Visual attention and perceptual grouping

MERCEDES BARCHILON BEN-AV and DOV SAGI
The Weizmann Institute of Science, Rehovot, Israel

and

JOCHEN BRAUN
California Institute of Technology, Pasadena, California

Perceptual organization is thought to involve an analysis of both *textural discontinuities* and *perceptual grouping*. In earlier work, we found that textural discontinuities were detected normally even when visual attention was engaged elsewhere. Here we report how perceptual grouping is affected when visual attention is engaged by a concurrent visual task. To elicit perceptual grouping, we used the Gestalt demonstrations of grouping on the basis of proximity and of similarity. Four tasks were investigated, some requiring the observer to discriminate between horizontal and vertical grouping, and some requiring the observer to merely detect the presence or absence of grouping. Visual attention was engaged at the center of the display by a form identification task. The detection of a textural discontinuity served as a control task. Concurrent form identification conflicted with all four grouping tasks, resulting in a significant reduction of grouping performance in each case. No performance reduction was observed when either form identification or grouping discrimination was combined with the detection of a textural discontinuity. These results suggest that perceptual grouping and form identification compete for visual attention, whereas the detection of a textural discontinuity does not.

The principles that govern perceptual organization form one of the most enduring subjects in vision research. One form of perceptual organization was studied by the Gestalt psychologists, who formulated laws of *proximity, similarity, continuity, common fate,* and *closure* to account for the phenomenon of *perceptual grouping,* which is how they termed the visual system's tendency to aggregate discrete stimulus elements into larger wholes (Koffka, 1935; Wertheimer, 1923). Another form of perceptual organization arises from the visual system's ability to "fast and effortlessly" segment a visual scene on the basis of texture, a phenomenon that has been termed *texture segregation* (Julesz, 1981; Watt, 1991). Aggregative and segregative forms of perceptual organization are believed to be related (Beck, 1982; Treisman, 1982), and they are both thought to contribute to the separation of figure from ground, thus preparing the way for visual recognition.

Although the events leading to visual recognition are little understood, the existence of two complementary types of visual processes is generally recognized, and the two types have been awarded the appellation *preattentive* and *attentive,* respectively (Julesz, 1981; Neisser, 1967;

Treisman & Gelade, 1980). In our view, these terms are apt, but perhaps too weighty in their implications. A typical preattentive process is reflexively activated by visual stimulation, whereas a typical attentive process depends not only on stimulation but also on the observer's adoption of a certain perceptual stance or disposition; it depends on visual attention. Since the adoption of one attentive stance or disposition seems to preclude, or at least obstruct, the simultaneous adoption of another, an alternative metaphor is to say that attentive processes are limited by the availability of a perceptual *resource*. Of preattentive processes, which are independent of the observer's attentive stance or disposition, one may say that they are *resource free*.

The subject of this paper is the relationship between perceptual grouping, on the one hand, and visual attention, on the other. Does perceptual grouping require full access to the resources of visual attention? Or does grouping occur even when attentive resources are unavailable (e.g., because they are committed elsewhere)? In earlier work on texture segregation, we found that this perceptual process is largely independent of the availability of the resources of visual attention (Braun & Sagi, 1990, 1991), and this result is once more confirmed here. Given the apparent similarity between perceptual grouping and texture segregation (Beck, 1982; Treisman, 1982), we wanted to investigate whether perceptual grouping is as independent of attentive resources as texture segregation appears to be.

As has been known at least since the demonstrations of the Gestalt psychologists (Koffka, 1935), the perceptual organization imposed by grouping processes can be

compelling. More recent demonstrations of the grouping phenomenon are Glass patterns (Glass, 1969; Glass & Perez, 1973), in which the collinearity of numerous randomly placed pairs of dots causes the formation of perceptual "grainlines" and a larger scale organization emerges from the collective of "grainlines." Another striking instance of grouping is the perceptual clusters formed in displays in which the convex- and concave-appearing tokens used to study shape-from-shading (Ramachandran, 1988) are mixed. Grouping processes may also account, at least in part, for numerous demonstrations involving motion coherence (e.g., Ramachandran & Anstis, 1985; Ullman, 1979a), as in the sudden disambiguation of a monocularly viewed three-dimensional object when the object begins to rotate (Ullman, 1979b) (cf. the Gestalt law of common fate).

One of the most extensive and continuous studies of grouping processes has been carried out by Beck (1966, 1967, 1972; Beck, Sutter, & Ivry, 1987; see also Olson & Attneave, 1970). He investigated the stimulus parameters that facilitate grouping on the basis of similarity and found that items of similar shape but different orientation are difficult to group together, whereas items differing in the arrangement, but not the orientation, of their constituent lines are grouped readily. Beck's findings indicate that grouping is governed by lower level mechanisms, which moreover possess characteristics similar to the mechanisms implicated in texture segregation (Julesz, 1981, 1986; Nothdurft, 1985).

Accordingly, current theories of perceptual organization tend to view perceptual grouping and texture segregation as closely related processes, operating either concurrently or in close succession (Beck, 1982; Beck, Prazdny, & Rosenfeld, 1983; Treisman, 1982; Treisman & Gormican, 1988). In theories that assume succession, grouping precedes segregation: Beck et al. (1983) suggest that segregation can occur between the global features that emerge through grouping of local features, and Julesz (1986) assumes that textural differences are computed between local texture elements, which in turn result from a grouping of nearby textons (line elements).

The relationship between visual attention and the processes underlying perceptual organization and figure-ground separation has for the most part eluded experimental efforts, but it has nevertheless been a frequent subject of theoretical considerations. Theories of early vision that distinguish between preattentive and attentive processes tend to place both texture segregation and perceptual grouping on the preattentive side (Julesz, 1986; Treisman, 1982). Other authors prefer to distinguish between situations in which attention is dispersed across the entire field of view and situations in which attention is focused or concentrated in one part (Beck & Ambler, 1973; Treisman & Gormican, 1988). These authors tend to assume that both texture segregation and perceptual grouping only require dispersed attention.

In previous work, we found that the detection or localization of a textural singularity is carried out normally even when visual attention is focused at a distant location in the field of view (the location of the target of a concurrent form identification task), suggesting that at least some tasks based on segregative mechanisms pose little or no demand for attentive resources (Braun & Sagi, 1990, 1991). More generally, we suspect that local perceptual salience within homogeneous and dense stimulus textures is attenuated by resource-inexpensive (or -free) processes everywhere except in the immediate vicinity of texture borders, thus permitting the resource-inexpensive (or -free) detection and localization of textural borders and singularities (Rubenstein & Sagi, 1990; Sagi, 1991).

As a result of this earlier work on texture segregation, we wondered whether it would be possible to demonstrate perceptual grouping in parts of the field of view from which attentive resources have been completely or mostly withdrawn. If possible, such a demonstration would give additional impetus to efforts to model the processes underlying perceptual grouping and texture segregation in similar ways—that is, as a feed forward cascade of linear filters with some nonlinearities interposed between successive stages (Beck, Graham, & Sutter, 1991; Fogel & Sagi, 1989; Malik & Perona, 1990; Rubenstein & Sagi, 1990). The opposite outcome—failure to observe perceptual grouping when attentive resources are absent or scarce—would suggest a fundamental difference between aggregative and segregative contributions to perceptual organization, and would place aggregative mechanisms at a later stage of visual processing than segregative ones.

To address the issue of whether or not perceptual grouping presents a demand on visual attention, we needed an experimental situation that would capture the essence of the grouping phenomenon and that would at the same time lend itself to the formulation of an objective psychophysical task. These requirements were met by some of the Gestalt demonstrations, in particular those involving proximity and similarity. Accordingly, our array identification tasks were based on these demonstrations.

In order to assess texture segregation, we used the detection of a one-element singularity in a dense background texture (Braun & Sagi, 1990, 1991). Tasks that involve identifying a shape defined by the boundaries of a foreground texture are considered elsewhere (Karni & Sagi, 1991). To engage the attentive resources of the observer, we made use of a form identification task, a choice for which there is ample precedent (Bergen & Julesz, 1983; Braun & Sagi, 1991; Kröse & Julesz, 1989).

Equipped with one task assumed to reflect perceptual grouping and another assumed to engage attentive resources, we asked observers to carry out both tasks concurrently. For each task, we compared the level of performance observed under the dual-task condition with the optimal level of performance obtained under a single-task condition. We expected this comparison to reveal whether or not there is competition for attentive resources (Norman & Bobrow, 1975; Sperling & Dosher, 1986) between the grouping task and the task meant to engage attention.

We investigated four tasks assumed to reflect perceptual grouping and observed a high degree of competition for attentive resources in all four cases. This suggests not
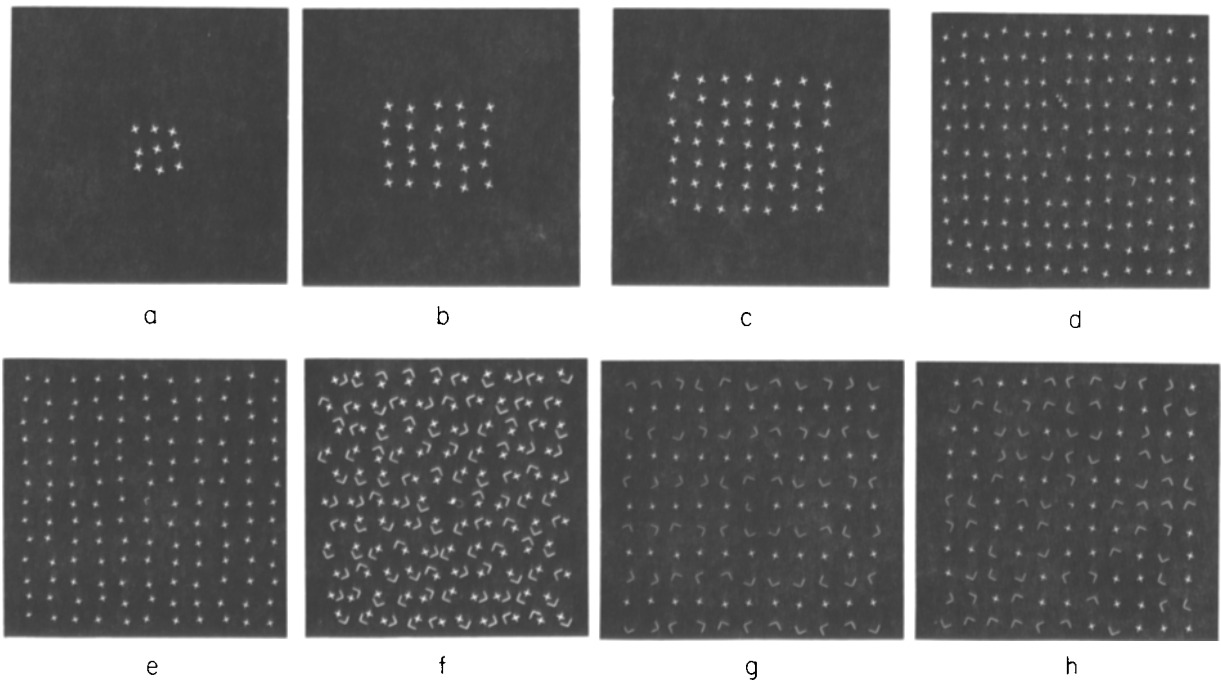
Figure 1. Stimulus and mask arrays. (a, b, and c) The three smaller sizes of stimulus arrays used in Experiment 1: 3×3, 5×5, and 7×7, respectively. The fourth size of array (11×13) used in Experiment 1 is not shown. (d, e) Possible stimulus arrays for Experiments 2, 3, and 4. (d) A T as the central element, an L present in the array of +s, and horizontal grouping in the array. (e) A Γ as the central element, no L present, and vertical grouping in the array of +s. (f) Mask array for Experiments 2, 3, 4, 5, and 6. (g) Stimulus array for Experiment 6; grouping on the basis of element similarity (in the horizontal direction). (h) Stimulus array for Experiment 6; absence of grouping. The stimuli for Experiment 5 are not shown.

only that perceptual grouping requires visual attention, but also that perceptual grouping extends to a higher level of visual processing than the related, preattentive process of texture segregation.

METHOD

Observers

Five practiced observers participated in the experiments. Three of them (S.W., R.K., and H.S.) were paid high school students and were naive as to the purpose of the study. The remaining observers (M.B., J.B.) were two of the authors. All enjoyed normal or corrected-to-normal visual acuity.

Apparatus

The stimuli were presented in a dark environment on a Hewlett-Packard 1310B oscilloscope (P31 phosphor). The oscilloscope was driven by custom-designed hardware (Smikt, 1989), which allowed real-time control of the stimulus properties. This image-generating system was controlled by a Sun 3/160 workstation. The screen resolution was 1,024×1,024 pixels, and a viewing distance of approximately 84 cm resulted in a display subtending approximately 15°×15° of visual angle.

Stimulus Patterns

The stimuli for Experiments 2-4 consisted of discrete pattern elements that were arranged as an array of either 11×13 or 13×11 rows and columns (Figures 1d, 1e). All array elements were randomly rotated, and, with one or two exceptions (see below), all were +s. The difference in the number of rows and columns was compensated by an opposite difference in the spacing between rows

and columns, respectively, so that the area occupied by the entire array was almost exactly square (16.5°×16.5°). The mean separation of array elements was 1.25° of visual angle along the more densely populated dimension (horizontal for the 11×13 and vertical for the 13×11 array). Along the other dimension, mean element separation was larger, namely, 1.5° of visual angle. The overall appearance of the array was characterized by the perceptual clustering of array elements along the more densely populated axis (proximity grouping). Specifically, the 11×13 array was organized perceptually into horizontal and the 13×11 array into vertical clusters. Accordingly, we will sometimes speak of horizontal and vertical array types (Figures 1a-1e).

In Experiment 1, arrays of four different sizes were used: 9 (3×3), 25 (5×5), 49 (7×7), and 143 (11×13) elements, corresponding to 4°×4°, 7°×7°, 9°×9°, and 15°×15° of solid visual angle, respectively. Mean element separations were those specified earlier.

In Experiments 5 and 6, different stimulus geometries were sometimes required, and therefore other mean element spacings were sometimes used: spacings of 1.88° and 1.25° produced arrays of 9×13 or 13×9 elements, spacings of 1.5° and 1.5° produced arrays of 11×11 elements, and spacings of 1.88° and 1.88° resulted in arrays of 9×9 elements. In Experiment 5, all elements other than the center element were +s. In Experiment 6, both +s and Ls were used (Figures 1g, 1h). Table 1 lists the various alternatives of array parameters employed in each experiment.

The complete set of array elements comprised +, L, a T (mirror reflection of the Hebrew letter daleth), and a Γ (mirror reflection of the Hebrew letter resh). The + and L elements were larger and consisted of line elements measuring 0.55° in length, while the T and Γ elements were smaller, containing line elements of 0.37° length. One element was displayed at every one position of

Table 1

Experiment	Horizontal Mean Separation, d_h	Vertical Mean Separation, d_v	Ratio d_h/d_v	Array size
1	1.25°/1.5°	1.5/1.25°	0.83/1.2	3×3
1	1.25°/1.5°	1.5/1.25°	0.83/1.2	5×5
1	1.25°/1.5°	1.5/1.25°	0.83/1.2	7×7
1, 2, 3, 4	1.25°/1.5°	1.5/1.25°	0.83/1.2	11×13/13×11
5	1.25°/1.5°/1.5°	1.5°/1.25°/1.5°	0.83/1.2/1.0	11×13/13×11/11×11
5	1.25°/1.88°/1.88°	1.88°/1.25°/1.88°	0.67/1.5/1.0	9×13/13×9/9×9
6	1.5°	1.5°	1.0	11×11

the array (positions being individuated as row n, column m). The element at each position was rotated randomly and displaced (jittered) from its nominal position by some amount d_x and d_y, where d_x and d_y were chosen randomly and independently from the interval $-0.188° < d_{x,y} < 0.188°$.

Visual Tasks

The array identification tasks used here are based on the Gestalt demonstrations of proximity and similarity grouping (Koffka, 1935). In each array type, elements are organized perceptually into elongated groups of horizontal or vertical orientation. The perceived elongated groups extend over most of the width and height of the display, and their discriminability grows with increasing image size (see Experiment 1). Since physical properties of the display control the strength and orientation of the perceptual organization, alternative display types can be used to pose an objective discrimination task. We assume that our array discrimination tasks are carried out on the basis of perceptual grouping, but of course we cannot simply rule out the possibility that observers employ other visual cues unrelated to grouping. To address this issue, we designed Experiment 1 to examine more closely the type of cue used in our first array identification task.

Array identification 1: Discrimination of proximity grouping (Experiments 1-4). The presentation of either vertical ($d_h/d_v = 1.2$) or horizontal ($d_h/d_v = 0.8$) array types, randomly and with equal probability, permitted us to ask observers to discriminate between these forms of organization. The observers were made aware of the two array geometries and the resulting bias in perceptual grouping, and they were instructed to report the perceptual organization of the stimulus array into either horizontal rows or vertical columns. Of course, in principle, the two array geometries are discriminable also on other grounds, without recourse to grouping.

Array identification 2: Detection of proximity grouping (Experiment 5). In this experiment, vertical ($d_h/d_v > 1.0$), horizontal ($d_h/d_v < 1.0$), and neutral ($d_h/d_v = 1.0$) array types were used, with respective probabilities of 25%, 25%, and 50%. This allowed us to pose a somewhat different task, which one might term a detection of grouping: observers were instructed to report either the presence of a clear perceptual organization (be it horizontal or vertical) or its absence. For Experiment 5A, milder differences in proximity were employed (11×13, 13×11, and 11×11 elements, $d_h/d_v = 0.8/1.2/1.0$), whereas for Experiment 5B, stronger differences in proximity were used (9×13, 13×9, and 9×9 elements, $d_h/d_v = 0.66/1.0/1.5$).

Array identification 3: Detection of similarity grouping (Experiment 6). This task was used to assess grouping on the basis of element similarity rather than element proximity. A square array of 11×11 elements was used with mean element separations of $d_h = d_v = 1.5°$ pixels. Vertical, horizontal, and neutral array types were formed by an appropriate distribution of +s and Ls across the array. In the vertical array type (25% of trials), +s and Ls formed alternating columns; in the horizontal array type (25% of trials), +s and Ls formed alternating rows; and in the neutral array type (50% of trials), +s and Ls were distributed in random fashion. As before, the observers were instructed to report either

the presence of a clear perceptual organization (vertical or horizontal) in the distribution of +s and Ls, or the absence of such an organization (neutral).

Singularity detection (Experiments 2-5). As a way to assess perceptual texture segregation, we chose the detection of a one-element singularity in a dense background texture. We view the detection of such a singularity as a limiting case of the segregation between a foreground and a background texture, in which the number of foreground elements is one. Increasing the number of foreground elements would merely serve to decrease the perceptual difficulty of this task. As a singular element we used an L, embedded in the array of +s. Segregation between textures composed of Ls and +s is considered to be fast and effortless compared with, say, segregation between Ls and Ts, which is thought to require scrutiny (Bergen & Julesz, 1983; Gurnsey & Browse, 1987; Julesz, 1981).

Accordingly, on half of the trials, a single randomly rotated L appeared within the array of +s. The components (line elements) of the L were identical to the components of the +s (only their relative positions were different). The L was restricted to one of the 24 array positions between 3.75° and 5.86° of eccentricity—that is, to the third concentric shell around the center of the array. The intermittent presence of this L allowed us to pose a *singularity detection* task, in which observers reported either the presence or the absence of an L from the stimulus array. In earlier work (Braun & Sagi, 1990, 1991, 1992), we have used textural singularities based on element orientation rather than element shape. Those stimuli had elements with continuous luminance distribution (2-D Gabor functions) rather than the discrete luminance distributions used here.

Form identification (Experiments 2-6). Form identification tasks in general, and letter identification tasks in particular, are thought to represent a significant demand for the attentive resources of an observer. The identification of a randomly rotated T or L, which is presented briefly (on the order of 50-100 msec) and then masked, has been used previously for the purpose of detaining visual attention and of reducing the availability of attentive resources for any concurrent, second visual tasks (Braun & Sagi, 1991; Kröse & Julesz, 1989).

In order to estimate the length of time for which a T/L identification affects availability of resources for any second task, one can measure the additional presentation time (stimulus onset asynchrony, or SOA) that an observer requires in order to perform a second T/L identification elsewhere in the field of view (Braun & Sagi, 1991). The value obtained in this way turns out to be on the order of 50-100 msec and, moreover, turns out to be roughly equal to the presentation time required by the first T/L identification. This outcome suggests that execution of a T/L identification engages attentive resources during roughly the entire presentation time. Similar conclusions regarding the temporal demand for attentive resources presented by a letter identification task were reached by Saarinen and Julesz (1991).

The question as to what fraction of attentive resources is taken up by a T/L identification is more difficult to answer. Taking a naive view, one might suppose that a T/L identification takes up

more than half of the available resources, since the remaining fraction of resources apparently does not sustain execution of a second T/L identification (on the assumption that both tasks need equal amounts of resources). Although the validity of this calculation is debatable, we feel that the result—that a T/L identification task engages between 50% and 100% of attentive resources during most of the presentation time—is probably not too far from the truth.

In the present study, either a Γ or a T was presented, randomly rotated, at the center of the stimulus array. The observers were required to report which letter (Γ or T) had occurred at the array center. The slight departure in shape from an L and a T increased the perceptual difficulty of the identification task (i.e., increased the SOA value required to achieve a given performance level) and permitted us to match the difficulty of our form identification task more closely to the respective difficulties of our array identification and singularity detection tasks.

Mask Patterns

To ensure that all relevant aspects of the stimulus pattern were masked effectively, a somewhat intricate mask array was generated (Figure 1f). The elements of the mask array combined the various alternative elements present in the stimulus array: the central mask element combined T and Γ, and other mask elements combined + and L. Perceptual grouping of the mask array was along the first or second diagonal, rather than horizontal or vertical. Preliminary testing with 3 subjects (M.B., J.B., and S.W.) verified that the mask array was comparably effective for the three main tasks: the average performance was 55%, 62%, and 58% correct at an SOA of 30 msec and 90%, 92%, and 83% correct at an SOA of 90 msec (form identification, array identification, and singularity detection, respectively).

Although the stimulus and mask were presented in rapid sequence, no percept of global apparent motion was observed. Accordingly, observers could not have identified the stimulus array on the basis of a motion percept. Mask arrays with horizontal or vertical grouping (which were not used) did give rise to a vivid motion percept. Although we did not pursue the issue, it was our impression that a motion percept resulted whenever the respective perceptual organizations of stimulus and mask patterns were similar (e.g., both were organized horizontally, or both were organized vertically).

The mask array was usually based on a square array of 10×11 (or 11×10) elements, with mean element separations of 1.5° in both dimensions. This array was then altered by sliding even and odd rows (or columns) past each other by 0.75°, aligning the array elements of each row (or column) *between* those of its neighbors. The total area occupied by the resulting array thus came to be approximately square (16.5° × 15.75°). Perceptually, the mask array tended to assume a diagonal organization.

With the exception of the center location, the same element was used at all locations of the mask array. This element consisted of a + *and* an L, slightly displaced with respect to each other (see Figure 1f). All mask elements of this type were rotated randomly, as well as jittered from their nominal positions. At the center location, an element in the shape of a ⫪ (the superposition of a T and a Γ) was used. This mask element was rotated in the same manner as the central element of the stimulus array. Matching orientations between the central element of the stimulus array and the central element of the mask array increased the strength of the mask for the form identification task.

This describes the mask pattern used for Experiments 2-4, as well as for Experiments 5A and 6. For Experiment 1, the central element was not different from all other elements of the mask array. For Experiment 5B, a mask array of 8×9 (or 9×8) elements was generated with mean element separations of 100 pixels (1.88°) and a relative displacement of 50 pixels (0.94°) between alternating rows or columns.

Masking, Visible Persistence, and Concurrent Tasks

The use of perceptual masking in a concurrent task paradigm is motivated by assumptions that perhaps deserve explicit mention.

A perceptual masking pattern is thought to act at the level of *visible persistence*, rather than *informational persistence* (Coltheart, 1980; Irwin & Yeomans, 1986). A characteristic of visible persistence is that it depends on intensity and duration of the stimulation, and that it decays within 100-300 msec of stimulus onset or offset.[1] Informational persistence is relatively independent of the parameters of stimulation and is thought to last longer. When perceptual masking is employed, visible persistence can be assumed to be a monotonically increasing function of SOA.

Accordingly, an effective masking pattern acting at the level of visible persistence is expected to permit perfect performance when presented 300 msec or more after stimulus onset or offset, but not more than chance performance when shown simultaneously with the stimulus. All combinations of stimulus and mask used in the present study approach chance performance for SOAs below 30 msec and perfect performance for SOAs above 150 msec, consistent with the assumption that masking limits visible persistence, rather than informational persistence, in each case.

In general, there are a number of reasons why a concurrent task paradigm can produce a conflict between tasks. However, for the tasks and task combinations investigated here, concurrent performance approaches the 100% correct level *when no masking is used* (visible persistence 100-300 msec). Accordingly, the main cause for interference between the tasks studied here would seem to be limited visible persistence.

Quantitative support for this assumption is supplied by earlier studies (Adini & Sagi, 1992; Braun & Sagi, 1991), in which we investigated the conflict between two concurrent form identification tasks as a function of visible persistence time (i.e., SOA). Substantial conflict was limited to SOAs below 150 msec. In fact, the SOA values required for each task (75% correct level) were roughly additive: 62 ± 3 msec for one (foveal), 100 ± 4 msec for the other (peripheral), and 145 ± 8 msec for one followed by the other (Braun & Sagi, 1991, Table 1). Other studies showed that the conflict between two form identification tasks disappears when the relevant information is presented sequentially, rather than concurrently (Sagi & Julesz, 1985b; Saarinen & Julesz, 1991).

Given that the visual tasks studied here interfere only when visible persistence is limited, the cause of interference must be sought in visual processes that require extended access to visible stimulus information. The most probable candidate for such a process is visual attention, for it is well known that competition for attentive resources intensifies when presentation time is limited (Kahnemann, 1973; Norman & Bobrow, 1975).

Procedure

Because we were interested in finding out how well observers would carry out various combinations of the three tasks posed by the stimulus, we conducted experiments under double-task conditions (Braun & Sagi, 1990), in which the observers attempted to perform two tasks concurrently, making simultaneous visual efforts with respect to two aspects of the stimulus pattern. To compare, we conducted experiments under single-task conditions, in which observers performed only one task and concentrated on one aspect of the stimulus. We emphasize that the double- and single-task conditions had identical stimulus and mask patterns.

Under double-task conditions, the two tasks involved were *ranked*—the observer was told to treat one task as primary and the other as secondary. This procedure tended to insulate the primary task from any performance decrement under double-task conditions, and to restrict any such decrement to the secondary task.

Data were collected in blocks of 50 trials. Each trial was preceded by a fixation mark at the center of the display. Once the ob-

server had indicated readiness by pressing the space bar on a standard keyboard, the trial sequence was initiated. This sequence consisted of a dark interval of randomly variable duration (from 20 to 140 msec), the stimulus presentation (10 msec), a dark interval of fixed duration (20 to 200 msec), and the mask presentation (100 msec). The rapidness of this sequence prevented a second eye fixation. The visual availability of the stimulus pattern was controlled by the duration of the second dark interval, the length of which determined the SOA (i.e., the interval between stimulus and mask onset).

The response of the observer consisted in typing 0 or 1 on the keyboard. Under single-task conditions, only one response was required, whereas two sequential responses were collected under double-task conditions. The order of responses was fixed, the first response pertaining to one task, and the second to another. Errors elicited feedback from the terminal bell.

Each session lasted approximately 1 h and comprised two or three periods devoted to one particular condition (such as form identification alone, or form identification and singularity detection together). Between periods, the conditions varied in no particular order. Each period devoted to one condition began with blocks of trials at the longest SOAs (110–150 msec) and ended with blocks of trials at the shortest SOAs (30–50 msec). Throughout the period, SOA values were sometimes reduced and sometimes repeated, but they were never increased.

Statistics

As mentioned earlier, a number of different tasks (form identification, singularity detection, array identification, second array identification, and third array identification), conditions (double and single task), and SOA values (30–150 msec) were investigated. For every task, condition, and SOA, several blocks of 50 trials were conducted. For every block, performance was obtained in terms of percent correct, averaged over positive and negative trials. A mean performance and its variance were obtained by averaging over all appropriate blocks and by computing the associated standard deviation and standard error values. These data were organized into psychometric functions specifying, for every task and condition, the dependence of performance on SOA.

To allow a quantitative comparison of the psychometric functions obtained for different tasks and/or conditions, every psychometric function was characterized in terms of a unique parameter, T, which was defined as:

$$T = \sum_{i=0}^{n} \frac{P_\infty - P_{(t_i)}}{P_\infty - P_0} \Delta t,$$

where the index i runs over all sampled SOA values t_i, $P_{(t_i)}$ is the performance (range $[P_0, P_\infty]$) at the SOA value t_i, Δ_t is the temporal spacing between the t_i, and P_∞ and P_0 are perfect performance and chance performance, respectively. As a consequence of this definition, the numerical value of the parameter T is monotonically related to the SOA value at threshold: as performance improves, the values for both threshold SOA and the parameter T decrease.

The main advantage of using the parameter T to characterize psychometric functions lies in the fact that its standard error can be computed in a straightforward manner on the basis of the observed standard errors of the $P_{(t_i)}$. This enables us to judge the statistical significance of the difference between two psychometric functions. To do so, we assume that two psychometric functions are significantly different if and only if the two associated T values are significantly different—that is, if ΔT is significantly different from zero.

RESULTS

Experiment 1: Integration Range for Discrimination of Proximity Grouping

Our first array identification task consists of discriminating between two alternative array types that are distinguished by slightly different horizontal (d_h) and vertical (d_v) mean separations of neighboring array elements. Perceptually, the array type with $d_h/d_v = 0.8$ is grouped in a horizontal fashion, whereas the array type with $d_h/d_v = 1.2$ is organized in a vertical manner (Figure 1; see also the Method section).

One can conceive of several visual strategies on which an observer could base the performance of this task. According to signal detection theory, each strategy would have the observer monitor one particular internal mechanism with an output signal that carries information about array type (the value of d_h/d_v). One plausible mechanism of this type is perceptual grouping, and, because of the sensitivity of grouping processes to proximity (Koffka, 1935), the grouping signal would be quite informative about array type. Equally plausible is any mechanism that merely signals the precise separation of neighboring array elements—in principle even the separation of just two neighboring elements would suffice—and that would therefore permit array identification. Other strategies are conceivable as well.

To learn more about the actual visual strategy used to carry out the array identification task, we determined the integration range of this task—the number of array elements at which performance saturates. As the considerations above suggest, this integration range may comprise as few as two or three elements—if array type is judged on the basis of the separation of a single pair of neighboring array elements—or as many as all elements in the array—if the judgment relies on processes that result in the perceptual organization of the entire array.

To determine integration range, we compared performance for four different array sizes: 9 (3×3), 25 (5×5), 49 (7×7), and 143 (11×13) elements, corresponding to 4°×4°, 7°×7°, 9°×9°, and 15°×15° of solid visual angle, respectively. The stimuli for this experiment, which, unlike the stimuli for subsequent experiments, were composed exclusively of + s, are illustrated in Figures 1a–1d. The observers were instructed to report the array type— that is, to discriminate horizontal ($d_h/d_v = 0.8$) and vertical ($d_h/d_v = 1.2$) arrays. One should note that the difference between the horizontal and vertical mean spacings ($d_h - d_v = 0.25°$) was larger than the positional variance or jitter ($-0.1875°-0.185°$).

For each array size, performance was established at a number of SOAs between 30 and 110 msec, in steps of 20 msec. The resulting psychometric curves are shown in Figure 2a for three observers. It is clear that performance improves with increasing array size. This outcome becomes even more evident when a suitable SOA is se-
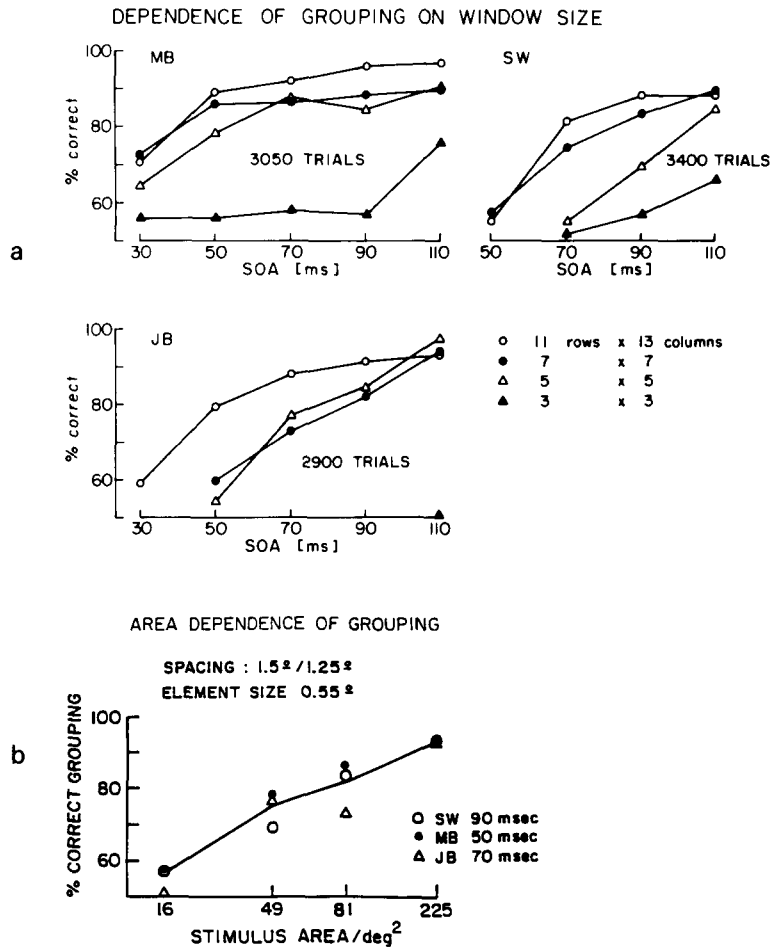
DEPENDENCE OF GROUPING ON WINDOW SIZE

Figure 2. Results for Experiment 1. (a) Performance of grouping discrimination (ordinate) as a function of stimulus onset asynchrony (SOA, abscissa) and of array size (four curves, one for each array size), for Observers M.B., S.W., and J.B. (b) Performance of grouping discrimination (ordinate) as a function of array size (abscissa), for one SOA value per observer (three curves, one for each observer). (Compare Table 2.)

lected for each observer (M.B., 50 msec; S.W., 90 msec; J.B., 70 msec) and performance at that SOA value is plotted against array size, as measured in units of solid visual angle (Figure 2b). All 3 observers exhibit similar behavior: performance is at chance level for the smallest stimulus size, improves as the stimulus size increases, and reaches its highest level only for the largest (11 × 13) array. It seems quite possible that even larger array sizes would have raised performance to even higher levels.

How does this performance increase compare with the increase that is expected on the basis of probability summation? Table 2 lists the observed performance levels, expressed as d', for each array size, along with the d' expected on the basis of probability summation. The expected d' is obtained from the next smaller or the next larger array size, according to the formula

$$d'_i = d'_{i\pm1} \sqrt{\frac{N_i}{N_{i\pm1}}},$$

where N_i denotes the number of units in the array. We assume that N_i equals the number of array elements. As is evident from Table 2, the increase in performance as array size increases is significantly larger than expected on the basis of probability summation for array sizes of 3×3 and 5×5, but it is comparable to (or smaller than) the expected value for array sizes of 7×7 and 11×13. These considerations suggest that the relevant perceptual unit for array identification is a portion of 5×5 elements or more, not individual elements or pairs of elements. Accordingly, a judgment about the precise separation between neighboring array elements is not the strategy used

Table 2

Array Size	Elements	Angle Size	Performance	d'	d'_i from d'_{i-1}
3×3	9	4°×4°	54.0±2.2	0.10±0.03	—
5×5	25	7°×7°	75.0±2.9	0.67±0.05	0.17±0.06
7×7	49	9°×9°	81.0±3.9	0.88±0.08	0.94±0.07
11×13	143	15°×15°	88.0±0.33	1.18±0.01	1.50±0.14

by observers in carrying out the array identification task. The strategy that is used instead is based on at least 5×5 elements, which is consistent with a strategy based on grouping processes.

Experiment 2: Form Identification and Discrimination of Proximity Grouping

This experiment and the following experiments were conducted to determine the compatibility or incompatibility of two visual tasks performed concurrently, with the aim of uncovering competition for visual attention between the two tasks. Several points of procedure that are common to all of the following experiments deserve to be reiterated.

1. Each experiment was devoted to a particular task combination and consisted of two single-task conditions, in which observers carried out each task alone, and one double-task condition, in which observers attempted to carry out both tasks together. Some experiments required identical single-task data for their constituent tasks; in these cases, one set of single-task data was used in the context of all such experiments.

2. For a given level of performance, most visual tasks require comparable stimulus presentation times (SOA values) when carried out alone; most tasks have comparable psychometric curves. (The array identification tasks in Experiments 5 and 6 were exceptions in this respect.) Although stimulus *presentation* time is of course not identical to stimulus *processing* time, this equivalence would nevertheless seem to hold at least for aspects of visual processing that are *rate-determining* for each task. Accordingly, when an observer successfully performs two tasks together, the rate-determining processes for each task must have been carried out concurrently.

3. Task compatibility or incompatibility was assessed by comparing the performance level of a given task *with* and *without* the concurrent task. In other words, the comparison of interest took place not between different types of tasks, but rather between different types of conditions (i.e., single- and double-task conditions) under which each task was carried out.

4. In all double-task conditions, the tasks involved were *ranked*, one being designated *primary* and the other *secondary*. The observers were instructed to perform optimally on the primary task and on the secondary task only insofar as this would not compromise performance on the primary task. Limiting the study to one extreme point of the attentional operating curve (Sperling & Dosher, 1986) reduced the variability of the outcome, and caused any conflict between tasks to be reflected in the performance of only one task, namely the secondary task.

In Experiment 2, the form identification task was combined with the first array identification task. A typical stimulus is shown in Figures 1d and 1e. The form identification required the observers to report the letter appearing at the center of the stimulus array. This letter took the form of either a ⊤ or a Γ, each with equal probability. The array identification required observers to report the array type, either horizontal (d_h/d_v = 0.8) or vertical (d_h/d_v = 1.2), both being equally probable.

Under single-task conditions (form identification alone, array identification alone) observers gave one response (0 or 1 on the terminal keyboard). Under the double-task condition (form identification and array identification combined), form identification was designated primary and array identification secondary. The observers gave two (serial) responses, with the first response pertaining to the primary task.

The results are presented graphically in Figure 3, and in numerical form in Table 3. Each graph summarizes the performance data obtained from 1 observer (M.B., S.W., J.B., or H.S.) on one of the two tasks (form identification or array identification), and juxtaposes his/her performance on this task under single-task (open circles) and double-task (filled circles) conditions. The graphs are based on 1,600–3,750 trials each. The values listed in Table 3 are the parameter T, which reflects SOA at threshold (see the Method section). For the primary task (form identification), performance is not expected to differ between single- and double-task situations. In fact, Observer H.S. exhibited a slight nonsignificant reduction in primary task performance, while Observer J.B. showed a small significant enhancement, perhaps due to an increase in vigilance under double-task conditions. Similar results were obtained for the performance of the *primary* task in Experiments 3–6.

On the secondary task (array identification), all observers suffered a large and highly significant (p < .005) performance reduction under double-task conditions. This was observed for all except the very largest (≤ 140-msec) SOA values. Evidently, concurrent execution of the primary task impairs the ability of observers to identify array type and interferes with the grouping processes that presumably mediate array identification under more favorable circumstances. It would appear, therefore, that perceptual grouping and form identification compete with each other for attentive resources.

Experiment 3: Form Identification and Singularity Detection

We have shown previously that the detection of a singularity in an otherwise uniform background texture is

CENTER 'T'/'L' DISCRIMINATION (PRIMARY TASK)

VERTICAL / HORIZONTAL (GROUPING) DISCRIMINATION (SECONDARY TASK)
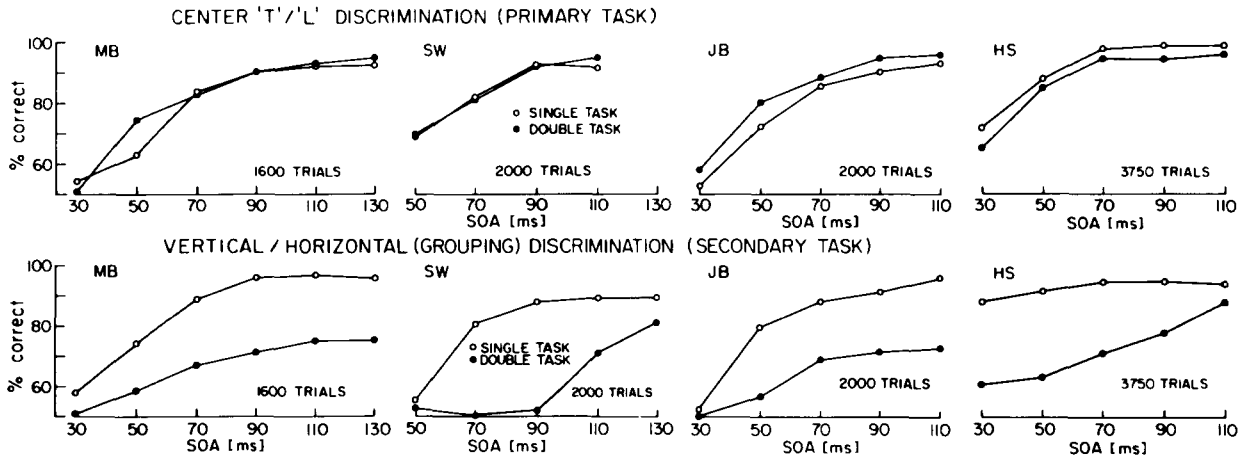
Figure 3. Results for Experiment 2. The experiment combined form identification and discrimination of proximity grouping. The upper row of graphs shows performance on form identification (ordinate) as a function of stimulus onset asychrony (SOA, abscissa), for Observers M.B., S.W., J.B., and H.S. The lower row of graphs shows performance on discrimination of proximity grouping as a function of SOA. Form identification was the primary task; grouping discrimination, the secondary task. Each graph compares optimal performance, obtained when the task was carried out alone (o = single-task situation) with the performance observed when the other task was carried out concurrently (• = double-task situation). The number of trials on which the graph is based is given at the lower right in each graph. In the observer average, performance on form identification does not differ significantly between single- and double-task situations, whereas performance on grouping discrimination is significantly reduced in the double-task situation (Table 3).

largely unaffected by the concurrent execution of a form identification task (Braun & Sagi, 1990, 1991, 1992). We took this outcome to show that the processes underlying texture segregation do not share attentive resources with form identification. Intrigued by the contrast between the earlier finding concerning texture segregation and the present one concerning perceptual grouping, we replicated the experiment of Braun and Sagi (1991) with the present stimulus pattern (Figures 1d and 1e). Due to the presence, in half of the trials, of a single L among the array of + s, this pattern permits us to combine the form identifi-

cation with a singularity detection task (see the Method section). Form identification involved reporting the letter (Γ or T) at the center of the stimulus array. Singularity detection involved reporting the presence or absence of an L among the array of + s. When present, the L appeared within an annular region around the center of the array (3.75°–5.86° eccentricity; see the Method section).

Observers responded once under the single-task (form identification alone, singularity detection alone) and twice under the double-task (form identification and singular-

Table 3
Form Identification and Discrimination of Proximity Grouping

	T Values (in Milliseconds)			
	Task			
Observer	Single	Double	Difference	Significance
Form Identification (Primary Task) T				
M.B.	70.00±2.17	65.53±2.44	−4.47±3.26	n.s.
S.W.	66.75±1.62	66.16±1.95	−0.59±2.53	n.s.
J.B.	63.44±1.94	51.40±1.93	−12.04±2.73	n.s.*
H.S.	37.20±3.14	43.20±1.88	6.00±3.66	n.s.
Average	59.35±7.50	56.57±5.61	−2.78±3.77	n.s.
Discrimination of Proximity Grouping (Secondary Task)				
M.B.	56.88±2.97	>100.67±5.27	>43.87±6.05	$p < .005$
S.W.	73.80±1.76	>118.02±2.93	>44.22±3.42	$p < .005$
J.B.	63.04±1.81	>93.12±1.37	>30.08±2.27	$p < .005$
H.S.	36.40±2.34	>77.80±4.16	>41.40±4.79	$p < .005$
Average	57.51±7.86	>97.40±8.36	>39.89±3.33	$p < .005$

*Significant enhancement, $p < .005$.

CENTER 'T'/'L' DISCRIMINATION (PRIMARY TASK)

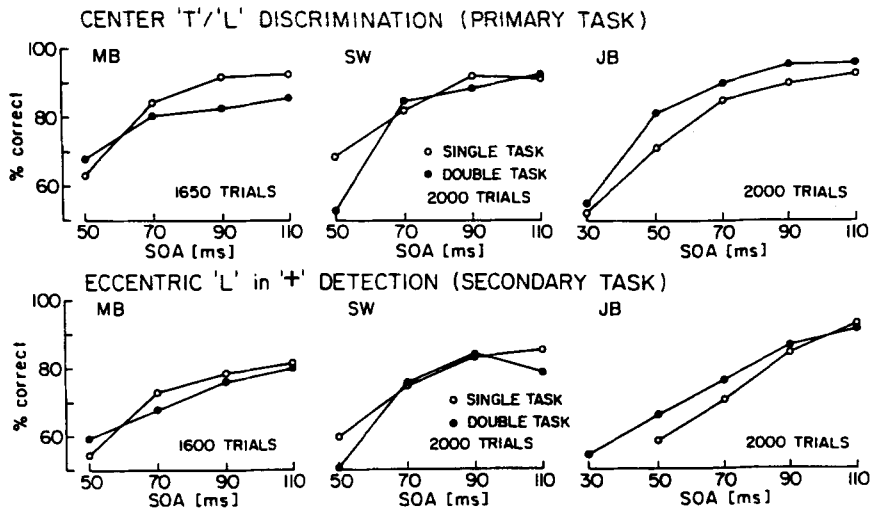ECCENTRIC 'L' in '+' DETECTION (SECONDARY TASK)

Figure 4. Results for Experiment 3, which combined form identification and singularity detection, for Observers M.B., S.W., and J.B. (o = single-task situation, • = double-task situation). Performance on neither task differs significantly between single- and double-task situations (Table 4).

ity detection combined) conditions. Under the latter condition, form recognition was designated the primary and singularity detection the secondary task. The first response pertained to the primary task. The results are presented graphically (Figure 4) as well as numerically (Table 4), with the format of presentation the same as before. The graphs are based on 1,600–2,000 trials each.

Performance on both the primary (form identification) and secondary (singularity detection) tasks was similar under the double- and single-task conditions. Double-task performance on singularity detection, where any competition for resources is expected to become apparent, was not significantly changed for Observer M.B., somewhat reduced ($p < .05$) for Observer S.W., and somewhat enhanced ($p < .01$) for Observer J.B., relative to the single-task performance in each case. Across observers, a paired t test revealed no significant change in performance.

The compatibility of concurrently executed form identification and texture segregation tasks demonstrates that there is no measurable competition for attentive resources between these tasks. Together with the outcome of the previous experiment, this result suggests that our array identification and singularity detection tasks differ substantially in the demand that they place on attentive resources. More generally, the two results suggest that fundamentally different types of processes mediate perceptual grouping and texture segregation, respectively.

Experiment 4: Discrimination of Proximity Grouping and Singularity Detection

The outcomes of the preceding two experiments showed that our array identification task conflicts with our form identification task, whereas our singularity detection task does not. If conflicts between these tasks can occur only

Table 4
Form Identification and Singularity Detection

| | T Values (in Milliseconds) | | | |
| | Task | | | |
Observer	Single	Double	Difference	Significance
	Form Identification (Primary Task)			
M.B.	68.60±2.00	74.96±1.84	6.36±2.72	$p < .01$
S.W.	66.75±1.62	72.44±1.38	5.69±2.12	$p < .01$
J.B.	63.44±1.94	50.49±2.18	−12.95±2.92	n.s.*
Average	66.26±1.51	65.96±7.77	−0.30±6.33	n.s.
	Singularity Detection (Secondary Task)			
M.B.	85.80±2.27	87.04±2.41	1.24±3.31	n.s.
S.W.	80.22±2.05	85.07±1.74	4.84±2.69	$p < .05$
J.B.	81.26±0.95	75.91±1.96	−5.35±2.18	n.s.†
Average	82.43±1.71	82.67±3.43	0.24±2.98	n.s.

*Significant enhancement, $p < .005$.   †Significant enhancement, $p < .01$.
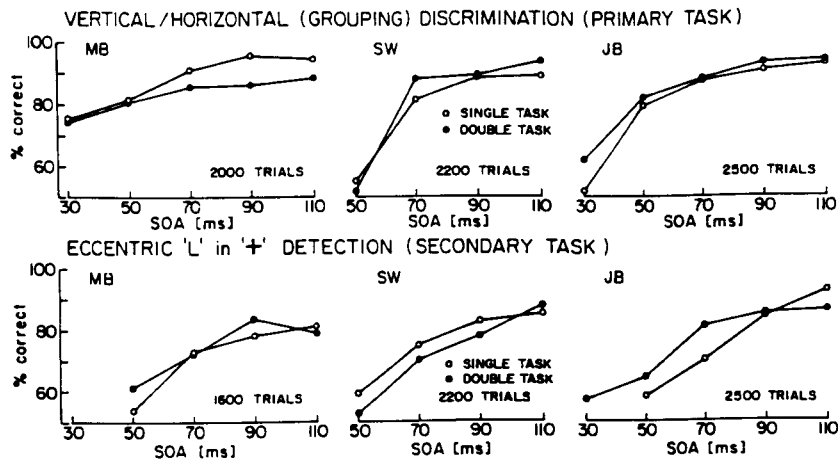
Figure 5. Results for Experiment 4, which combined discrimination of proximity grouping with singularity detection, for Observers M.B., S.W., and J.B. (○ = single-task situation, • = double-task situation). Performance on neither task differs significantly between single- and double-task situations (Table 5).

on one level, namely, the level of competition for attentive resources, as we would like to assume, then one would expect the singularity detection to be compatible not only with the form identification but also with the array identification task.

To ascertain whether or not this expectation would be met, we conducted an experiment in which array identification was combined with singularity detection. The stimulus was identical to the one used before (Figures 1d and 1e), and the procedure remained unchanged as well: array identification required a report on array type (horizontal, $d_h/d_v = 0.8$, or vertical, $d_h/d_v = 1.2$), whereas singularity detection required a report on the presence or absence of an L that could appear in an annular region around the center of the array. The observers responded once under the single-task condition (array identification alone, singularity detection alone) and twice under the

double-task condition (array identification and singularity detection combined). In the double-task situation, array identification was the primary task and the first response pertained to it.

The results are presented graphically (Figure 5) and numerically (Table 5) in unchanged format. The graphs are based on 1,600–2,500 trials each. All observers show nearly identical performance under the single- and double-task conditions for both the primary (array identification) and the secondary (singularity detection) tasks. As in earlier experiments, Observer J.B. experienced a significant ($p < .005$) performance enhancement under the double-task conditions, presumably because of increased vigilance. Small significant ($p < .05$) reductions in double-task performance occurred for Observer M.B. in the case of the array identification and for Observer S.W. in the case of the singularity detection. Across observers,

Table 5
Discrimination of Proximity Grouping and Singularity Detection

	T Values (in Milliseconds)			
	Task			
Observer	Single	Double	Difference	Significance
Discrimination of Proximity Grouping (Primary Task)				
M.B.	46.80±2.02	54.46±2.66	7.66±3.34	$p<.05$
S.W.	73.80±1.76	71.39±2.09	−2.41±2.74	n.s.
J.B.	63.04±1.81	52.45±2.05	−10.59±2.73	n.s.*
Average	61.21±7.85	59.43±6.01	−1.78±5.28	n.s.
Singularity Detection (Secondary Task)				
M.B.	85.80±2.27	81.50±2.39	−4.30±3.30	n.s.
S.W.	80.22±2.05	87.67±3.14	7.44±3.75	$p<.05$
J.B.	81.26±0.95	69.99±1.66	−11.26±1.92	n.s.*
Average	82.43±1.71	79.72±5.18	−2.71±5.46	n.s.

*Significant enhancement, $p < .005$.

a paired *t* test revealed a small significant enhancement in the performance of both tasks under the double-task conditions.

The outcome shows that there exists no measurable conflict between the concurrent performance of the array identification and the singularity detection tasks. This result is consistent with the hypothesis that, in any combination, the tasks investigated here conflict only at one level.

Experiment 5: Form Identification and Detection of Proximity Grouping

It is possible that the array identification task in Experiments 1-4 involves two separate steps: establishing the perceptual organization of the stimulus array, and discriminating between the two possible forms of perceptual organization (horizontal and vertical grouping). If this was indeed so, there would be more than one way of explaining why array identification requires visual attention (Experiment 2); namely, the demand for attentive resources

could be due to the first step, the second step, or both. To get around this concern, we repeated Experiment 2 with an alternative array identification task, chosen in such a way as to retain the first step (establishing perceptual organization) and to remove the second step (discriminating between orientations).

Specifically, we used an array identification task in which observers reported merely the presence or absence of perceptual organization. To this end, the stimulus was modified to present the observer with three array types: a horizontal type with $d_h/d_v < 1.0$ (25% of the trials), a vertical type with $d_h/d_v > 1.0$ (25% of the trials), and a neutral type with $d_h/d_v = 1.0$ (50% of the trials). The task of the observer consisted of discriminating between perceptually organized (horizontal or vertical) array types and the neutral array type.

Using ratios of mean element spacing of $d_h/d_v = 0.8$, 1.2, and 1.0, this grouping detection task turned out to be more difficult than the grouping discrimination task used before.[2] In order to permit observers to perform at
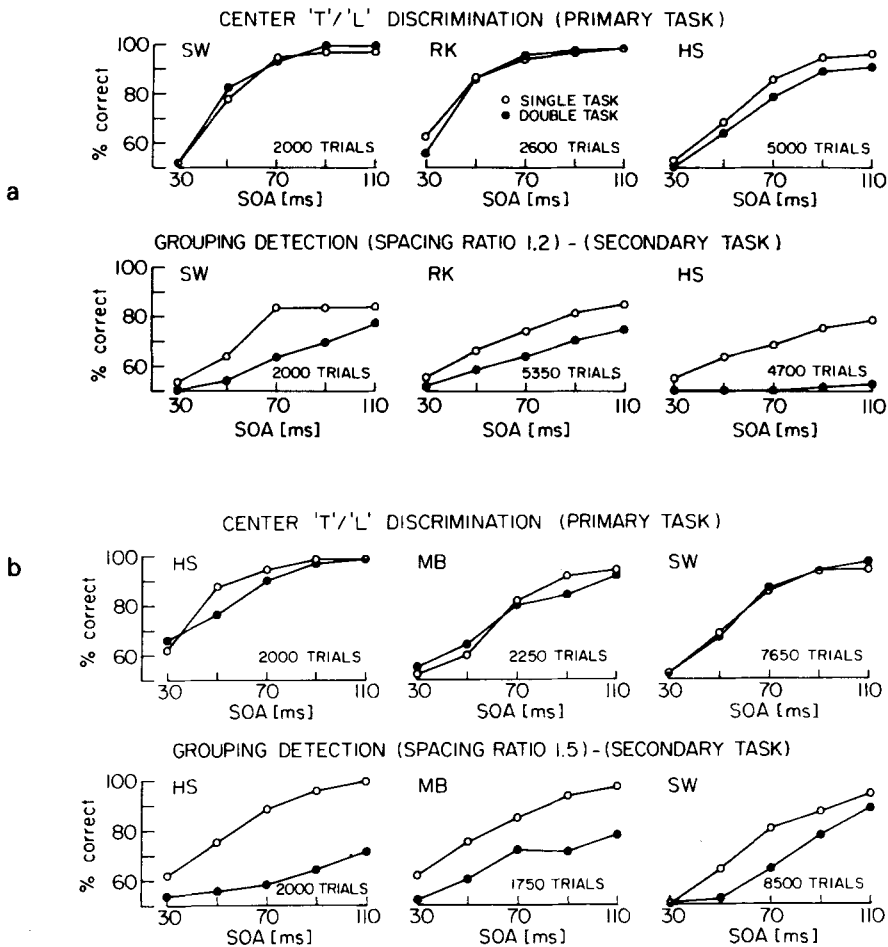

Figure 6: Results for Experiment 5, which combined form identification with detection of proximity grouping, in an array with a ratio of element spacing of 1.2 (a) , and of 1.5 (b), for Observers S.W., R.K., and H.S. (o = single-task situation, • = double-task situation). Performance on grouping detection is significantly lower in the double-task situation (Tables 6 and 7).

levels that were comparable to those in previous experiments, we repeated the experiment with a larger difference in element proximity, using ratios of mean element spacing of d_h/d_v = 0.66, 1.5, and 1.0 (see the Method section). This change had the additional effect of making comparable the respective single-task performance levels for array identification and form identification tasks. In all other respects, the experiment was analogous to Experiment 2. As before, form identification required observers to report the letter (T or Γ) appearing at the array's center and was designated the primary task. In the double-task situation, the first response pertained to the form identification task. The results are represented graphically in Figures 6a and 6b and numerically in Tables 6 and 7. The graphs are based on 2,000–5,350 trials (Figure 6a) or on 1,750–8,500 trials each (Figure 6b).

Performance on the primary task (form identification) was similar under single- and double-task conditions, although for the smaller proximity differences (d_h/d_v = 0.8, 1.2, or 1.0) a small significant ($p < .01$) reduction in double-task performance was obtained from Observer H.S. and a small significant ($p < .05$) enhancement was obtained from Observer S.W. On the secondary task (second array identification), all observers experienced a large and highly significant ($p < .005$) reduction in the double-

task performance. This was true for the smaller (d_h/d_v = 0.8, 1.2, 1.0) as well as for the larger (d_h/d_v = 0.66, 1.5, 1.0) proximity differences.

The results show that execution of the primary task severely impairs the ability of observers to carry out the modified array identification task, which involves only the detection, not the discrimination, of perceptual organization. The outcome suggests that the process of perceptual grouping, which we assume underlies the performance of the array identification, poses a demand for attentive resources that competes with a demand from the form identification.

Experiment 6: Form Identification and Detection of Similarity Grouping

The main finding reported so far is that the execution of tasks assumed to rely on perceptual grouping conflicts with the concurrent execution of a form identification task. To extend the generality of this finding, we investigated a third type of array identification task, in which element *similarity* rather than proximity is the factor responsible for the organization (both objective and perceptual) of the stimulus array. Preliminary experiments confirmed that similarity-based array identification is based on a large part of the stimulus array; that is, the task is not carried

Table 6
Form Identification and Detection of Proximity Grouping (r = 1.2)

| | T Values (in Milliseconds) | | | |
| | Task | | | |
Observer	Single	Double	Difference	Significance
	Form Identification (Primary Task)			
S.W.	54.40±1.90	49.80±1.76	−4.60±2.59	n.s.*
R.K.	45.79±1.58	48.25±2.10	2.46±2.63	n.s.
H.S.	60.44±2.84	71.73±3.33	11.29±4.38	$p<.01$
Average	53.54±4.25	56.59±7.58	3.05±4.60	n.s.
	Detection of Proximity Grouping (Secondary Task)			
S.W.	72.40±2.57	>94.00±4.09	>21.60±4.82	$p<.005$
R.K.	74.84±2.16	>92.71±1.80	>17.88±2.82	$p<.005$
H.S.	82.64±1.88	>120.52±2.16	>37.88±2.86	$p<.005$
Average	76.63±3.09	>102.41±9.06	>25.79±6.14	$p<.05$

*Significant enhancement, $p < .05$.

Table 7
Form Identification and Detection of Proximity Grouping (r = 1.5)

| | T Values (in Milliseconds | | | |
| | Task | | | |
Observer	Single	Double	Difference	Significance
	Form Identification (Primary task)			
H.S.	43.80± 2.59	49.20± 3.02	5.40±3.98	n.s.
M.B.	69.36± 2.75	70.75± 2.23	1.39±3.54	n.s.
S.W.	63.26± 1.85	62.65± 1.43	−0.61±2.34	n.s.
Average	85.47±32.43	87.53±31.85	2.06±1.77	n.s.
	Detection of Proximity Grouping (Secondary Task)			
H.S.	52.00± 1.83	>99.20± 2.54	>47.20±3.12	$p<.005$
M.B.	55.80± 1.79	>87.17± 2.71	>31.37±3.25	$p<.005$
S.W.	70.39± 2.10	>86.45± 2.19	>16.06±3.03	$p<.005$
Average	86.06±25.43	>117.61±25.05	>31.54±8.99	$p<.05$
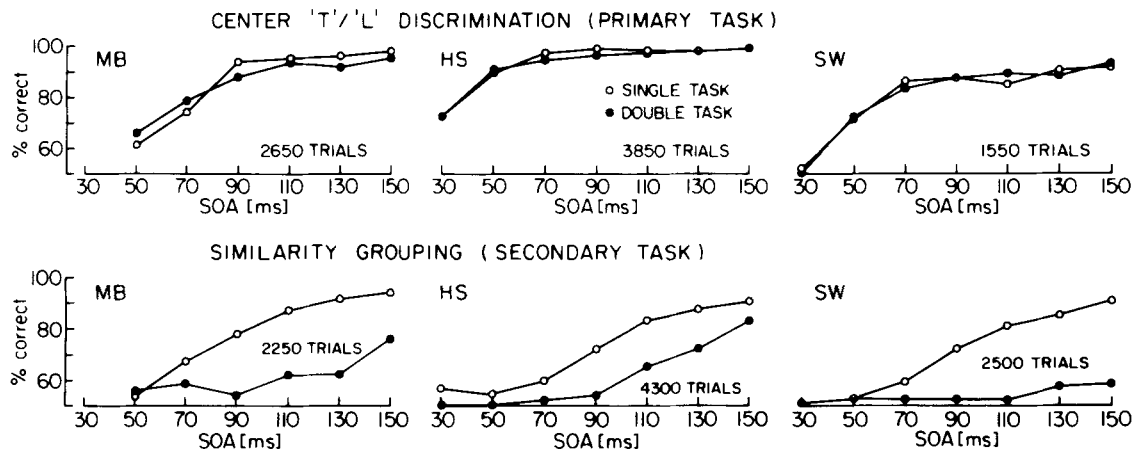
Figure 7: Results for Experiment 6, which combined form identification with detection of similarity grouping, for Observers M.B., H.S., and S.W. (o = single-task situation, • = double-task situation). Performance on grouping detection is significantly lower in the double-task situation (Table 8).

Table 8
Form Identification and Detection of Similarity Grouping

| | T Values (in Milliseconds) | | | |
| | Task | | | |
Observer	Single	Double	Difference	Significance
	Form Identification (Primary Task)			
M.B.	72.77± 1.79	75.47± 2.72	2.69±3.26	n.s.
H.S.	39.54± 1.66	41.88± 2.75	2.34±3.21	n.s.
S.W.	75.20± 2.65	76.00± 3.59	0.80±4.47	n.s.
Average	62.50±11.50	64.45±11.29	1.94±0.58	n.s.
	Detection of Similarity Grouping (Secondary Task)			
M.B.	92.20±3.28	>132.47±3.06	>40.27±4.49	$p<.005$
H.S.	98.39±3.23	>130.36±3.65	>31.97±4.87	$p<.005$
S.W.	103.32±1.95	>149.07±6.87	>45.75±7.14	$p<.005$
Average	97.97±3.22	>137.30±5.92	>39.33±4.01	$p<.005$

out by scrutinizing and identifying the shapes and positions of a few individual array elements (results not shown).

To this end, we presented a stimulus pattern consisting of L s and + s arranged in a square array of 11×11 elements. Three types of organization were used: a vertical type with L s and + s forming alternating columns, a horizontal type with L s and + s forming alternating rows, and a neutral type with L s and + s distributed randomly (Figures 1g and 1h). At the center of the array, a smaller element appeared in the shape of either a T or a smaller Γ.

The array identification task consisted of reporting the presence (but not the type) of organization in the stimulus array (horizontal or vertical array), or its absence (neutral array). The form identification task involved, as before, a report on the shape (T or Γ) of the central array element. In the double-task situation, form identification was the primary task and the observer's first response pertained to this task. The outcome is presented graphically

in Figure 7 and numerically in Table 8. Individual graphs are based on 1,550–4,300 trials each.

For the primary task (form identification), performance is similar under double- and single-task conditions. Statistical analysis (t test) of the temporal threshold data shows a highly significant difference ($p < .005$) between the single-task and the double-task conditions of the array identification task. It is also clear that the present variant of detecting organization within a stimulus array finds itself in conflict with the execution of a form identification task. This outcome suggests that perceptual grouping on the basis of similarity requires access to attentive resources.

DISCUSSION

The general aim in the present study was to investigate whether or not visual attention is required for perceptual grouping. To address this issue, we had observers attempt to carry out two visual tasks concurrently: a task assumed

to reflect perceptual grouping, and a task assumed to engage visual attention. Thus, visual attention was expected to be unavailable, or available only to a small extent, for the presumed grouping task. The outcome, obtained with four variants of grouping, was a dramatic reduction in the performance of the presumed grouping tasks. In an initial control experiment, the task assumed to engage visual attention was combined with a task assumed to reflect texture segregation, and in a second control experiment, the presumed grouping task was combined with the presumed texture segregation task. Neither combination led to a significant impairment of any of the tasks involved. This demonstrates that only some, and not all, combinations of tasks give rise to a conflict. Specifically, the problematic combinations appear to be those in which one task involves perceptual grouping (rather than texture segregation) and the other task engages visual attention (rather than relying on preattentive mechanisms).

· If the stated assumptions concerning the nature of our tasks are correct, the results would seem to show that perceptual grouping requires visual attention. Since this would contradict some of the current thinking on perceptual grouping, we will examine once more the basis for these assumptions, beginning with the task meant to engage visual attention.

Form Identification Task and
Attentive Resources

A considerable literature (Bergen & Julesz, 1983; Braun & Sagi, 1990, 1991; Duncan, 1979, 1985; Eriksen & St. James, 1986; Eriksen & Yeh, 1985; LaBerge, 1983; LaBerge & Brown, 1989; Sagi & Julesz, 1985a, 1985c; Treisman, 1985; Treisman & Gormican, 1988; Treisman and Paterson, 1984) suggests that a form identification task such as the Γ/T discrimination used here is resource limited (Norman & Bobrow, 1975; Sperling & Dosher, 1986), in the sense that the performance level directly reflects the amount of attentive resources devoted to the task. When an observer is presented with two simultaneous form identification tasks in separate parts of the field of view, the two tasks conflict, and performance declines in one task or the other (Adini and Sagi, 1992; Braun & Sagi, 1991; Duncan, 1984). Specifically, if either of the two tasks is performed optimally, performance on the other task is impaired drastically and recovers only if additional viewing time (on the order of 50-100 msec) is made available (Braun & Sagi, 1991). Such "serial" behavior is considered characteristic of tasks that require visual attention (Julesz, 1981; Treisman and Gelade, 1980).

Experiments like these suggest that performing a form identification task at the optimal level produces a concentration of attentive resources (on the item to be identified), and, as long as performance remains optimal, it can be inferred that the presence of other items in the field of view, and/or the performance of other visual tasks with respect to such items, does not weaken or dilute this concentration of resources. Of course, sufficiently conspicuous secondary item(s) could conceivably produce a more even distribution of attentive resources and could force performance of the form identification below its optimal level. However, our stimulus array did not have this effect on form identification: neither presence/absence of the stimulus array (around the central Γ/T target) (Barchilon Ben-Av, 1992) nor performance/nonperformance of a secondary task was observed to significantly affect performance on the Γ/T discrimination.

Accordingly, it appears safe to assume that performance on the Γ/T discrimination made attentive resources less available, if at all, for secondary tasks with respect to the array around the Γ/T target. Whether the reduction in attentive resources was complete or incomplete (but substantial) is not critical to the interpretation of our main result, which consists of a *differential* effect on tasks assumed to reflect perceptual grouping and on tasks assumed to reflect texture segregation.

Array Identification Tasks and
Perceptual Grouping

Aiming to measure perceptual grouping, we employed four tasks that involved the discrimination of different types of arrays, composed of randomly rotated + s and, in one type of array, ∟s. Differential proximity or, in one type of array, differential similarity between neighboring array elements elicited strong perceptual organization in either the horizontal or the vertical direction. In other types of arrays, the differences in proximity or similarity were not distributed systematically among array elements and failed to elicit perceptual organization. Specifically, the four tasks were (1) the discrimination between horizontal and vertical organization due to proximity; (2) the discrimination between organization based on proximity and no organization; (3) the same as Task 2, but with larger differences in proximity; and (4) the discrimination between organization based on similarity and no organization.

Several arguments support the hypothesis that our array identification tasks reflect perceptual grouping.

1. Although the physical differences between array types would, in principle, have allowed discrimination on the basis of only a few (two or three) neighboring array elements, psychophysical performance was not based on local cues but on a larger part of the array, measuring roughly 5×5 elements or more (Experiment 1).

2. Element proximity and similarity, the two factors manipulated by the physical attributes of the stimulus arrays, are known to be strong determinants of perceptual grouping (Koffka, 1935).

3. The introspective appearance of the stimulus arrays was dominated by perceptual grouping.

4. The four tasks are similarly affected by manipulation of array size (Barchilon Ben-Av, 1992) and by restrict-

ing the availability of attentive resources, suggesting that performance on all four tasks is based on similar perceptual mechanisms.

Singularity Detection Task and Texture Segregation

To measure texture segregation, we used a task involving the detection of singularity (a single, randomly rotated L) in a dense background texture (the array of randomly rotated + s). The juxtaposition of textures based on Ls and + s has frequently been used in texture research (Bergen & Julesz, 1983; Julesz & Kröse 1988; Williams & Julesz, 1991). Both psychophysical evidence (Gurnsey & Browse, 1987) and quantitative models of texture segregation (Fogel & Sagi, 1989; Malik & Perona, 1990; Rubenstein & Sagi, 1990) suggest that the same mechanisms affect texture segregation on the basis of differences in shape (e.g., between line figures such as Ls and + s) and differences in orientation or spatial frequency. Accordingly, there can be little doubt that our L in + detection task reflects texture segregation.

Pairwise Compatibility or Incompatibility of Visual Tasks

We report here on six situations in which we have assessed the compatibility of two visual tasks performed concurrently. In each situation, the two tasks were ranked into a primary and a secondary task and observers were instructed to perform the primary task at optimal levels. For the secondary task, a range of outcomes was observed. These were quantified in terms of the increase in threshold SOA relative to the optimal performance level, ΔT (see the Method section). The distribution of outcomes contained values of ΔT that were not significantly different from zero (0.2 ± 3.0 msec and -2.7 ± 5.5 msec, observer averages) and values of ΔT that were significantly larger than zero (39.9 ± 3.3, 25.8 ± 6.1, 31.5 ± 9.0, and 39.3 ± 4.0 msec, observer averages). Accordingly, the outcomes can be described in terms of either the compatibility (ΔT around 0 msec) or the incompatibility (ΔT around 25–40 msec) of the two visual tasks in question.

Compatibility was observed between the form identification (as the primary task) and the singularity detection (as the secondary task), and also between an array identification (primary task) and the singularity detection (secondary task). Incompatibility was observed between the form identification (primary task) and all four types of array identifications (secondary task). Additional experiments have shown the incompatibility of array identification and form identification under the opposite task ranking (array identification primary, form identification secondary) (Barchilon Ben-Av, 1992).

The pattern of outcomes is consistent with two simple hypotheses: (1) that incompatibility occurs if and only if two tasks compete for certain resources, and (2) that form identification and array identification present a demand for the resources in question, but that singularity detec-

tion does not. Since form identification is known to engage visual attention, one is compelled to conclude that the resources at issue are almost certainly the resources of visual attention. If one assumes further that array identification reflects perceptual grouping, it follows that perceptual grouping presents a demand for visual attention.

When singularity detection was the secondary task (there were two such situations), compatibility between primary and secondary tasks was observed, suggesting that singularity detection presents no measurable demand for visual attention. This conclusion is consistent with previous work in which we have been unable to discern a requirement for attentive resources in tasks involving texture segregation (Braun & Sagi, 1990, 1991, 1992; Karni & Sagi, 1991; Sagi & Julesz, 1985a, 1985c). In our view, these results show conclusively that certain texture segregation tasks do not involve visual attention to a measurable degree and rely entirely, or virtually entirely, on preattentive mechanisms.

Implications for Perceptual Grouping

To the best of our knowledge, the experiments reported here constitute the first direct evidence for the necessity of visual attention in perceptual grouping. A previous report showed that the pace of visual search, which is assumed to reflect the pace of successive allocations of visual attention, depends strongly on whether or not display items are perceptually organized into groups (Treisman, 1982). Apparently, grouping allows attention to be allocated to more items at a time and thus accelerates visual search. From our point of view, it is interesting that the number of attentive fixations appears to be determined by the number of groups in the display, which is consistent with the possibility that the groups are formed one by one with the help of visual attention. However, Treisman (1982) assumes that groups are established preattentively at the very beginning of, rather than throughout, the search process, noting that "the theories all agree that perceptual grouping occurs automatically and in parallel, without attention or scrutiny," and that "this preattentive organization should then affect all subsequent stages of processing." We believe that this view, which is representative of current thinking on perceptual grouping, must be reconsidered in the face of the evidence presented here.

The relation between texture segregation and perceptual grouping must, in our opinion, be reexamined as well. The notion that "perceptual grouping is a special case of texture segregation" (Treisman, 1982) stems from what is known about the stimulus features that do, and those that do not, support perceptual grouping and texture segregation. Features such as luminance, orientation, or spatial frequency are effective in the context of both perceptual grouping and texture segregation, whereas features such as the curvature of line elements or the arrangement of line elements are not (Beck, 1966, 1967, 1972; Beck et al., 1991; Beck et al., 1987; Treisman, 1985; Treisman & Gormican, 1988). Findings of this type suggest that texture segregation as well as perceptual grouping

uses information that is represented at an early stage of visual processing. Notwithstanding these similarities, the fact that the two types of tasks behave so differently in the face of competition for visual attention shows that there also exists a fundamental difference between them. Indeed, it is difficult to conceive of a difference more fundamental than the dramatic imbalance in attentive requirements observed here.

How can one account for the fact that texture segregation and perceptual grouping are so similar and so different at the same time? A closer look at the two processes' presumed perceptual functions suggests a possible answer (with a little help from hindsight). In the case of the visual processing of stimulus texture, a significant part of the function would seem to be the registration of textural boundaries and lines of discontinuity. Once registered, the pattern formed by these lines can be used for several purposes, as, for example, to help direct shifts of visual attention (Koch & Ullman, 1985), or to reduce the number of possible perceptual organizations of the field of view (Beck, 1982; Beck et al., 1983). In any case, it seems clear that texture segregation is an "edge-based" operation (Mumford, Kosslyn, Hillger, & Herrnstein, 1987). Perceptual grouping, on the other hand, can lend cohesion to a region (or several regions) within the field of view and clearly seems to be a "region-based" operation (Mumford et al., 1987). In addition to perceptual cohesion, however, grouping also creates an internal structure or organization in the affected region. A grouped area is not merely a perceptual unit; it has a perceived direction, pattern of flow, or some other internal organization. This global aspect of grouping—"global" in the sense that it applies to the grouped region in its entirety—does not seem to have a counterpart in texture segregation. Therefore it seems conceivable that *this aspect alone* requires visual attention.

Implications for Visual Attention

How is our view of visual attention and its function affected by the finding that attention is necessary to perceptual grouping? In an attempt to answer this question, we consider two hypothetical functions of visual attention: the selective function and the organizing function. By *selective function*, we refer to the idea that visual attention selects a small subset from the large number of stimuli crowding the field of view, thus reducing the amount of information confronting subsequent processes (e.g., LaBerge & Brown, 1989; Van Essen, Anderson, & Felleman, 1991; Wise & Desimone, 1988). By *organizing function*, we refer to the notion that visual attention not only selects, but also structures what it has selected, creating some sort of organized whole out of the stimulus elements that have been singled out (e.g., Beck, 1982; Beck et al., 1983; Treisman, 1982).

In view of the visual operations that, as we have found, require visual attention—namely, form identification and perceptual grouping—it seems clear that postulating a selective function for visual attention may explain why form identification requires visual attention, but not why per-

ceptual grouping does so. In order to account for both observations, we are compelled to postulate an organizing function for visual attention. The nature of this organizing function is difficult to gauge from present evidence. Because of the size of the region on which grouping performance is based (Experiment 1), it is tempting to invoke the dispersed form of visual attention that has been postulated by several authors (Eriksen & St. James, 1986; LaBerge & Brown, 1986). Independently of the spatial distribution of visual attention, it is conceivable that the act of attentive organizing is precisely that step which, in the view of the Gestalt school, separates figure from ground. Finally, it should not go unremarked that, at least in the case of perceptual grouping, visual attention appears to work with information about luminance, orientation, and spatial frequency, among other features. This suggests that attentive (and not only preattentive) processes can operate in the context of an early representation of the stimulus.

REFERENCES

ADINI, Y., & SAGI, D. (1992). Parallel processes within the 'spot-light' of attention. *Spatial Vision*, 6, 61-77.

BARCHILON BEN-AV, M. (1992). *The role of similarity, proximity and attention on visual grouping*. Doctoral dissertation, Weizmann Institute of Science, Rehovot, Israel.

BECK, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics*, 1, 300-302.

BECK, J. (1967). Perceptual grouping produced by line figures. *Perception & Psychophysics*, 2, 491-495.

BECK, J. (1972). Similarity grouping and peripheral discriminability under uncertainty. *American Journal of Psychology*, 85, 1-19.

BECK, J. (1982). Textural segmentation. In J. Beck (Ed.), *Organization and representation in perception* (pp. 285-317). Hillsdale, NJ: Erlbaum.

BECK, J., & AMBLER, B. (1973). The effects of concentrated and distributed attention on peripheral acuity. *Perception & Psychophysics*, 14, 225-230.

BECK, J., GRAHAM, N., & SUTTER, A. (1991). Lightness differences and the perceived segregation of regions and populations. *Perception & Psychophysics*, 49, 257-269.

BECK, J., PRAZDNY, K., & ROSENFELD, A. (1983). A theory of textural segmentation. In J. Beck & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 1-38). New York: Academic Press.

BECK, J., SUTTER, A., & IVRY, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics, and Image Processing*, 37, 299-325.

BERGEN, J. R., & JULESZ, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, 303, 696-698.

BRAUN, J., & SAGI, D. (1990). Vision outside the focus of attention. *Perception & Psychophysics*, 48, 45-58.

BRAUN, J., & SAGI, D. (1991). Texture-based tasks are little affected by a second task which requires peripheral or central attentive fixation. *Perception*, 20, 483-500.

COLTHEART, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, 27, 183-228.

DUNCAN, J. (1979). Divided attention: The whole is more than the sum of its parts. *Journal of Experimental Psychology: Human Perception & Performance*, 5, 216-228.

DUNCAN, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501-517.

DUNCAN, J. (1985). Visual search and visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 85-106). Hillsdale, NJ: Erlbaum.

ERIKSEN, C. W., & ST. JAMES, J. D. (1986). Visual attention within

and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, **40**, 225-240.

ERIKSEN, C. W., & YEH, Y.-Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 583-597.

FOGEL, I., & SAGI, D. (1989). Gabor filters as texture discriminator. *Biological Cybernetics*, **61**, 103-113.

GLASS, L. (1969). Moire effect from random dots. *Nature*, **223**, 578-580.

GLASS, L., & PEREZ, R. (1973). Perception of random dot interference patterns. *Nature*, **246**, 360-362.

GURNSEY, R., & BROWSE, R. A. (1987). Micropattern properties and presentation conditions influencing visual texture discrimination. *Perception & Psychophysics*, **41**, 239-252.

IRWIN, D. E., & YEOMANS, J. M. (1986). Sensory registration and informational persistence. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 343-360.

JULESZ, B. (1981). Textons, the elements of texture perception and their interactions. *Nature*, **290**, 91-97.

JULESZ, B. (1986). Texton gradients: The texton theory revisited. *Biological Cybernetics*, **54**, 464-469.

JULESZ, B., & KRÖSE, B. (1988). Visual texture perception: Features and spatial filters. *Nature*, **333**, 302-303.

KAHNEMANN, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

KARNI, A., & SAGI, D. (1991). Where practice makes perfect in texture discrimination—evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences*, **88**, 4966-4970.

KOCH, C., & ULLMAN, S. (1985). Shifts in visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, **4**, 219-227.

KOFFKA, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace.

KRÖSE, B., & JULESZ, B. (1989). The control and speed of shifts of attention. *Vision Research*, **29**, 1607-1619.

LABERGE, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 371-380.

LABERGE, D., & BROWN, V. (1986). Variations in size of the visual field in which targets are presented: An attentional range effect. *Perception & Psychophysics*, **40**, 188-200.

LABERGE, D., & BROWN, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, **96**, 101-124.

MALIK, J., & PERONA, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, **7**, 923-932.

MUMFORD, D., KOSSLYN, S. M., HILLGER, L. A., & HERRNSTEIN, R. J. (1987). Discriminating figure form ground: The role of edge detection and region growing. *Proceedings of the National Academy of Sciences*, **84**, 7354-7358.

NEISSER, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

NORMAN, D. L., & BOBROW, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, **7**, 44-64.

NOTHDURFT, H. C. (1985). Sensitivity for structure gradients in texture discrimination tasks. *Vision Research*, **25**, 1957-1968.

OLSON, R., & ATTNEAVE, F. (1970). What variables produce similarity-grouping. *American Journal of Psychology*, **83**, 1-21.

RAMACHANDRAN, V. S. (1988). Perception of shape from shading. *Nature*, **331**, 163-166.

RAMACHANDRAN, V. S., & ANSTIS, S. M. (1985). Perceptual organization in multistable apparent motion. *Perception*, **14**, 135-143.

RUBENSTEIN, B. S., & SAGI, D. (1990). Spatial variability as a limiting factor in texture discrimination tasks: Implications for performance asymmetries. *Journal of the Optical Society of America A*, **7**, 1632-1643.

SAARINEN, J., & JULESZ, B. (1991). The speed of attentional shifts in the visual field. *Proceedings of the National Academy of Sciences*, **88**, 1812-1814.

SAGI, D. (1991). Spatial filters in texture segmentation tasks. In B. Blum (Ed.), *Channels in the visual nervous system: Neurophysiology, psychophysics and models* (pp. 397-424). London: Freund Publishers.

SAGI, D., & JULESZ, B. (1985a). Detection versus discrimination of visual orientation. *Perception*, **13**, 619-628.

SAGI, D., & JULESZ, B. (1985b). Fast non-inertial shifts of attention. *Spatial Vision*, **1**, 141-149.

SAGI, D., & JULESZ, B. (1985c). "Where" and "what" in vision. *Science*, **228**, 1217-1219.

SMIKT, O. (1989). *VDG*. Unpublished manuscript, The Weizmann Institute of Science, Department of Applied Mathematics, Rehovot, Israel.

SPERLING, G., & DOSHER, B. (1986). Strategy and optimization in human information processing. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and performance* (Vol. 1, pp. 2.1-2.65). New York: Wiley.

TREISMAN, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 194-214.

TREISMAN, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, & Image Processing*, **31**, 156-177.

TREISMAN, A., & GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

TREISMAN, A., & GORMICAN, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, **95**, 15-48.

TREISMAN, A., & PATERSON, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 12-31.

ULLMAN, S. (1979a). The interpretation of structure from motion. *Proceedings of the Royal Society London: Series B*, **203**, 405-426.

ULLMAN, S. (1979b). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

VAN ESSEN, D. C., ANDERSON, C. H., & FELLEMAN, D. J. (1991). Information processing in the primate visual system: An integrated systems perspective. *Science*, **255**, 419-423.

WATT, R. (1991). Seeing textures. *Current Biology*, **1**, 137-139.

WERTHEIMER, M. (1923). Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung*, **4**, 301-350. Available in abridged English translation (1958): Principles of perceptual organization. In D. C. Beardsler & M. Wertheimer (Eds.), *Reading in perception* (section 8). New York: Van Nostrand.

WILLIAMS, D. W., & JULESZ, B. (1991). Filters versus textons in human and machine texture discrimination. In H. Wechsler, *Neural networks for human and machine perception*. Orlando, FL: Academic Press.

WISE, S. P., & DESIMONE, R. (1988). Behavioral neurophysiology: Insights into seeing and grasping. *Science*, **242**, 736-741.

NOTES

1. As it is used here, the term *visible persistence* subsumes both the "visible persistence" of Coltheart (1980), which lasts for 100-200 msec after stimulus *onset*, and the "visual analog representation" of Irwin and Yeomans (1986), which is thought to last for 150-300 msec after stimulus *offset*.

2. This is not surprising, given that grouping detection requires observers to discriminate between ratios 0.8 or 1.2 on the one hand and 1.0 on the other (ratio of ratios: 1.2), whereas grouping discrimination requires merely the discrimination between ratios 0.8 and 1.2 (ratio of ratios: 1.5).