35 years. All subjects were acquainted with reaction time tasks, and three subjects, including the authors, had previously served in Experiment 1.

Stimuli. The stimulus material consisted of slides which were photographed from computer-generated drawings in black and white. Each slide showed a capital letter. The letter was either normal (positive) or rotated π rad in the picture plane (negative), and the letter type was either *A, B, C, D, E, F, G, J, K, L, P, Q, R, T, U, V,* or *Y*. The letters appeared in four fixed size formats with linear size ratios of 1:2:6:9 (see Figure 3).

The slides were arranged in a standard sequence in which the first-order probability of transitions from one size format to the same size format was .75 regardless of letter types and orientations. Specifically, the sequence contained 1,153 positions. Discounting the initial position, normal letters were contained in half of the sequential positions and rotated letters in the rest. For any of the four size formats, normal letters in that format were immediately preceded by other letters in the same format for a total of 108 times. They were preceded by letters in

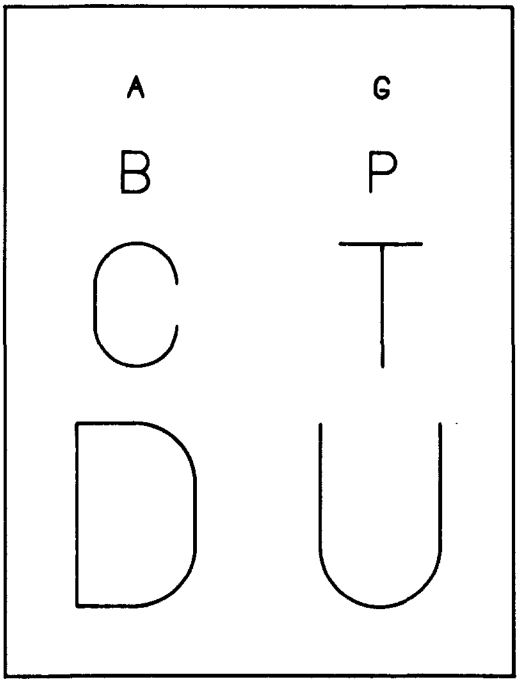

Figure 3. Examples of stimulus characters in different size formats (1:2:6:9) used in Experiment 2.

different formats for a total of 36 times, namely, 12 times for each of the three formats. Exactly the same was true for rotated letters in any format. Otherwise, sequential positions for normal versus rotated letters were chosen randomly. Finally, for any sequential position, the letter type was drawn at random from the ensemble above with the single constraint that the same type was never used for two positions in immediate succession.

Procedure. Subjects served individually in four experimental sessions in which the standard sequence of 1,153 slides was run forwards, backwards, backwards, and forwards, respectively. The experimental sessions were preceded by a practice session of about 30 min in which the same type of stimulus material was employed. During the practice session, the subject was carefully informed about the statistical properties of the stimulus sequence with respect to transitions between size formats. After the experimental sessions, the subject was asked to report upon his strategies for performing the task.

Trials were blocked within sessions. A block of approximately 77 trials began when the subject pressed a starting key which released the exposure of the first stimulus with a latency of 2 sec. The task was to decide "as quickly as possible" whether the stimulus presented was an upright letter. If it was, the subject pressed a button on his right; otherwise he pressed a button on his left. Reaction time was measured from stimulus onset. When a response button was pressed, stimulus exposure terminated with a latency of .5 sec, and after a fixed blank intertrial interval of 2 sec, the next stimulus was exposed. The task was self-paced between blocks. The first two trials in a new block repeated the last two trials in the previous one, whence the first two reactions were not recorded.

All stimulus letters were centered on the projection screen, where the largest format spanned about .18 × .12 rad and the smallest format about .02 × .013 rad. Viewing conditions and apparatus were otherwise the same as in the previous experiment.

Results

Subjective reports. Though some of the subjects felt unable to report upon their perceptual strategies, most of the subjects reported that their performance was determined by the sequential structure of the task: Following the presentation of a letter in a given size format, they were perceptually prepared for letters in the same format. If the stimulus letter appeared in another format, they had to adjust themselves to that format in order to recognize the letter. This feeling was especially pronounced for grossly different formats.

Reaction times. Reactions with latency above 1,500 msec were not analyzed, which eliminated 15 trials from a total of 36,864 trials. The individual error rates ranged between 3% and 9%, which seemed acceptable. Only correct responses entered the analysis of reaction times.

The *stimulus format* of a letter in the stimulus sequence was defined as the relative size of the format of the letter, the value being 1, 2, 6, or 9. The *cued format* of a stimulus letter was defined as the stimulus format of the immediately preceding letter in the stimulus sequence.

Table 1 shows mean reaction times for positive and negative responses, and their mean, as functions of stimulus format and cued format across subjects and sessions. For any format combination, positive reactions were faster than negative reactions; the positive–negative difference was rather stable over format combinations, ranging between 28 and 56 msec, with a mean of 40 msec. Across response types, the pattern was as follows: For each stimulus format, reaction time was shortest when the cued format equalled the stimulus format, increasing approximately monotonically with divergence in both directions. Similarly, for any value of cued format, reaction time was shortest when the stimulus format took this value, increasing monotonically with divergence in both directions. Finally, to a first rough approximation, the reaction time for a given combination of cued format and stimulus format was about the same as the reaction time for the reverse combination.

The main diagonal of the matrix of pooled reaction times in Table 1 shows that when the cued format equalled the stimulus format, mean reaction time was a U-shaped function of the size format, with minimum at format value 2. The variation along the diagonal spanned about one fourth of the total range of the reaction times in the matrix. Taking this variation into account, the *reaction time increment* for a given combination of cued format (f) and stimulus format (g) was computed as the mean reaction time for $f \times g$ minus the mean reaction time for format combination $g \times g$. Furthermore, the *size ratio* associated with a letter in the stimulus sequence was defined as the linear ratio between the cued format and the stimulus format such that this ratio was either 1, 1.5, 2, 3, 4.5, 6, or 9.

Table 1
Mean Reaction Time (in msec) as a Function of Stimulus Format, Cued Format, and Type of Response (Positive vs. Negative) in Experiment 2

Stimulus format	Cued format 1	2	6	9
	Positive			
1	445	468	488	501
2	453	447	470	478
6	479	472	454	472
9	498	490	479	459
	Negative			
1	496	500	537	535
2	490	490	498	515
6	508	528	494	517
9	543	523	516	503
	Overall			
1	471	484	513	518
2	472	469	484	497
6	494	500	474	495
9	521	507	498	481

Note. Data are for correct responses.

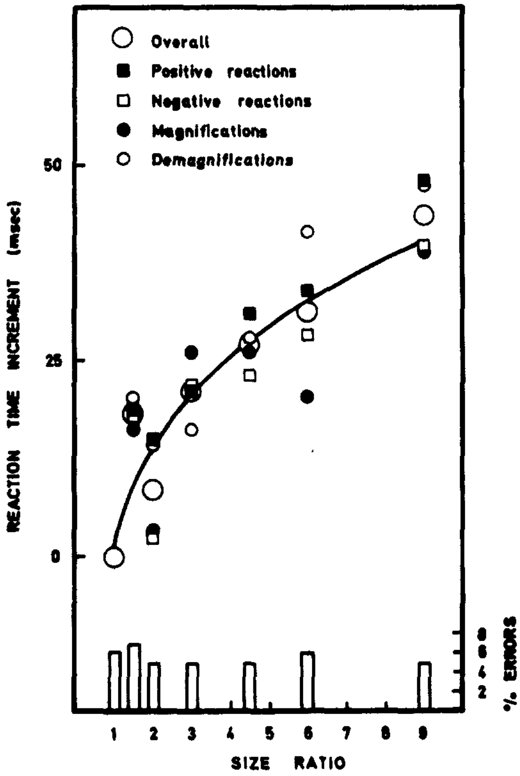

Figure 4. Mean reaction time increments for correct positive and negative responses, their mean, and the means for correct responses after format magnifications and demagnifications as functions of linear size ratio of cued size format and stimulus format in Experiment 2.

Figure 4 shows mean reaction time increments as functions of size ratio for positive and negative reactions, for reactions after format magnifications and demagnifications, and overall. A single logarithmic function through the point (1, 0) provides a reasonable fit to the data.[2] As shown in

---

[2] Mathematical simplicity favors the choice of logarithmic relations for fitting the data in Figure 4, since this is the only type of nonconstant

the bottom panel of Figure 4, the mean rate of errors was approximately constant over values of size ratio. The error rates were also about the same for positive and negative reactions and for cases of magnification and demagnification of size formats.

Individual data were subjected to a median-based statistical analysis. For each subject, session, and response type, the increment in any reaction time associated with format combination $f \times g$ was computed by subtracting the median reaction time for format combination $g \times g$ with the subject, session, and response type concerned. For each subject and session, then, a minimum chi-square logarithmic curve through the point (1, 0) was fitted to median reaction time increment as a function of size ratio. Goodness of fit was evaluated by testing the hypothesis that for each response type and size ratio, the probability that a reaction time increment fell above the fitted curve was .5 for both format magnifications and format demagnifications. Each of the 32 minimum chi-square curves was found to increase. Overall, the logarithmic fits were satisfactory, $\chi^2(736) = 732.2$, $p = .53$.

Discussion

In the present experiment, stimulus recognition was presumably achieved by comparison against visual representations in long-term memory. The results support the hypothesis that in this case, pattern recognition presupposed that the subject's perceptual reference system (i.e., the current interpretation of the standard reference system for long-term representations of visual patterns) was scaled to the size format of the stimulus pattern. The suggested account by scale transformations is as follows:

1. At any time, the subject's perceptual reference system was adjusted to letters of a certain size format, the format currently assumed. At the beginning of a trial, the assumed format approximated the cued format, which was the format of the immediately preceding letter of the stimulus sequence.
2. When the stimulus letter was exposed, the size of the letter was computed prior to the recognition of the letter. If the size format diverged from the format assumed, letter recognition presupposed that the perceptual reference system was rescaled.
3. Rescaling of the perceptual reference system was realized as a gradual transformation by which the assumed format changed towards the format of the stimulus letter. The time taken by this scale transformation was roughly proportional to the logarithm of the linear size ratio of cued format and stimulus format.

Whereas the results of Experiment 2 are readily explained in terms of scale transformations, a plausible account purely in terms of image transformations is difficult to envisage. Consider first the idea of stimulus normalization. The assumption that recognition is based on transforming a visual image of the stimulus to a fixed standard format could possibly explain temporal effects of absolute stimulus format. The gross effects in Experiment 2, however, were associated with the relationship between the stimulus format and the cued format, and these effects are not explainable by the normalization hypothesis, as the cued format was not a constant standard format. Generally speaking, the two processes of stimulus normalization (size-transforming a stimulus image to fit the scale of a given reference system) and scale transformation (size-transforming the scale of the reference system to fit the stimulus image) are complementary to each other. The unequal power of these processes in accounting for the present data arises from the fact that only scale transformation can logically take place before the stimulus has been presented.

Another theoretical attempt to explain the results of Experiment 2 in terms of image transformations might assume that the positive items were retained as a stack of visual images in short-term memory. If

---

continuous functions t such that $t(x \times y) = t(x) + t(y)$, where the arguments are arbitrary size ratios.

this stack was preset to fit the cued format, and recognition was achieved by matching the stimulus against the stack, appropriate image transformations (of stimulus or stack) would be called for. However, explanations along such lines may be rejected a priori, since the positive set was presumably much too large to be contained in visual short-time memory.

A size-scaling explanation of the results of Experiment 2 must apparently refer to scale transformations, while reference to image transformations is not needed. Since this conclusion does not depend on our previous interpretations of other experiments, it is tempting to ask whether the previous studies may possibly be reinterpreted in light of the new findings. Thus, if the previous results on simultaneous (Bundesen & Larsen, 1975) and successive (Experiment 1) matching could be explained by scale transformations, any reference to image transformations might be avoided.

Suppose, for the sake of argument, that the matching task was performed by first encoding one of the stimulus patterns in long-term memory and then matching the other pattern against this memory representation. Assume that the long-term encoding, as well as the matching, presupposed that the perceptual reference system was adjusted to fit the size format of the stimulus pattern in question. If so, a scale transformation by which the assumed format changed from the size of the first stimulus to the size of the second stimulus would be implied. By substituting scale transformations for image transformations, a partial account of the matching reaction time data is thus available. Against this type of interpretation, however, the following objections can be made. First, the nature of the matching tasks previously employed does suggest that short-term representations, rather than long-term representations, should be used for matching, and that suggestion was clearly supported by the introspective reports of retaining the first stimulus as a visual image during the interstimulus interval in the successive-matching task. Second, the (linear) reaction time functions obtained in the matching experiments were so different from the (logarithmic) functions found in the character-recognition task of Experiment 2 that these functions are not likely to be explained by the same process of scale transformation.

Experiment 3

The suggested account of visual size invariance in terms of image and scale transformations would be strongly supported if contrasting roles of the two types of size scaling could be evidenced in the performance of a single experimental task. Further, by separating the effects of image and scale transformations in a given experiment, one would expect to gain valuable insight into the specific processing strategy at work in the experimental situation.

Consider the memory-scanning task that was developed and refined by Sternberg (1966, 1969). In this task, the subject memorizes a short list of items defining the positive set of stimuli. When a test stimulus is presented, the subject must indicate as rapidly as possible whether it is contained in the positive set. In typical fixed-set procedures, the same positive set is used for a block of many consecutive trials, each of which consists only of warning signal, test stimulus, and response. A simple serial case is obtained when the response-stimulus interval is fixed within blocks and the warning signal is omitted. Formally, each block of trials in the serial task becomes a special case of the paradigm of Experiment 2 if (a) the stimulus ensemble consists of characters in different size formats, (b) the sequence of formats is governed by appropriate transitional probabilities, and (c) size is disregarded in the definition of the positive set. By analogy with Experiment 2, scale transformations might be expected to serve for size-invariant recognition in this situation. On the other hand, it is reasonable to suppose that image-operations may be more efficient, and hence take over, when stimulus repetitions occur: If a visual image persists from the preceding stimulus presentation, a repeated character can be recognized as such by being matched against the image, and the previous (stored) response may at once be repeated. The supposition accords with results from several studies of sequential effects in choice reaction time (see Bertelson, 1965; Eichelman, 1970; Rabbitt, 1968; Smith, 1968; and the review in Kornblum, 1973). When a character is repeated in a new size format, the hypothesized strategy requires a process of image transformation. These considerations suggest that a suitable version of the Sternberg task may serve to contrast the effects of image and scale transformations in a simple situation.

In typical experimental conditions, mean reaction time in the memory-scanning task is found to increase approximately linearly as a function of positive set size. The rates of increase for positive and negative reactions are about equal. To account for these results, Sternberg (1966) developed a well known model in which an encoded representation of the test stimulus is serially compared with memory representations of the items in the positive set. The comparison is exhaustive, even for positive test stimuli, and reactions are based on decisions as to whether or not matches have occurred. The serial-exhaustive scanning model has its problems (e.g., Corballis, 1975; Wickelgren, 1975), but the basic conception still seems plausible (cf. Sternberg, 1975). A question of major interest, then, concerns the nature of those internal representations among which comparisons are assumed to be made. Existing data (Posner, 1973; Sternberg, 1967, 1969) suggest that in typical visual experiments, the test stimulus is encoded as a refined visual image, which is subsequently compared against images of the positive items held in visual short-term memory. If so, introduction of size variation of test stimuli into typical experimental conditions may primarily be expected to call on image transformations, not scale transformations.

A different pattern of results has emerged from experiments in which subjects had extended practice with the same fixed sets and response consistency prevailed. Response consistency means that for each subject and all trials, each item in the stimulus ensemble consistently requires only a positive or only a negative reaction. Under these conditions, the reaction time functions become flatter and more closely approximated by logarithmic functions than by linear ones (Kristofferson, 1972; Ross, 1970; Simpson, 1972; Swanson & Briggs, 1969).[3] The effect of practice with a given set of characters is highly specific to the set employed, but it transfers across character cases (Ross, 1970).

The finding of specific transfer of training to positive sets which are nominally (but not visually) equivalent to the practiced ones is very suggestive. It argues against the possibility that the effect of positive set size is generated at a level that is lower than that at which conceptual codes are linked to memory representations of sensory patterns. Specifically, the effective representation of the positive set can hardly be a collection of visual features, nor a stack of visual images. The introduction of size variation of test stimuli in memory-scanning experiments employing well-practiced small fixed sets and response consistency would therefore be expected to call on scale transformations. Provided that visual-image matching takes over on repetition trials, this type of task should serve for contrasting image and scale transformations in a single setting, which was the main purpose of Experiment 3.

A second aim of this experiment was to elucidate the specific processing strategy used in the selected sort of memory-scanning task with well-practiced fixed sets. By the above arguments, the type of size scaling evidenced should help to converge upon the level of processing at which the functionally effective representation of the positive set is located. Let a *descriptor* be a

---

[3] In some of the experiments considered (Kristofferson, 1972; Ross, 1970), the positive sets were also nested (i.e., each positive set contained all the members of smaller positive sets), but nesting is not decisive (Simpson, 1972; Swanson & Briggs, 1969).

memory unit in which one or more long-term representations of sensory patterns are connected to a given conceptual code (compare, e.g., the "conceptual-store nodes" proposed by Atkinson, Herrmann, & Wescourt, 1974). Scale transformations, then, are assumed to mediate comparison of stimulus patterns against memory representations at the level of descriptors. A pattern of reaction times indicating scale transformations would accordingly suggest that the reactions were contingent upon processing at the descriptor level. If so, the functionally effective representation of the positive set should be located at or beyond this level.

Another converging operation is required if we wish to discriminate between positive-set representations at and beyond the descriptor level. Suppose the composition of the positive set is only specified at a level beyond the descriptors. Two subcases have some plausibility. First, the positive-set representation could be located in verbal short-term memory, and second, the location could be in another division of long-term memory forming some sort of "event-knowledge store" (cf. Atkinson et al., 1974). In either case, the representation of the positive set is assumed to be nonvisual in nature. Hence, the hypothesized process of comparing an encoded version of the test stimulus against members of the positive set should not be sensitive to the visual similarity between probe and targets. Visual confusability between members of the stimulus ensemble could influence stimulus encoding at the level of descriptors, but this influence should be independent of the definition of the positive set. Thus, unless the composition of the positive set is somehow specified at (or below) the descriptor level, effects on reaction time due to visual similarity between members of a given stimulus ensemble would not be expected to depend on the definition of the positive set. On the other hand, if the positive set is effectively represented at (or below) the descriptor level, where visual comparisons are made, effects on reaction time due to visual similarity might be expected to depend critically on the composition of the positive set.

In sum, by adding confusion data to results on size scaling, we hoped to converge upon the memory location of the positive-set representation used in the selected sort of memory-scanning task employing well-practiced small fixed sets and response consistency.

Method

Stimuli. The stimulus slides were photographed from computer-generated drawings similar to those employed in Experiment 2. Each slide showed a normal capital letter in one of three fixed size formats with linear size ratios of 1:2:9.

The positive set was either *A, B, C, AB, AC, BC*, or *ABC*. The negative set always consisted of letters *D* through *Z*. For each of the seven positive-set conditions, a stimulus sequence was generated such that the first-order probability of transitions from one size format to the same size format was .75. The number of positions in a sequence was 384, 512, and 1,152 for positive set sizes of 1, 2, and 3, respectively. Each of the seven sequences was constructed to fulfil the following conditions as exactly as possible: (a) Positive and negative stimuli were equally frequent, as were the different positive letters. (b) For any positive letter, the three size formats were exemplified equally frequently. (c) For any size format, each of the positive letters in that format was immediately succeeded by letters in the same format with a frequency of .75, and the remaining immediate successors were divided equally among the other two formats. (d) Conditions b and c remained satisfied if cases of positive letters immediately preceded by letters in a different format were disregarded. (e) Conditions b, c, and d were also satisfied when cases of stimulus repetition (with respect to letter type) were considered separately. (f) The set of negative letters as a whole satisfied the analogs of conditions b, c, and d. (g) Negative stimulus repetitions did not occur. In other respects, the sequence was random.

Each of the stimulus sequences for the seven positive-set conditions was divided into blocks of about 75 consecutive members. Duplicates of the last two members of any block were added to the beginning of the following one (if any). The total set of 51 blocks was finally arranged in a counterbalanced order which defined the standard sequence of stimuli.

Subjects and procedure. Seven subjects were drawn at random from those who had served in previous experiments. Each subject participated in two experimental sessions during which the blocked standard sequence of 3,928 slides was run forwards and backwards, respectively. Prior to each block of trials, the composition of the positive set was orally announced by the experimenter. The subject was instructed to decide as rapidly as possible whether stimulus letters belonged to the positive set. The composition of the negative set was never made explicit. The task was self-paced between blocks, except that a 1-hour break was requested in the middle of each session. Apparatus and procedure were otherwise exactly the same as in Experiment 2.

Results

Reactions with latency above 1,500 msec were not analyzed, which eliminated 12 out of 53,564 trials. The individual error rates ranged between 2% and 11%. Only correct responses were analyzed with respect to latency.

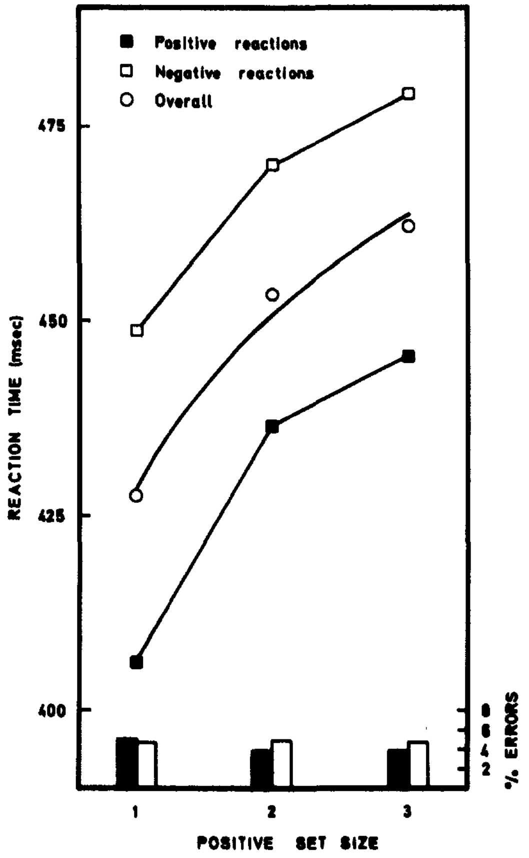

Figure 5. Mean reaction times for correct positive and negative responses and their mean as functions of positive set size in Experiment 3. (Bottom panel shows rates of false alarms [solid bars] and misses [open bars].)

The overall effect of positive set size is summarized in Figure 5. While positive reactions were faster than negative ones, mean reaction time was an increasing, negatively accelerated function of positive set size for each type of response (see upper and lower curves). By least squares logarithmic regression, the rate of increase with positive set size was 25 msec per $\log_2$ unit for positive reactions and 19 msec per $\log_2$ unit for negative reactions. The interaction between positive set size and response type was significant.[4] Across response types, mean reaction time was approximately a logarithmic function of positive set size with a slope of 22 msec per $\log_2$ unit (see middle curve). Error rates were rather stable over set sizes and response types, ranging between 3.9% and 5.1% (see bottom panel).

Mean reaction times for stimulus-nonrepetition trials are shown in Table 2 for each value of positive set size, for each of the nine combinations of cued format and stimulus format, and for each type of response. Restricting the analysis to stimulus-nonrepetition trials raised the mean latency of positive reactions by about 8 msec and lowered the rate of increase with positive set size by about 2 msec per $\log_2$ unit for the positive reactions; otherwise, the reaction time pattern in Figure 5 was not appreciably affected.

For each value of positive set size and each type of response, mean reaction time varied systematically with the relation between cued format and stimulus format (see Table 2). For each stimulus format, reaction time tended to be shortest when the cued format equalled the stimulus format; with

---

[4] The following convention is adopted in this article: When an effect is reported to be significant, and nothing else is indicated, it is implied that this effect was significant at the .01 level by a sign test based on averaged results for each subject and session. In the present case, for example, a least squares logarithmic regression of mean reaction time as a function of positive set size was made for each type of response and for each subject and session. For any of the 14 combinations of Subjects × Sessions, the rate of increase in latency per log unit was higher for positive reactions than for negative ones, which has a probability below .01 by a two-tailed sign test.

Table 2

Mean Reaction Time (in msec) as a Function of Stimulus Format, Cued Format, Positive Set Size, and Type of Response (Positive vs. Negative) for Stimulus-Nonrepetition Trials in Experiment 3

Stimulus format	Positive cued format 1	Positive cued format 2	Positive cued format 9	Negative cued format 1	Negative cued format 2	Negative cued format 9
			Set size 1			
1	413	423	450	451	454	473
2	404	407	439	450	449	449
9	430	431	414	458	441	441
			Set size 2			
1	444	452	461	473	475	486
2	443	438	454	474	463	487
9	465	464	443	483	480	465
			Set size 3			
1	447	458	474	480	479	507
2	452	445	463	480	476	490
9	479	468	452	488	479	473

Note. Data are for correct responses.

divergence in either direction, reaction time tended to increase. Further, for any value of cued format, reaction time tended to be shortest when the stimulus format took this value, increasing with divergence in either direction. For cued format equal to stimulus format, the variation in reaction time as a function of size format was less consistent, though the difference in latency between format values 1 and 2 (maximum and minimum, respectively) was significant by sign test across set sizes and response types ($N = 84$).

Based on the data in Table 2, panel A in Figure 6 shows mean reaction time increment as a function of linear size ratio between cued format and stimulus format. The function is approximated by a least squares logarithmic curve through the point (1, 0). Three different breakdowns of the function are illustrated in panels B, C, and D. As indicated in panel B, the effects of size ratio and positive set size were approximately additive. However, as shown in panel C, mean reaction time increment with divergence between cued format and stimulus format was higher for positive reactions than for negative ones. Finally, as shown in panel D, mean reaction time increment was slightly higher for demagnification of size formats than for magnification of size formats. The interaction between size ratio and response type was significant by sign test ($N = 252$), while the interaction of size ratio with the factor of magnification versus demagnification appeared random ($N = 252$, $x = 132$).

The interaction between size ratio and response type with respect to speed of reactions was accompanied by interaction with respect to accuracy. With divergence between cued format and stimulus format, the mean rate of misses increased from 4.6% to 5.6%, while the rate of false alarms decreased from 4.4% to 3.9%. Across response types, however, error rate was almost constant over values of size ratio, ranging between 4.5% and 5.1%.

Individual data for stimulus-nonrepetition trials were subjected to a median-based statistical analysis similar to that employed in Experiment 2. For each subject and session, a minimum chi-square logarithmic curve through the point (1, 0) was fitted to me-

Table 3

Mean Reaction Time (in msec) as a Function of Stimulus Format, Cued Format, and Positive Set Size for Positive Stimulus-Repetition Trials in Experiment 3

Stimulus format	Cued format 1	Cued format 2	Cued format 9
		Set size 1	
1	389	407	470
2	387	385	429
9	454	433	383
		Set size 2	
1	406	415	454
2	433	399	436
9	456	451	397
		Set size 3	
1	404	386	457
2	421	403	437
9	458	449	398

Note. Data are for correct responses.

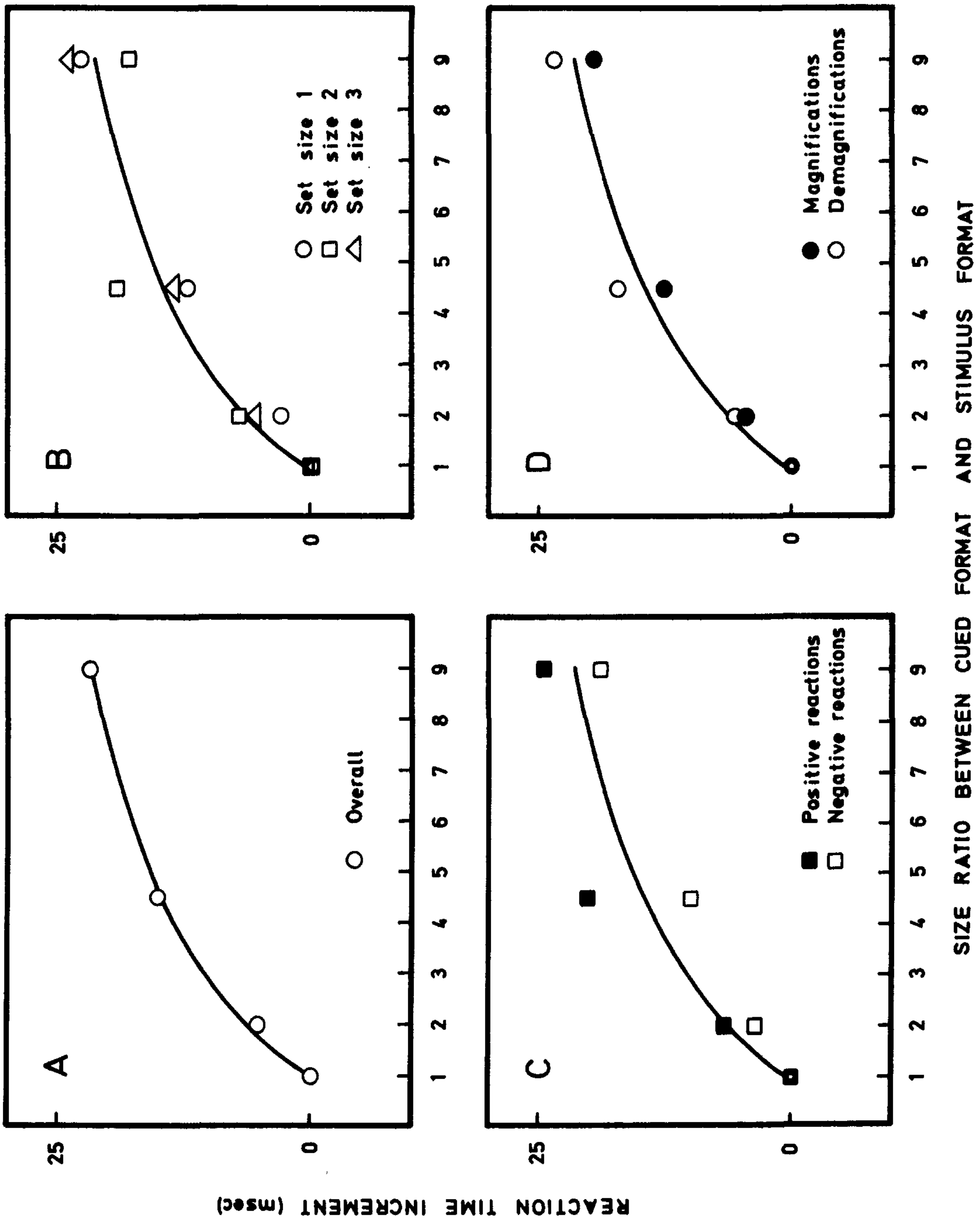

Figure 6. Mean reaction time increments for correct responses to stimulus nonrepetitions as functions of linear size ratio of cued size format and stimulus format in Experiment 3. (A. mean across all conditions; B: means for positive set sizes 1, 2, and 3; C: means for positive and negative responses; D: means for responses after format magnifications and demagnifications.)

dian reaction time increment as a function of size ratio. Goodness of fit was evaluated by testing the hypothesis that for each response type, set size, and size ratio, the probability that a reaction time increment fell above the fitted curve was .5 for both format magnifications and format demagnifications. As should be expected from previous indications of interaction between size ratio and response type, the fits were not acceptable, $\chi^2(490) = 631.9$, $p < 10^{-4}$. For each subject and session, then, two separate minimum chi-square logarithmic curves through the point (1, 0) were fitted to median reaction time increment as a function of size ratio: one for positive reactions, one for negative reactions. In each case, goodness of fit was evaluated by testing the hypothesis that for each set size and size ratio, the probability that a reaction time increment fell above the fitted curve was .5 for format magnifications as well as for format demagnifications. Overall, these fits were acceptable; for positive reactions, $\chi^2(238) = 267.2$, $p = .09$; for negative reactions, $\chi^2(238) = 243.1$, $p = .40$; in total, $\chi^2(476) = 510.3$, $p = .13$.

Mean reaction time for stimulus-repetition trials is shown in Table 3 as a function of stimulus format, cued format, and positive set size. The effect of positive set size was much less for stimulus repetitions than for stimulus nonrepetitions. Across format combinations, the repetition reaction times averaged 397, 411, and 410 msec for set sizes 1, 2, and 3, respectively. The difference in latency between set size 1 and set sizes 2 and 3 was significant by sign test ($N = 14$). The interaction of positive set size with the factor of repetition versus nonrepetition was significant whether positive or negative nonrepetitions were considered.

Mean reaction time for stimulus repetitions also varied systematically with the relation between cued format and stimulus format (see Table 3). Across values of positive set size, mean reaction time for any stimulus format was shortest when the cued format equalled the stimulus format, increasing monotonically with divergence in either direction. Similarly, for any value of cued format, mean reaction time was shortest when the stimulus format took this value, increasing with divergence in either direction. Across values of size ratio, mean reaction time for format magnifications exceeded that for format demagnifications by about 6 msec; this difference was not significant by sign test ($N = 126$). For size ratio equal to 1, mean reaction time showed some decrease with increasing size format. The difference in latency between format values 1 and 9 (maximum and minimum, respectively) was about 7 msec, and this difference was significant by sign test ($N = 42$).

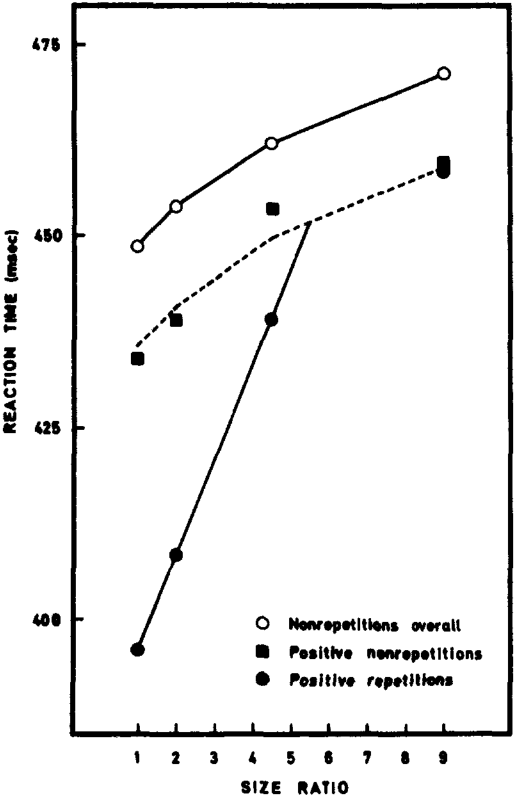

Figure 7. Mean reaction times for correct responses to positive stimulus repetitions, positive stimulus nonrepetitions, and stimulus nonrepetitions pooled across response types as functions of linear size ratio of cued size format and stimulus format in Experiment 3.

Figure 7 shows mean reaction time as a function of size ratio for (positive) stimulus repetitions, for positive stimulus nonrepetitions, and for stimulus nonrepetitions pooled across response types. The curve for pooled nonrepetitions was shifted downwards by

Table 4
Mean Reaction Times (in msec) and Error Rates (in percent) as Functions of Positive Set Size for G, O, *and* P *in Visual-Confusability Conditions and for Negative Letters Overall in Experiment 3*

Stimulus letters	Reaction time set size			Error rate set size		
	1	2	3	1	2	3
G, *O*, and *P*	507	504	508	13.4	9.8	8.2
All negatives	447	467	476	5.1	3.9	4.4

Note. Data are for linear size ratio of stimulus format and cued format equal to 1. Reaction time data are for correct responses.

some 12 msec to give a rough fit to the function for positive nonrepetitions, whereas the lower part of the function for positive repetitions was fitted by a straight line segment with a slope constant of 12.3 msec. It may be noted that the general shape of the reaction time function for positive repetitions would not be affected by plotting reaction time increments instead of reaction times.

As is evident from Figure 7, the reaction time function for positive stimulus repetitions was grossly different from those for nonrepetitions. For size ratio equal to 1, mean reaction time for positive stimulus repetitions was 38 msec shorter than the mean for positive nonrepetitions. Over size ratios 1, 2, and 4.5, the positive repetition reaction times showed a steep linear increase, approaching the function for positive nonrepetitions. Finally, for size ratio equal to 9, the reaction time for positive stimulus repetitions was almost the same as the reaction time for positive nonrepetitions.

The task demanded that reactions to stimulus repetitions should be response repetitions. The effect of response repetition per se was evaluated by comparing response-repetition reaction times on stimulus-nonrepetition trials with the corresponding response-nonrepetition reaction times. The main result was that across format combinations and across set sizes 2 and 3, positive mean reaction time for stimulus nonrepetitions was about 9 msec longer for response repetitions than for response nonrepetitions. Similarly, across format combinations and set sizes, negative mean reaction time was 12 msec longer for response repetitions than for response nonrepetitions. Since stimulus repetitions without response repetitions did not occur for correct reactions, the possibility of interaction between the two types of repetition could not be tested. However, the effect of stimulus repetition was clearly not reducible to that of response repetition.

Stimulus letters *G*, *O*, and *P* were selected a priori for the analysis of visual confusions. As indicated in Figure 3, *G* was generated by adding a single stroke to *C*, *O* was generated from *C* by a smooth completion, and *P* was generated from *B* by deletion. The analysis of reactions to *G*, *O*, and *P* was restricted to cases associated with a size ratio equal to 1. With this restriction, the number of correct reactions to *G*, *O*, and *P* totaled about 2,500.

For set sizes 1 and 2, negative reactions to *G* and *O* were significantly slower when *C* was a member of the positive set (confusability condition) than when *C* was not (nonconfusability condition). Similarly, negative reaction times for *P* were significantly lengthened when *B* was a member of the positive set. For the three negative letters, the latency difference between confusability and nonconfusability conditions averaged about 43 msec across set sizes 1 and 2. A test was conducted to determine whether reaction time for *G*, *O*, and *P* in confusability conditions depended on whether or not the critical positive letter (i.e., *C* for *G* and *O*; *B* for *P*) was presented on the preceding trial. In either case however, mean reaction time for *G*, *O*, and *P* across set sizes 1, 2, and 3 was approximately 506 msec.

Table 4 shows mean reaction time and false alarm rate as functions of positive set size for *G*, *O*, and *P* in confusability conditions and for negative letters overall; both analyses were restricted to cases associated with a size ratio equal to 1. For any value of positive set size, the reactions to *G*, *O*, and *P* were much slower than the average for negative reactions, and the rate of false alarms was higher. Furthermore, while the average negative reaction time increased with positive set size, the reaction time for *G, O,* and *P* was almost constant. Finally, whereas the average rate of false alarms was rather stable over values of positive set size, the false alarm rate for *G, O,* and *P* showed a systematic decrease with increasing set size.

Discussion

The overall effect of positive set size in Experiment 3 (see Figure 5) accords with previous findings from memory-scanning experiments using response-consistent fixed-set procedures with well-practiced small positive sets. The data most closely parallel those obtained by Kristofferson (1972) and Ross (1970) after extended practice with initially unfamiliar positive sets. Thus, by least squares logarithmic regression of Kristofferson's data for the last experimental sessions (results for Days 31–36, estimated from Figure 3 in Kristofferson, 1972), the rate of increase with positive set size was about 28 msec per $\log_2$ unit for positive reactions and about 19 msec per $\log_2$ unit for negative reactions. Ross (1970) reported that the rate of increase was nearly the same for positive as for negative reactions, averaging some 22 msec per $\log_2$ unit for his last session.

Image and scale transformations. The main purpose of the present experiment was to try to contrast the roles of image and scale transformations in a single setting. Disregarding stimulus-repetition trials, it was expected that size-invariant recognition would be achieved by scale transformations. Accordingly, from the findings in Experiment 2, mean reaction time increment was predicted to increase logarithmically with the size ratio between cued format and stimulus format, and the rate of increase was predicted to be the same for each value of positive set size, for each type of response, and for magnification versus demagnification of size formats. These predictions were roughly confirmed by the data (see Figure 6), except that there was a significant interaction between size ratio and response type.

With divergence between cued format and stimulus format, reaction time increments were higher for positive reactions than for negative ones. The observed interaction may possibly be explained by hypothesizing a certain measure of response bias towards negative reactions, and against positive reactions, when discrepancy was detected between the size format assumed for a given stimulus presentation and the actual format of the stimulus. The suggested hypothesis was supported by the fact that with divergence between cued format and stimulus format, the mean rate of misses increased and the rate of false alarms decreased.

As regards the stimulus-repetition trials, reactions could presumably be based on matching the repeated character against a visual short-term image persisting from the preceding stimulus presentation. With this matching procedure, size-invariant recognition should be obtained by means of image transformations. From the findings in previous experiments, it was accordingly expected that, as long as reaction time for stimulus repetitions was shorter than that for nonrepetitions, it should increase linearly with the value of size ratio, and the function should be comparatively steep. This prediction was confirmed by the results (see Figure 7), though a slight effect of size format per se was noted for size ratio equal to 1.

The suggested interpretation implies that image and scale transformations could go on in parallel. Since the occurrence of stimulus repetitions was not predictable, prerecognition decisions to initiate the image-transforming procedure could not be contingent upon the type of trial (repetition vs. nonrepetition). Therefore, even though reactions on nonrepetition trials were supposedly based on scale transformations, recognition by image matching must also have been attempted on these trials, at least when the preceding stimulus was a positive letter. The theoretical possibility of serial organization such that scale transformations were only initiated when image matching had already failed may be excluded; that proposal would imply a higher rate of increase in reaction time with size ratio for nonrepetitions than for repetitions, which goes against the evidence. The obvious conclusion is that image and scale transformations were performed in parallel such that reactions were based on the scale-transforming procedure unless the image-transforming procedure completed with a faster match. This model of a race between the two types of processes also fits the observation that reaction time for repetitions never exceeded that for nonrepetitions.

The hypothesis that different types of processes underlay reactions on stimulus-repetition and stimulus-nonrepetition trials, respectively, received independent support from the fact that the effect of positive set size differed between these cases. As expected from the above account, the effect of set size was much less for repetitions than for nonrepetitions. The reason that set size did influence reaction time on repetition trials significantly may be that with smaller set size, the probability of repetitions increased, whence subjects were more strongly induced to attend to the short-term stimulus-images and thus to preserve the information content of these such that reaction time to repetitions was reduced (cf. Posner et al., 1969, Experiment 3).

In conclusion, the results of the current Experiment 3 support the contention that both scale and image transformations were at work in the experiment. Reactions to stimulus repetitions were normally based on matching stimulus patterns against visual short-term images by means of image transformations. The remaining reactions were based on matching stimulus patterns against long-term representations by means of scale transformations. Apparently, the two types of size scaling could go on in parallel, each with its own time course.

Visual confusions. The next question concerns the possible memory locations of the positive-set representations used in reacting to nonrepetitions. Given that performance was based on size-invariant recognition achieved by scale transformations, and assuming that scale transformations serve for comparison of stimulus patterns against memory representations at the level of descriptors, it is strongly suggested that a functional specification of the positive set was located at or beyond the descriptor level. However, in order to discriminate between positive-set representations at and beyond the descriptor level, other types of evidence must be considered. The analysis of visual confusions showed lengthened negative reaction times for correct responses to stimulus characters that were visually similar to members of the positive set. When *C* was positive, for instance, negative reaction times for *O* were lengthened. Apparently, the specific composition of the positive set influenced processing at levels where visual comparisons were made. Assuming that visual comparisons were not used beyond the level of descriptors, it follows that the composition of the positive set was specified at or below that level. The combined evidence on size scaling and confusions tends to suggest that the functionally effective representation of the positive set was located at the level of descriptors. It is also possible, however, that the composition of the positive set was functionally specified at each of several memory locations, for example, at the descriptor level as well as beyond that level.

Extant dual-representation models for memory-scanning experiments are based on the notion of familiarity processing (Atkinson & Juola, 1973, 1974; Juola, Fischler, Wood, & Atkinson, 1971; Swanson, 1974). A case for such a model might be made as follows. If, say, each time *C* is presented, the *O*-descriptor is partly activated such that the familiarity value associated with *O* is increased, then the familiarity value of *O* should tend to be higher when *C* is a member of the positive set than when *C* is not. On the usual assumptions, the higher the familiarity value of *O*, the more negative reactions to *O* should then require an extended memory search, rather than being based on fast decisions from familiarity values. Hence, when *C* is positive, negative reaction times for *O* should be lengthened.[5] The explanation works formally so far, but it may be rejected by considering the implication that correct reactions to *O* should

---

[5] A similar type of explanation was discussed by Atkinson and Juola (1973, 1974).

be slowest on those trials immediately following presentations of *C*.

The simpler assumption that a descriptor-level representation of the positive set underlay performance on nonrepetition trials in Experiment 3 is actually compatible with all the findings reported on visual confusions. In particular, when a negative stimulus letter which is selectively similar to one member of the positive set is considered in visual confusability conditions, constant negative reaction times and decreasing false alarm rates with increasing positive set size (see Table 4) can be predicted from a parallel random walk model of descriptor processing.

General Discussion

The reported series of experiments on visual recognition seems to demonstrate the occurrence of two processes of size scaling, which were tentatively identified as mental-image transformation and perceptual-scale transformation. The data suggest that (a) image and scale transformations are discriminable by their temporal courses, (b) the role of image transformation is mainly restricted to matching performance relying on visual short-term memory, (c) size-invariant pattern recognition is normally achieved by scale transformation when recognition is based on comparing stimulus patterns against visual representations in long-term memory, and (d) in special conditions, the two types of size scaling can go on in parallel, each with its own time course.

The agreement of results with expectations tends to support the initial supposition that visual pattern recognition is based on position-wise comparison of stimulus patterns with memory representations. Nevertheless, alternative conceptions of the recognition process might accommodate the present data by suitable ad hoc assumptions. Assume, for concreteness, that stimulus patterns are encoded as size-invariant feature lists (cf. Blakemore & Campbell, 1969; Milner, 1974) or structural descriptions (cf. Sutherland, 1968) before being compared with (size-invariant) memory representations. Rather than serving to establish an adequate positional correspondence between stimulus patterns and memory representations, size scaling should then be a preliminary operation in stimulus encoding; it should be necessary to adjust the perceptual procedures to the size format of the stimulus before abstracting the features or descriptions to be used for comparison against memory specifications. A perceptual reference system could thus be interpreted as a scalable reference system for size-invariant structural description, or a scale transformation could even be interpreted as a scalar tuning of size-specific feature-detecting mechanisms. Further elaboration might provide for the evidence of two different types of size scaling in visual pattern recognition.

References

Atkinson, R. C., Herrmann, D. J., & Wescourt, K. T. Search processes in recognition memory. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium.* Potomac, Md.: Erlbaum, 1974.

Atkinson, R. C., & Juola, J. F. Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV.* New York: Academic Press, 1973.

Atkinson, R. C., & Juola, J. F. Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 1): *Learning, memory, and thinking.* San Francisco: Freeman, 1974.

Bertelson, P. Serial choice reaction time as a function of response versus signal-and-response repetition. *Nature,* 1965, *206,* 217–218.

Blakemore, C., & Campbell, F. W. On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology,* 1969, *203,* 237–260.

Bundesen, C., & Larsen, A. Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance,* 1975, *1,* 214–220.

Corballis, M. C. Access to memory: An analysis of recognition times. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V.* New York: Academic Press, 1975.

Duda, R. O., & Hart, P. E. *Pattern classification and scene analysis.* New York: Wiley, 1973.

Eichelman, W. H. Stimulus and response repetition effects for naming letters at two response-stimulus intervals. *Perception & Psychophysics,* 1970, *7,* 94–96.

Hochberg, J. Attention, organization, and consciousness. In D. I. Mostofsky (Ed.), *Attention:*

Contemporary theory and analysis. New York: Appleton-Century-Crofts, 1970.

Juola, J. F., Fischler, I., Wood, C. T., & Atkinson, R. C. Recognition time for information stored in long-term memory. *Perception & Psychophysics,* 1971, *10,* 8–14.

Kornblum, S. Sequential effects in choice reaction time: A tutorial review. In S. Kornblum (Ed.), *Attention and performance IV.* New York: Academic Press, 1973.

Kristofferson, M. W. When item recognition and visual search functions are similar. *Perception & Psychophysics,* 1972, *12,* 379–384.

Milner, P. M. A model for visual shape recognition. *Psychological Review,* 1974, *81,* 521–535.

Minsky, M. Steps toward artificial intelligence. *Proceedings of the Institute of Radio Engineers,* 1961, *49,* 8–30.

Neisser, U. *Cognitive psychology.* Appleton-Century-Crofts, 1967.

Noton, D., & Stark, L. Eye movements and visual perception. *Scientific American,* 1971, *224*(6), 34–43.

Posner, M. I. Coordination of internal codes. In W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press, 1973.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monograph,* 1969, *79*(1, Pt. 2).

Pribram, K. H., Nuwer, M., & Baron, R. J. The holographic hypothesis of memory structure in brain function and perception. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2): *Measurement, psychophysics, and neural information processing.* San Francisco: Freeman, 1974.

Rabbitt, P. M. A. Repetition effects and signal classification strategies in serial choice-response tasks. *Quarterly Journal of Experimental Psychology,* 1968, *20,* 232–240.

Ross, J. Extended practice with a single-character classification task. *Perception & Psychophysics,* 1970, *8,* 276–278.

Simpson, P. J. High-speed memory scanning: Stability and generality. *Journal of Experimental Psychology,* 1972, *96,* 239–246.

Smith, M. C. Repetition effect and short-term memory. *Journal of Experimental Psychology,* 1968, *77,* 435–439.

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

Sternberg, S. Two operations in character recognition: Some evidence from reaction-time measurements. *Perception & Psychophysics,* 1967, *2,* 45–53.

Sternberg, S. Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist,* 1969, *57,* 421–457.

Sternberg, S. Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology,* 1975, *27,* 1–32.

Sutherland, N. S. Outlines of a theory of visual pattern recognition in animals and man. *Proceedings of the Royal Society, Series B,* 1968, *171,* 297–317.

Swanson, J. M. The neglected negative set. *Journal of Experimental Psychology,* 1974, *103,* 1019–1026.

Swanson, J. M., & Briggs, G. E. Information processing as a function of speed versus accuracy. *Journal of Experimental Psychology,* 1969, *81,* 223–229.

Wickelgren, W. A. Dynamics of retrieval. In D. Deutsch & J. A. Deutsch (Eds.), *Short-term memory.* New York: Academic Press, 1975.

Received February 14, 1977 ■