

COMPLEXITY EFFECTS IN FORM DETECTION¹

WILLIAM R. UTTAL and THELMA E. TUCKER
The University of Michigan, Ann Arbor, MI

(Received 8 August 1975; in revised form 15 July 1976)

Abstract—This study examines the effect of pattern complexity on the detectability of dotted target patterns presented in dotted visual noise. The effects of complexity on performance in this detection task were shown to be substantial. Discrepancies between the psychophysical scores and the prediction of a previously described autocorrelation theory of form detection (Uttal, 1975) point to specific deficiencies in the theory's sensitivity to particular classes of geometrical form.

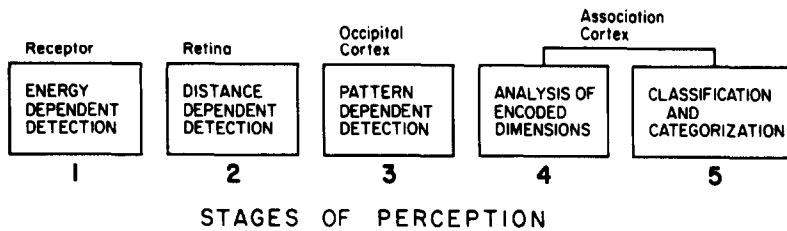
Despite the enormous interest in form perception over the past century (see, for example, the 2583 references cited in Zusne, 1970), there is still nothing that can be identified as a comprehensive and compelling theoretical description of even a subportion of this important area of human performance. A number of writers have suggested that the main difficulty in establishing a form perception theory is that we still have not identified which parameters of the stimulus are important in the perception of form.

There are two reasons for the inability to remove this deficiency in form perception theory. The first reason is on the stimulus side of the problem. Form, though certainly a dimension of stimulus variation, is unlike acoustic frequency or photic wavelength, in that it is not a unidimensional independent variable with a well-established set of physical instruments and procedures for its quantification. The second reason

is a complementary obstacle on the response side of the problem. Form "perception" is a multifaceted behavior that includes a large number of different psychological functions. Several quite different processes are often simultaneously assayed in a single experimental design and the subject, therefore, is often trying to master several perceptual tasks at once. Thus the problem of definition is great in this area. Form is ill-defined; we often do not have good control over the independent variable. Form perception is also ill-defined; we usually do not abstract a task from the total form perception process that is sufficiently pure to enable each experiment to deal with only a single component of this complex behavior.

The separation of the component subprocesses of perception so that they can be assayed by relatively independent experimental tests is an important preparatory step for the solution of the entire form perception problem. One attempt to systematically analyze perception into a series of stages or subprocesses is presented in Fig. 1. This figure depicts a model that breaks up perception into five sequential and separate stages. The first stage involves those aspects

¹ This research was supported by a research grant from the National Institute of Mental Health (1 R01 MH24016) and by an NIMH Research Scientist Award (5 K05 MH29941) to the first author.



1. Luminosity, temporal, chromatic effects on receptor response
2. Lateral summation and inhibition dependent on distance
3. Detection based on pattern organization
4. Information processing in which stimuli only serve to convey multidimensional information
5. Cognitive encoding and response linkage

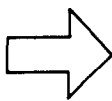


Fig. 1. Hypothetical analysis of the stages of visual perception from Uttal (1975). The arrow indicates the stage at which the present experiment is directed. Anatomical loci are suggestive of the locations at which such processing might occur, but at the higher levels these associations are speculative.

of perception sensitive to the energetics of the stimulus and the photochemistry of the receptor cells. The second stage is heavily dependent upon the lateral interconnections that occur between geometrically adjacent parts of the receptor. The physiological mechanisms underlying these two initial stages are probably located in the peripheral retina and are sensitive to stimulus geometry.

The third stage is probably mediated by a central nervous mechanism that is also highly dependent upon the geometry of the stimulus, but in a way that is quite distinct from the physiological interactions underlying the second stage. The fourth stage is one in which the geometry of the stimulus is no longer important, but rather the information dimensions borne by various aspects of the stimulus serve as cues or codes to higher levels of symbolic or semantic processing. The dimensionality of the transmitted cues, rather than the geometry of the stimulus, plays the key role at this stage of processing. The fifth and final stage involves the categorization and classification of the stimulus pattern. Implicit in this multistage model is the assumption that each stage can be explored in isolation only when the thresholds for all previous stages have been substantially exceeded and the particular task utilized as an experimental probe does not require processing at any higher level.

In the present study, attention is concentrated on Stage 3 subprocesses by the selection of a highly constrained experimental task—dotted target detection in a visually noisy dotted masking field. In this manner, a form of pattern detection can be explored that is mainly sensitive to geometry and centrally localized. Thus it transcends to a considerable degree those visual effects introduced by either receptor physics and biochemistry or by lateral interactions, on the one hand, and keeps the effects of the more complex cognitive learning, dimensional, or classification processes constant, if not minimal, on the other.

The other part of the problem, the specification and quantification of exactly what is meant by the term "form" in the study, is handled in a different manner. There is still no unidimension along which form can be ordered, and notational systems that attempt to provide a systematic algebra for form (e.g. Leeuwenberg, 1971) generally prove to be disappointing upon close examination. There are, however, instances in which points along a single dimension of form have been either ordered or reduced to simplified situations in which the ordering is more or less obvious.

In the present study, we deal with one such simplification: complexity. The purpose of this study is to explore the effect of highly constrained stimulus sets, varying in a particular kind of complexity, on what we believe to be mainly a Stage 3 form detection task. The experiments described in the present study represent a continuation of work that has been reported elsewhere (Uttal, 1975) in which the effects of a number of other dimensions of form on this Stage 3 perceptual subprocess were studied.

The study of complexity as an experimental variable in form perception theory has been of interest to psychologists for a considerable time. Chipman (1972) reviews many of the studies that have been carried out using one or another test of form perceptibility to evaluate the effects of complexity. However,

in the only previous study, to our knowledge, that used a detection task (French, 1954), complexity was varied by manipulating the number of randomly positioned dots in the target display. In the present study, the number of dots remains constant. The present study thus explores the effects on detection of a form of organization that is more precisely *complexity* rather than *numerosity*.

The autocorrelation model presented in the discussion is the contemporary stage of development of an idea that has long had appeal to students of form perception. The history of this development and its relation to other similar models is spelled out in detail in Uttal (1975).

METHOD

Subjects

Undergraduate students at The University of Michigan were used as *Ss* in the experiments to be reported in this paper. Each was paid a fixed stipend for a daily 1-hr session, and none reported any evidence of visual abnormalities, although some wore corrective lenses. All *Ss* became highly trained and served for at least one entire academic semester during which they participated in more than one part of the experiment. *Ss* were always highly familiar with the target stimuli prior to the collection of the experimental data.

Apparatus

Stimuli were generated on the face of a cathode ray oscilloscope under the control of a small digital computer. Each stimulus was composed of a dotted target pattern embedded within variable amounts of randomly positioned masking dots.

Target stimulus patterns were prepared in advance by the experimenter and then stored in the computer for recall as requested by the experimental control program. A random number algorithm was used by this control program to counterbalance the number and order of the stimuli to be presented within each *S's* hourly session. The pattern of random masking dots was recalculated after each trial so that a new mask was used in each trial. The computer also controlled all response acquisition and data analysis procedures.

The *S* viewed the CRT screen from a distance of 33 cm. The viewing region on the face of the oscilloscope on which the random masking dots were presented was 5.6° by 5.6°, while stimulus patterns were smaller, typically subtending an angle of about 3° by 3°. The luminance of each plotted dot was about 0.1 m². The compartment in which the study was run was otherwise dark and sound-proofed.

The *S's* head was constrained in a fixed viewing position by a forehead switch. Unless this forehead switch contact was closed, two hand-held response switches were inactive and no responses could be entered into the computer.

Stimuli were plotted using an intensification pulse with a duration of 8 μsec to the Z-axis input of the oscilloscope. However, the duration of each dot was determined more by the persistence of the P-15 phosphor of the oscilloscope cathode ray tube than by this electronic pulse. According to the manufacturer's specifications, the P-15 phosphor has a persistence such that the light output is reduced to 0.1 of 1% of its initial brightness after 50 μsec. The apparent persistence of this briefly illuminated spot, of course, is even further elongated by the properties of the observer's visual system. However, the total physical duration of each stimulus was defined by the number of masking and target dots involved and varied between 1 and 5 msec.

Procedure

The experimental procedure involved a two-alternative, forced-choice paradigm in which the *S* was required to specify (by depressing one or the other of the two hand-held switches) which of two sequentially presented stimuli (separated by 1.0 sec) contained one of the target patterns. Both stimuli contained a common random pattern of dotted visual masking dots. The *S* was not required to say which of the targets he detected. The nontarget stimulus contained, in addition to the common masking dot pattern, a number of randomly placed dots equal to the number in the target pattern scattered over the same 3° by 3° region in which the target pattern was placed to equate overall brightness and dot numerosity.

Following the depression of either one of the response switches, the *S* received information concerning the correctness or incorrectness of his choice by means of a plus or a minus sign plotted on the oscilloscope screen. This feedback signal, which lasted for 0.5 sec, also served as a fixation cue; there was no other fixation cue, other than the stimulus itself, present in the intertrial or interstimulus intervals. The depression of the response switch and the presentation of the feedback signal was followed automatically by the presentation of the two stimuli of the next trial. Pacing, therefore, was a function of the rate at which the *S* responded.

All performance was measured as the percentage of the total number of presentations for which the target was correctly detected. A score of 50% represents chance performance in this two-alternative task. These data were accumulated for each *S* and punched on a paper tape at the end of each session. At the end of the day, a secondary data analysis program was used to summarize the results from all *Ss*.

Stimulus patterns ranked in complexity were generously supplied to us by Susan Chipman (Chipman, 1972). The Chipman patterns, as we shall refer to them, had been arbitrarily designed and then scaled by her using a magnitude estimation procedure in which *Ss* were asked to judge complexity by assigning a number to each stimulus. The original Chipman patterns were checkerboard patterns, but, for the purposes of the present experiment, the patterns were converted to equivalent dot patterns with a constant horizontal and vertical spacing of 37.7' of visual angle. The actual target patterns used are shown in Fig. 2. Table 1 shows the complexity scores in rank order for each of the Chipman patterns.

The 30 target patterns were presented in random order on 14 successive days in which the masking dot densities were 5, 10, 15, 20, 25, 30 and 35. On the first 7 days, the masking dot densities were presented in this ascending order. On the second 7 days, the masking dot densities were presented in descending order. Data from corresponding pairs of sessions were then pooled to reduce the possibility of any sequence effect biasing the results.

Since each stimulus was presented approx 35 times to each of the three *Ss* during two sessions, each data point plotted in Fig. 3 represents a sample of approximately 210 trials. Further pooling of the data for all masking dot densities resulted in Fig. 4 in which each point represents an approximate sample size of 1470 trials.

RESULTS

The results of the experiment are shown in Fig. 3. In this figure, the horizontal axis indicates each pattern's rank of judged complexity (see Table 1), and

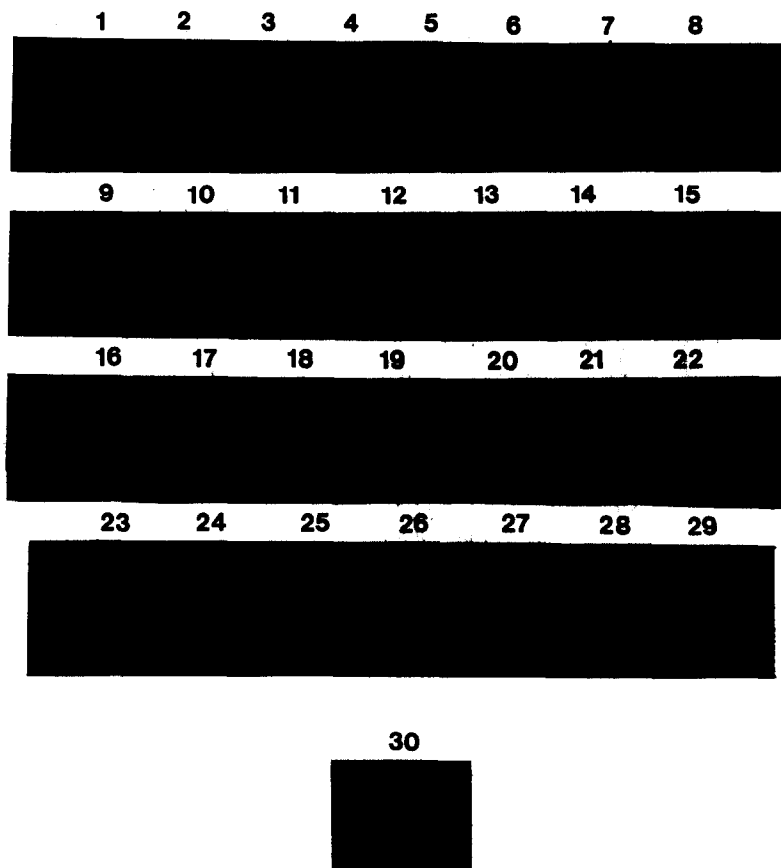


Fig. 2. The patterns used in the experiment.

Table 1. Ranking of the patterns in order of judged complexity according to Chipman (1972)

Rank order of complexity	Pattern No.	Complexity score
1	1	0.4
2	13	0.7
3	9	1.4
4	28	2.2
5	11	3.5
6	10	4.0
7	27	5.3
8	4	5.6
9	29	5.7
10	8	7.0
11	7	9.7
12	14	12.1
13	2	15.7
14	6	16.0
15	15	20.9
16	12	22.2
17	26	22.2
18	23	23.0
19	5	25.1
20	21	29.2
21	30	30.5
22	17	31.9
23	20	32.4
24	24	33.1
25	22	34.3
26	25	36.2
27	18	36.2
28	16	36.7
29	3	38.8
30	19	44.9

trend at all utilized masking dot densities. To more precisely metricize the degree of association between the rank of judged complexity (lowest complexity is ranked 1) and ranked detectability scores, a Spearman rank order correlation (ρ) of -0.89 was computed for the average scores for each pattern pooled for all densities of masking dots. This is not surprising since both the detection scores and judged complexity rankings would be expected to be highly correlated. The fact that two different tasks both rank ordered the stimuli in the same way does support the validity of the ranking and obviates the necessity to separately judge the complexity of our modified Chipman stimuli. A plot of this grand average is shown in Fig. 4. Thus we can see that as the complexity of the figure increases, the detectability drops in a manner indicating a high degree of association between the two behavioral measures of detection and complexity.

Another characteristic of this data is the relatively (compared to much of our previous work) low density of masking dots at which the performance begins to deteriorate. Some of the less detectable stimuli were seen at what must be considered to be chance levels at masking dot densities as low as 35 dots. This is a masking dot density that would have been almost completely ineffectual in masking some of the other targets that we have used in other experiments (e.g. see Uttal, 1975). On the other hand, the less complex patterns were highly resistant to masking and showed little decline in performance at the highest masking dot density used. Clearly, complexity is a powerful determinant of the relative susceptibility to masking and produces a strong effect at all and any masking dot densities.

the vertical axis indicates the percentage of the total number of trials in which it was correctly detected. The family of curves is parametric with masking dot density. Each curve exhibits a gradual decline in performance as the complexity increases and, although the data are not totally monotonic, there is a clear

DISCUSSION

The dotted target detection paradigm, used in this work and elsewhere, suggested a mathematical-neurophysiological model based upon an autocorrelation

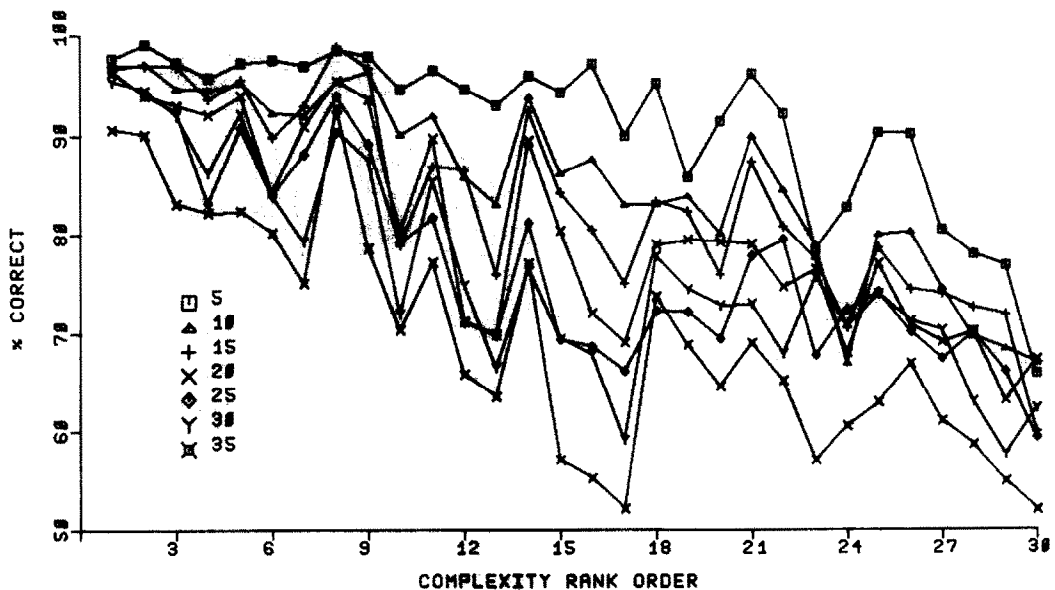


Fig. 3. The results plotted in order of increasing complexity judgments. Increases in complexity produce a decrease in detectability at all masking dot densities. The family of curves is parametric with the number of masking dots.

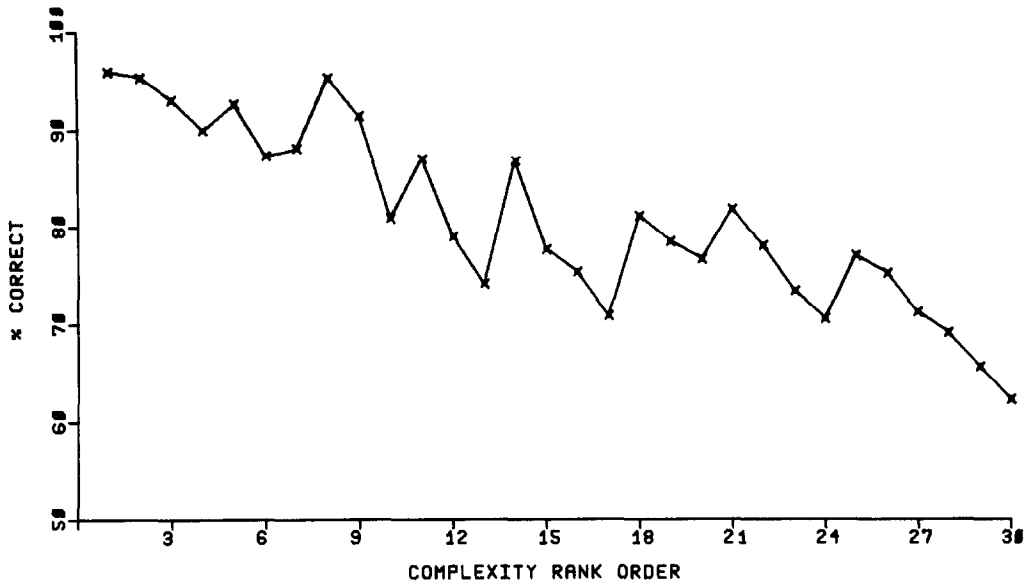


Fig. 4. The results with all data for all masking dot densities pooled and plotted as a function of complexity. This produces a rank order correlation coefficient of -0.89 .

transform presumably underlying this stage of form perception. This model has been fully developed in the previously cited monograph (Uttal, 1975), but to briefly summarize its details, it should be noted that the model is based upon only a few assumptions. First, it is assumed that Stage 3 detection processes of this sort are handled by information processing mechanisms that are specifically dependent upon the geometrical relations of the parts of the stimulus pattern. Second, it is assumed that the geometrical relations of the target stimulus may be transformed by the autocorrelation function:

$$A_c = \iint f(x, y) \cdot f(x + \Delta x, y + \Delta y) dx dy \quad (1)$$

from the x, y space to the $\Delta x, \Delta y$ autocorrelation space, where $f(x, y)$ is the original stimulus and $f(x + \Delta x, y + \Delta y)$ is a shifted replica of it. A sample of this transformation is shown in Fig. 5. The new pattern generated in the $\Delta x, \Delta y$ space, the autocorrelogram, has the pictographic geometrical qualities of the stimulus quantified in such a way that simple numerical metrics like peak height and interpeak spacing represent the original geometric properties such as straightness and organization. Therefore, a relatively simple formula can be applied to the autocorrelogram to produce a single numerical figure of merit. The simple, non-unique and relatively arbitrary formula we use is of the form:

$$F_m = \sum_{i=1}^I \sum_{j=1}^N \frac{(A_i \cdot A_j)}{d_{ij} N} \quad (i \neq j) \quad (2)$$

where A_i and A_j are the amplitudes of a pair of peaks, d_{ij} is the distance between the two peaks and N is the number of peaks in the transformed space of the autocorrelogram. This figure of merit serves the important role of providing a single numerical estimate to us, the experimenters, of the geometrical properties of the stimulus pattern but it is not actually a part of the process we believe is carried out in the brain.

There, the state of the autocorrelogram itself is the end product.

The main purpose of the following discussion will be to apply the model to the results of the present

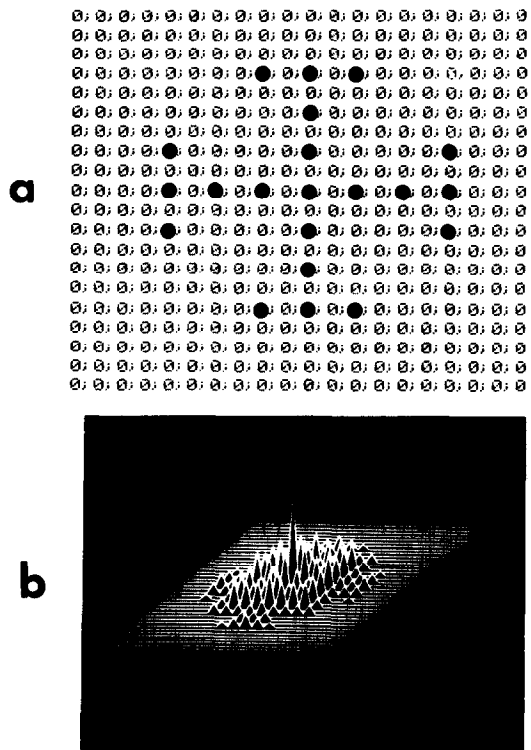


Fig. 5. (a) Sample of a pattern (a Maltese cross, not part of the stimulus set of this experiment) that is fed into the autocorrelation program. (b) Graphic display of the autocorrelogram of the pattern in (a). The amplitude of the peaks and the interpeak spacing in the $\Delta x, \Delta y$ space of the resulting autocorrelogram are functionally dependent on the geometry of the input pattern.

study. To apply the model, a simulated form of each of the stimuli shown in Fig. 2 was prepared for autocorrelation transformation (carried out on the same minicomputer on which the experiment was run) in the form of matrices like that shown in Fig. 5a. The computer then applies the autocorrelation transform represented by equation (1) to produce a set of patterns like that shown in Fig. 5b. The output of the first program—the autocorrelogram—is the input for a second program that produces the numerical figure of merit by evaluating equation (2).

Table 2 indicates respective rank order (highest figure of merit is ranked 1), the pattern numbers, and the figures of merit for the Chipman stimuli used in this experiment. Figure 6 plots the percentage of correctly detected patterns against the rank order of the calculated figures of merit. A Spearman rank order correlation of -0.52 between the two ranked sets was also computed.

It is obvious that the prediction in this case is less satisfactory than the prediction obtained from judged complexity. The model predicts much less well (-0.52 compared to -0.89) than does the set of complexity estimates provided by Chipman.

However, the discrepancies are few and fairly specific and thus indicate the particular weaknesses of the theory. The major discrepancies between the predictions of the model and the psychophysical results are for items rank ordered 22 and 29 (patterns 7 and 4), which are underranked by the figure of merit expression, and for items rank ordered 5 and 8 (patterns 2 and 26), which are overranked by the figure of merit expression. With regard to patterns 4 and 7, it appears that the expression for the figure of merit is unable to deal properly with forms that are composed of lines placed near the periphery of the field. The model underestimates the ability of the subject to see patterns separated by several degrees. Patterns 2 and 26, on the other hand, seem to possess some property that spuriously raises their figure of merit but that does not help in the psychophysical detection

task. The fact that these two patterns are orderly groups of dots with diagonal organization probably explains this discrepancy. These two kinds of discrepancies may, however, provide clues for the improvement of the expression for the figure of merit.

Table 2. Ranking of the patterns in order of the figure of merit [equation (2)]

Rank order of figure of merit [equation (2)]	Pattern No.	Figure of merit
1	1	9896
2	13	9595
3	11	9437
4	9	6680
5	2	6503
6	10	6480
7	28	5429
8	26	5410
9	27	5278
10	5	5268
11	30	5226
12	29	4656
13	22	4574
14	6	4489
15	15	3967
16	21	3723
17	23	3661
18	20	3658
19	3	3593
20	18	3478
21	12	3474
22	7	3449
23	16	3387
24	17	3338
25	25	3219
26	14	3170
27	8	3022
28	24	2929
29	4	2854
30	19	2695

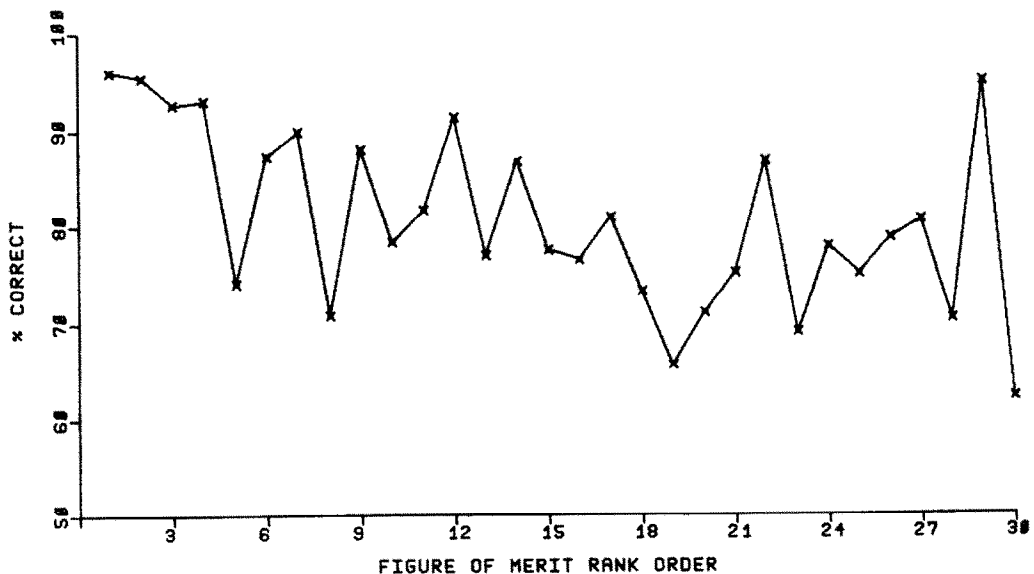


Fig. 6. The pooled results plotted in order of the figure of merit of the patterns as calculated with equation (2). This produces a rank order correlation coefficient of -0.52 .

It is interesting to note, but of course no solution to the problem, that if these four discrepant patterns are removed from the set of stimuli then the rank order correlation of the predictions by the model and the psychophysical results would improve from -0.52 to -0.81 , a value quite close to that obtained with the judgments of complexity. Considering that the model is purely based upon only a few precisely defined parameters of the autocorrelogram, while the complexity judgment is a subjective ranking based upon a large number of unspecified parameters, we believe the progress made so far is a relatively satisfactory achievement, and a step in the direction of a more comprehensive theory of form detection.

The autocorrelation theory explanation of this restricted aspect of visual perception—form detection—has a number of implications that we believe to be especially important.

(1) The autocorrelation model suggests that form detection can be carried out entirely by processes performed solely on the input stimulus. It is an alternative to those theories that invoke some sort of a template or reference pattern.

(2) The model suggests that homogeneous networks of unspecialized neurons can exhibit much the same selectivity to stimuli, on the basis of a general form of information processing, as highly specialized neurons do by narrowly defined kinds of feature filtering.

(3) The model suggests that a reasonable explanation is possible for the commonly observed fact that people seem to be more sensitive to the pattern of arrangement of the parts of a stimulus than to the nature of the parts. In spite of the widespread popularity of feature filtering theories in contemporary psychology, it just may be that the Gestalt psychologists were more correct in their emphasis on global patterns of organization than has recently been thought to be the case. One contribution this study makes is to provide a possible quantitative link between the Gestalt and Information Processing perspectives.

REFERENCES

- Chipman S. F. (1972) The perceived complexity of visual patterns. Doctoral thesis, Harvard University.
- French R. S. (1954) Pattern recognition in the presence of visual noise. *J. exp. Psychol.* **47**(1), 27–31.
- Leeuwenberg E. L. J. (1971) A perceptual coding language for visual and auditory patterns. *Am. J. Psychol.* **84**(3), 307–349.
- Uttal W. R. (1975) *An Autocorrelation Theory of Form Detection*. Lawrence Erlbaum, Hillsdale, N.J.
- Uttal W. R. (1969) Masking of alphabetic character recognition by dynamic visual noise (DVN). *Percept. Psychophys.* **6**, 121–128.
- Zusne L. (1970) *Visual Perception of Form*. Academic Press, New York.