Perception & Psychophysics
1995, 57 (5), 682-691

Combining image degradations in a recognition task

WILLIAM R. UTTAL, TODD BARUCH, and LINDA ALLEN
Arizona State University, Tempe, Arizona

Six experiments are reported that investigate the effect on form recognition performance of combining three kinds of stimulus degradations: local area averaging of intensities, low-pass spatial frequency filtering, and random dot visual interference. The effects are shown to be more complicated than previously reported in simple demonstrations. The complexity of the results suggests that models based on single stimulus attributes such as energy or spatial frequency spectrum probably cannot account for the data. Eclectic theories that invoke combinations of redundant processes may be necessary for describing visual recognition phenomena, even within the limited domain examined in this study.

If there is any single image that has made the transition from the esoteric experimental studies of perceptual phenomena to become a public icon, it must be the coarsely sampled or "blocked" picture of Abraham Lincoln that was originally presented in the classic paper by Harmon and Julesz (1973). This degraded picture, shown in Figure 1a, of the face of the great president has become the prototype for paintings by Salvador Dali and the model method for such prosaic tasks as the obscuring of the face of arrested suspects in the evening television news.

Harmon and Julesz (1973) concluded on the basis of their demonstration that the blocked image shown in Figure 1a "can be recognized more easily after blurring the sampled and quantized image" (p. 1194) by filtering out selected high-frequency components. (See Figure 1b.) Hereafter, this is referred to as the *Harmon and Julesz phenomenon*. The particular theoretical explanation that they proposed in order to explain this phenomenon was that the energy in the high spatial frequencies, which had been created at the edges of the averaged regions by the blocking operation, masked or inhibited recognition of the face. They went on to assert that removal of a critical band of that high-frequency information by a spatial frequency filter would selectively improve the recognizability of the face. Hereafter, this is referred to as the *Harmon and Julesz theory*.

There are two main issues revolving around the Harmon and Julesz report that we wish to deal with separately. First, we wish to consider the generality of the phenomenon itself. Second, we wish to consider the current status of their theory. Our interest in the generality of the phenomenon is driven by our general goal of understanding the effects of combined image degradations on perception. It was specifically stimulated by an earlier study from our laboratory (Uttal, Baruch, & Allen, 1995) in which we described the effects of combining image degradations in a *discrimination* paradigm.

The successive blocking and blurring method utilized by Harmon and Julesz (1973) in their recognition experiment may be considered to be the antecedent of that discrimination experiment. However, both our earlier work and the present study explore a wider range of the salient types of degradations (including random dot visual interference, block sizes, and low-pass filter values) than has previously been reported. In our two studies, we also used a different type of stimulus material—aircraft silhouettes—rather than the faces that they used.

The absence of full-scale parametric studies of this and related visual phenomena is a serious deficiency of much of the previous work in the field. It is rare (e.g., Bachmann, 1991) when even a single one of the three degrading variables that we manipulate is presented at other than a few selected values. Clearly, isolated demonstrations do not assay the full complexity of the phenomena encountered when degradations are combined. Not so clear is the fact that demonstrations that sample only a few points of a continuum may lead to highly idiosyncratic, nongeneralizable, and even inappropriate theoretical models.

The results of our previous study (Uttal et al., 1995) indicated that in a discrimination task there was no improvement in perceptual performance comparable to the Harmon and Julesz (1973) phenomenon. Subjects were not able to discriminate two different objects better when they were sequentially blocked and then low-pass filtered than when they were just blocked; sequential degradations always progressively worsened discrimination performance. For a discrimination task, therefore, the Harmon and Julesz phenomenon does not obtain.

The question then arises—To what can the difference between the outcomes of the two studies (Harmon & Julesz, 1973; Uttal et al., 1995) be attributed? Is it the task, the stimulus materials, the more extensive range of independent variables, the order of combination, or some other aspect of the procedure?

This research was supported by Grant F49620-92-J-0176 P0001 from the United States Air Force Office of Scientific Research. I am grateful to my colleague Takeo Watanabe for his comments during the preparation of this article. Correspondence should be addressed to W. R. Uttal, Perception Laboratory, Departmment of Industrial and Management Systems Engineering, Arizona State University, Tempe, AZ 85287-5906 (e-mail: aowru@asuvm.inre.asu.edu).

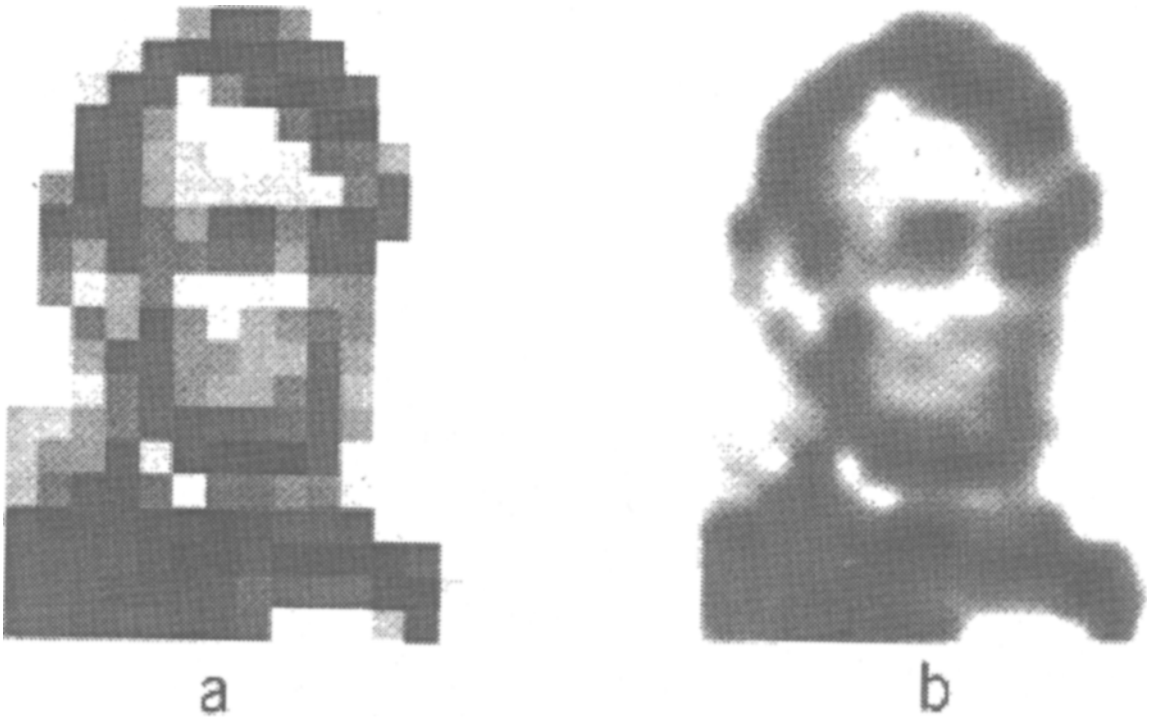

Figure 1. The classic icon: Harmon and Julesz's (1973) sequential degradation of the face of Abraham Lincoln: (a) The original portrait after "blocking"—intensity averaging over square regions. (b) The blocked picture after low-pass filtering with a cutoff limit just above the highest frequency present in the original picture. The Harmon and Julesz phenomenon is the purported improvement in recognition of picture b as opposed to picture a.

To answer this question fully, it is necessary to carry out at least three follow-up experiments. The first is to use the same stimulus material that we used previously (the aircraft silhouettes) in a recognition task. The second is to use faces in an expanded recognition task, and the third is to use faces as stimuli in a discrimination task.

Considering the outcome of all of these studies, it would be important to ask whether other experimental conditions, paradigms, and results provide justification for maintaining the frequency domain based model as the context in which analyses should be embedded. Harmon and Julesz's explanation is essentially a "bottom-up" model that assumes that the processing of the spatial frequency components of the stimulus is the critical element. However, a fuller, parametric investigation of these phenomena may suggest that there are "top-down" processes, collectable under the rubric of *perceptual organization*, that may be better descriptions and explanations of the data obtained so far.

In the present article, we report the results of the first of these proposed follow-up experiments, in which aircraft silhouettes were used as a stimulus in a recognition experiment. In preview, we wish to highlight one particularly surprising result. In the earlier article from this laboratory (Uttal et al., 1995), it was reported that the order of the degradations was not important. That is, a stimulus could be blocked and then low-pass filtered, *or* low-pass filtered and then blocked, with exactly the same results in a discrimination task. In other words, the degradations were commutative. We also explored comparable order effects in the *recognition* task of the present study and, very surprisingly, discovered that *either order of imposing the two degradations produced* improvement! Though the magnitudes of the effects were not the same, the qualitative improvement observed when blocking followed spatial frequency filtering is contrary to the predictions of the Harmon and Julesz (1973) theory.

It is now clear that combining image degradations produces much more complicated perceptual effects than was originally suggested. The results obtained in the present study tap deeper into those complexities, but certainly they do not fully answer all the questions concerning the interactive effects of spatial frequency filtering, blocking, visual interference, task, stimuli, sequential order, or any of the other parameters that are relevant to this process.

Beyond the results reported here, there are other indications that our science has not yet produced a satisfactory theory of these combined effects. The explanation of the phenomenon offered by Harmon and Julesz on the basis of masking or inhibition by particular bands of high spatial frequency information of low spatial frequency information has, for example, also been challenged by Morrone, Burr, and Ross (1983) and by Durgin and Proffit (1993).

The purpose of this article is to provide a more complete empirical description of the perceptual effects of combining degradations in a recognition task and to take some steps toward resolving these contradictory results. Given the complexity of the results obtained in our studies when degradations were combined, it is unlikely that we will resolve all issues and answer all questions. However, an improved understanding of both the empirical and the theoretical situations may emerge from the results of the six experiments reported here.

METHOD

Subjects

Between 6 and 8 subjects were used in each of the experiments reported in this article. Each subject had normal or corrected vision and was paid an hourly stipend plus a bonus for completion of each experiment. All were well-trained observers rather than casual participants. Unfortunately, because of the extended duration over which the series of six experiments was carried out, the same groups of subjects did not serve in all the experiments. Some *inter*experimental differences between otherwise equivalent conditions, therefore, appear in our findings. All conclusions drawn and analyses made, however, are based on *intra*experimental differences.

General Procedure

Six experiments were carried out. Three forms of image degradation were used in various combinations either in pairs (Experiments 1, 2, 3, and 5) or in triples (Experiments 4 and 6) to evaluate their combined effect on the recognition of aircraft silhouettes. In Experiments 1 and 2, we sought to determine the effect of visual interference on either filtered or blocked stimuli. The determination of the effects of combining spatial frequency filtering and blocking is the main goal of Experiments 3 and 5. These two experiments differed in the order in which the other two degradations were applied. The effect of the addition of the visual interference was measured in Experiments 4 and 6, which were also opposite in the order in which the degradations were applied. Table 1 summarizes the conditions used in each of the six experiments and the overall design of this study.

The experimental procedure used in this study was fully automated. Subjects signed into each session by typing their names on the computer keyboard. This initiated a sequence of actions in which the experiment assigned for that session was loaded and the computer was configured to present the appropriate stimuli.

The subjects were instructed to identify the single stimulus presented in each experimental trial. A master stimulus list consisting of photographs of the 12 undegraded aircraft silhouettes was visible adjacent to the computer display throughout the experiment. A trial consisted of a sequence of visual displays on the CRT. The subject was first presented with a fixation stimulus consisting of the four, dimly lit, outline corners of the viewing region. This was followed by a 500-msec blank period. The stimulus display was then presented for a nominal 100 msec.[1] Following another 500-msec blank period, the four dim outline corners briefly appeared again on the display, instructing the subject to respond by depressing 1 of the 12 keys on the top row of the computer keyboard. When the subject responded, the fixation corners for the next trial were displayed and the cycle was repeated.

Table 1
The Order of Degradation in the Experiments

Experiment			
1	Filtering	Interference	
2	Blocking	Interference	
3	Blocking	Filtering	
4	Blocking	Filtering	Interference
5	Filtering	Blocking	
6	Filtering	Blocking	Interference

Because of the highly automated nature of our laboratory and the rapid exchange of stimulus materials between trials, large numbers (approximately 325) of trials were executed in each session. The obtained data, always measured in terms of the percentage of correct recognitions, were pooled only across the days of a particular experiment in order to provide the final values plotted on our figures. Within each experiment, all conditions were presented in random order each day to balance out any possible sequence effects. The stimulus conditions used for any trial were determined by random selection with replacement. We also included appropriate control conditions, as will be described later for each experiment.

Stimuli

The stimuli used in this recognition study were the same 12 solid aircraft silhouettes that were used in the preceding study; they are shown in Figure 1 of that paper (Uttal et al., 1995). These silhouettes were captured into the computer memory by a video camera and were subsequently processed by various combinations of the degrading algorithms to be described later. Each of the silhouettes was oriented so that the aircraft was pointing up toward the 12:00 position. All were approximately 1° of visual angle in height. One degree on our display subtended approximately 39 pixels horizontally and 37 pixels vertically. A stimulus silhouette was presented once in each experimental trial; it was located at the center of the viewing region. Since this is a 12-alternative recognition task, the chance performance level was 8.33%.

Our choice of these particular stimuli was determined by several factors. First, we wished to compare the results of this experiment with the results of the earlier discrimination experiment, in which the same aircraft silhouettes were employed (Uttal et al., 1995). Second, we specifically wished to use a different class of stimuli than did Harmon and Julesz (1973), in order to determine whether there would be a stimulus type effect. Third, although it may seem a priori that these stimuli may have different properties from those of the gray-scale image of a face, it should be remembered that the original binary silhouettes are transformed into gray-tone images by the degradation processes. Both blocking and spatial frequency transforms introduce nonhomogenities into the resulting images: As block size becomes larger and the cutoff frequencies become lower, there is more and more structure within the boundaries of what had originally been the homogeneous internal region of the silhouette. This is illustrated by Figure 2, which shows a sample set of stimuli used in this experiment.

The contrast of the images produced by the degradations is complex and cannot be simply expressed for each of the conditions of our experiment (which pool across all stimulus forms). The contrast depends on the particular stimulus form as well as on the specific degradations. When visual interference is added, the contrast situation is simpler. The peak intensity is the intensity of the individual interference dots (maximum screen luminance), and the least intensity is the dark background as determined by the veiling light. But this contrast value is irrelevant, since the stimulus image is varying in the complex ways just mentioned.

Stimulus Degradations

The stimuli were degraded by applying the "averaging over a region" algorithm (blocking), by blurring the image by removing spatial frequencies higher than a prespecified cutoff frequency from the spatial frequency spectrum of the Fourier transformed image (filtering), or by embedding the stimulus in randomly illuminated pixels (visual interference). A more detailed discussion of the three forms of visual interference is presented in the preceding paper (Uttal et al., 1995).

In addition to that discussion, in Figure 3 we now also provide a plot of the shape of the Butterworth filters used in the present experiment. These graphs illustrate the extreme sensitivity of the vi-

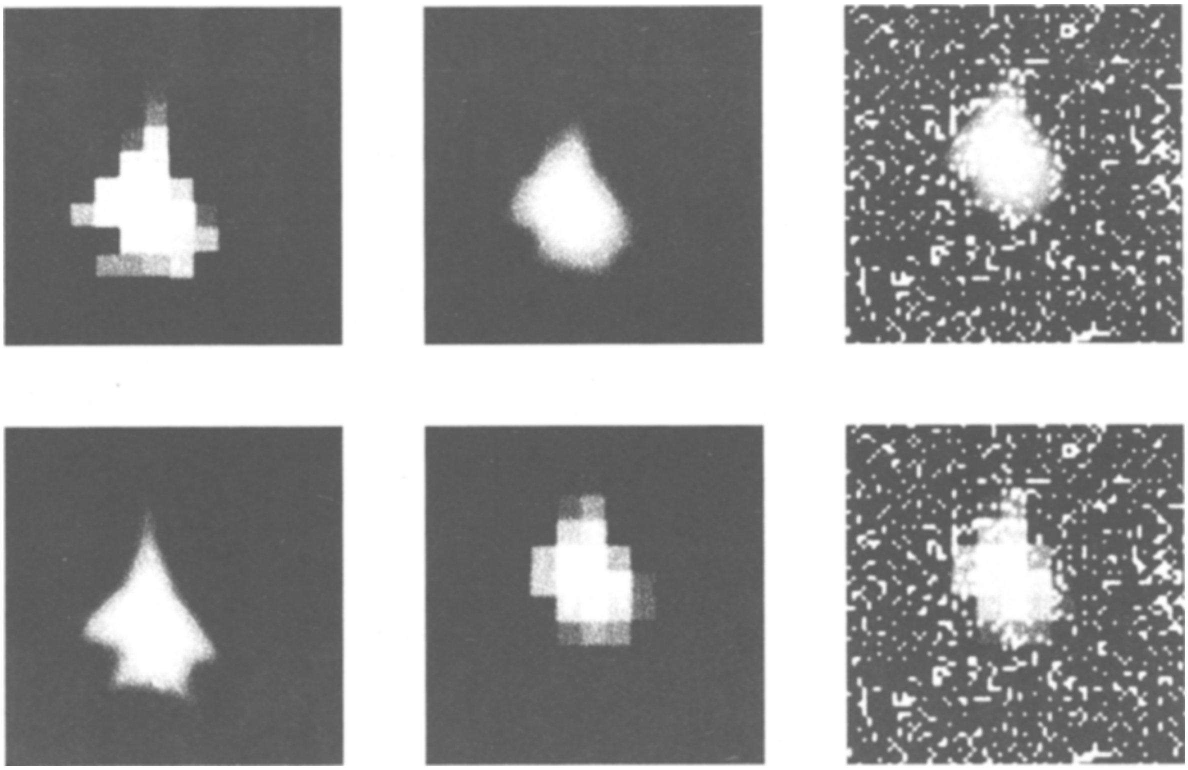

Figure 2. Sample stimuli showing the effect of combining degradations. From left to right, the upper row of three pictures shows the progressive effect of blocking (with a square averaging region of five pixels on a side), low-pass spatial frequency filtering (with a nominal cutoff limit of 1.48 cycles/deg), and finally with 20% random dotted visual interference superimposed. The same parameters of degradation are used in the lower three pictures, but the blocking follows the low-pass filtering. All these pictures have been enlarged to show the local detail of the degradation effects. To appreciate what the subjects saw, one must place these at a distance so that the stimulus object subtends 1° of visual angle.

sual system to the nominal cutoff limits of these filters. However, it must be appreciated that the dominant frequencies of the silhouettes lie in a narrow range where the four curves are very steep.

The values of the nominal cutoff limits of the Butterworth filters used in this study were chosen in pilot studies that showed a very sharp dropoff in recognition performance beyond the lowest value (1.30 cycles/deg). For nominal cutoff limits higher than that value, relatively little change occurred when only this type of degradation was applied. When filtering was combined with the blocking or interference, however, the joint effect was quite large and consistently different for each of the cutoff limits, as will be shown later in this article.

Apparatus

The experiments reported in this article were carried out on IBM-PC–compatible work stations with 486 Intel processors operating at 33 MHz. Observers were seated with their heads constrained by a chinrest so that their eyes were 64.7 cm from the face of the display. The entire experimental procedure was controlled by a computer program that randomly selected the stimuli, prepared the stimulus presentation sequence for each trial, collected each observer's responses (i.e., which key on the top row of the computer keyboard was depressed), and then performed a preliminary analysis of the data obtained in each hour-long session. If a recognition error was made, auditory feedback of the correct answer was given by a computer speech-generating system through earphones.

The viewing region was 5.08 × 5.08 cm (4.49° × 4.49° of visual angle) in extent. The CRT display itself (Tatung Model CM14SBS) was a raster scan, 34-cm (diagonal measurement) CRT with a full screen of 1024 × 768 pixels. The experimental room was indirectly lit by an incandescent bulb so that approximately 1 cd/m^2 fell on the screen as an ambient veiling light (Baker & Braddick, 1985;

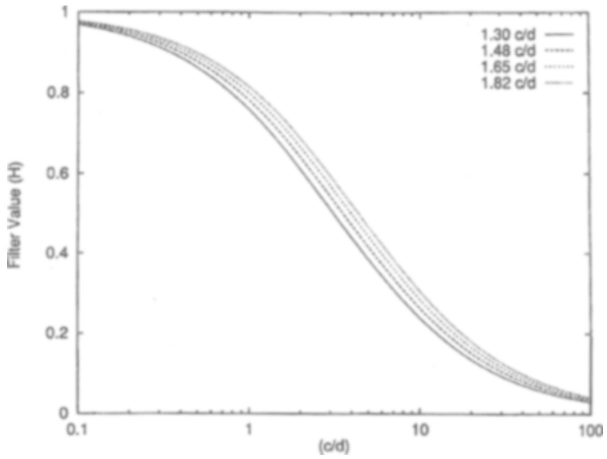

Figure 3. The shape of the four Butterworth filters used in this study. Though the curves are very similar, the differences between them have a substantial effect on recognition because of the relationship between the dominant spatial frequencies of the stimuli and the steep portions of the curves between 1 and 2 cycles/deg.

Farrell, Pavel, & Sperling, 1990; Groner, Groner, Muller, Bischof, & Di Lollo, 1993). The veiling light was measured by determining the amount of light reflected from a sheet of white paper at the surface of the display with a Tektronix J17 photometer equipped with a J1803 photometric sensing head.

The ambient veiling light also provided a constant lighting environment that stabilized the adaptation level of the subjects between trials. As a general calibration procedure, the luminance of a test pattern consisting of fully illuminated screen (i.e., all pixels set to white) was adjusted each day to 100 cd/m^2 with the veiling light present. Since illumination levels and stimulus durations are not critical in this recognition experiment, no attempt was made to precisely define either stimulus or perceptual durations or the luminance of the individual pixels of the stimulus.[2]

EXPERIMENTS AND RESULTS

Experiment 1: The Combined Effect of Spatial Frequency Filtering and Visual Interference

Experiment 1 was designed to determine the effects of the visual interference on the recognizability of stimuli that had been blurred by low-pass spatial frequency filtering. This was the first of the six different combination type experiments used in this study.[3] Pilot studies had indicated that the range of visual interference densities that was effective in this recognition study was much lower than the densities used in the previous discrimination study (Uttal et al., 1995). Therefore, only densities of 10% and 20% were used in this experiment. Control conditions in which only the 10% and 20% visual interference levels were combined with the unfiltered aircraft silhouettes were also included in the experimental design.

The stimuli were low-pass spatial frequency filtered to blur the images. Nominal cutoff frequencies of 1.82, 1.65, 1.48, and 1.30 cycles/deg of visual angle were selected as the values of this independent variable.

Figure 4 displays the results of Experiment 1. All data points in this and all subsequent figures contain standard error bars as measures of the variability of our results. The general pattern of results is a progressive decline in recognition performance as the cutoff limit of the low-pass filter decreases—that is, as higher and higher spatial frequencies are progressively removed. A nearly constant 15% difference between the curves representing the two densities of visual interference was obtained. The shape of the curves suggests a simple additivity of the independent effects.

Since each daily session in Experiment 1 consisted of approximately 325 trials and 6 subjects participated for four daily sessions, each point on this curve represents a mean performance score based on approximately 780 trials.

Experiment 2: The Combined Effect of Blocking and Visual Interference

In Experiment 2, visual interference was superimposed upon stimuli that had been blocked by averaging over square regions varying from 2 to 7 pixels on the side in steps of one pixel. The same two levels of visual interference were used as in Experiment 1—10% and 20%. Controls, in which these two visual interference levels were applied to unblocked images, were also inserted into the experimental design.

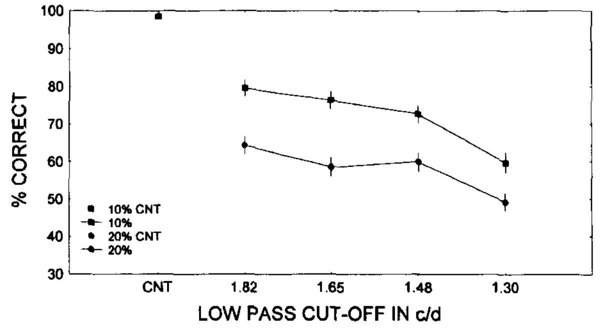

Figure 4. The results of Experiment 1. The two curves represent the results for stimuli that were low-pass filtered with the nominal cutoff limits specified along the horizontal axis. Ten percent and 20% random visual interference were added, respectively, to the stimuli, producing the two data curves. The control conditions are unfiltered stimuli embedded in the same two densities of visual interference. Recognition performance declines as the cutoff frequency is lowered. The joint effect of interference and spatial frequency filtering appears to be additive, with no sign of any interaction between the two forms of degradation. Standard error bars have been added in this and all subsequent figures. If the bars are not visible, the standard error was smaller than the symbol size.

The results of Experiment 2, shown in Figure 5, indicate that the differential effect of the two levels of visual interference was less and more irregular than that observed in Experiment 1. Nevertheless, the effects also appear to have been mainly additive as in Experiment 1. Adding visual interference to either of the other two types of degradation (blocking or filtering) always produced lower recognition performance for the 20% than for the 10% densities.

The absolute effect of increasing block size, however, was more detrimental than spatial frequency filtering. If one compares the results of Experiment 1 with those of Experiment 2, blocking can also be seen to have had a larger influence than spatial frequency filtering in the ranges of these variables that were used. This result is opposite to that obtained in the previously reported discrimination study. This comparison illustrates the task dependence of this type of experiment and suggests that the perceptual processes are not determined solely by the spatial frequency components of the stimulus. As we shall see later in this article, other task-related differences were also found.

Since each daily session in Experiment 2 consisted of approximately 325 trials and 7 subjects participated for four daily sessions, each point on this curve represents a mean performance score based on approximately 650 trials.

Experiment 3: The Combined Effect of Blocking Followed by Low-Pass Spatial Frequency Filtering

Experiment 3 was the closest analogue in the present study to the original Harmon and Julesz (1973) phenomenon itself. The aircraft silhouette stimuli were first blocked

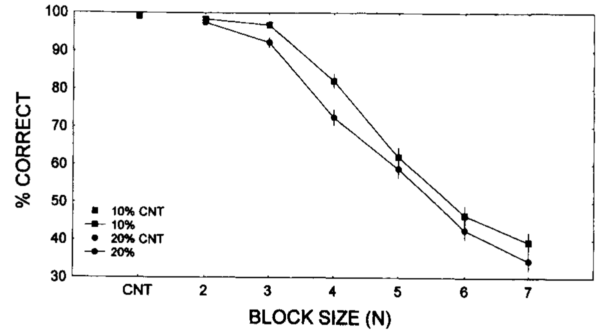

Figure 5. The results of Experiment 2. The two curves represent the results for stimuli that were blocked by averaging intensities across the square region sizes specified along the horizontal axis. (*N* is the length of a side of the block measured by the number of pixels.) Ten percent and 20% random visual interference were also added, respectively, to the stimuli, producing the data represented by the two curves. The control conditions are for unblocked stimuli embedded in the same two densities of visual interference. Recognition performance declines as the size of the block is increased. The joint effect of interference and spatial blocking also appears to be additive, with no sign of any interaction between the two forms of degradation.

and then low-pass spatial frequency filtered. The sizes of the averaging regions used to block the stimuli were 3, 5, and 7 pixels on the side, a subset of those used in Experiment 2.[4] The nominal values of the cutoff limit of the spatial frequency filter used in this experiment were the same as in Experiment 1—1.82, 1.65, 1.48, and 1.30 cycles/degree, respectively. In addition to the 12 combinations of these two factors, two sets of control conditions were necessary in order to make the results understandable. The first set consisted of stimuli that had only been area averaged at three block sizes, but not subsequently low-pass filtered. The second control consisted of stimuli that had only been low-pass filtered at four levels.

The results of Experiment 3 are presented in Figure 6. The control values for the condition in which the stimuli were only low-pass filtered are shown as a cluster of overlapping points. These control values were very high in all four graphs, indicating that the stimuli were recognized virtually perfectly when they were only low-pass filtered at any of the four nominal cutoff limits used in this experiment.

The solid line in Figure 6 represents the results obtained from the other set of control conditions—the stimuli that were only blocked and not subsequently low-pass filtered. Blocking alone produces a substantial decrement in recognition performance; performance progressively falls as the block size increases. Since no visual interference was used in this experiment, these values are higher than those obtained in Experiment 2. The four broken lines depict the results for the conditions that were blocked and then low-pass filtered at four spatial frequency limits. These data have been plotted as a function of the size of the averaging region used to block the stimuli.

A comparison of the broken lines with the solid lines in Figure 6 provides confirmation of the presence of the Harmon and Julesz (1973) phenomenon under at least some conditions. The performance levels for these aircraft silhouette stimuli, which were blocked and then filtered, are higher than those for the controls that were only blocked, except in the case of the smallest ($n = 3$) block size that was used. The outcome for the aircraft silhouettes is therefore the same as that reported for face stimuli in the demonstration by Harmon and Julesz. This result was found for all four values of the low-pass filter, with little change as a function of the nominal cutoff limit—a not unexpected outcome, given the very broad spectrum of the small dots of the interference.

However, as was previously shown (Uttal et al., 1995), when a *discrimination* task is used, the results are completely different. In that case, the additional image transformation produced by filtering of a blocked image results in a reduction in performance. In the present *recognition* experiment, to the contrary, *increased* recognizability resulted from the same sequence of degradations and the associated progressive image transformation. Therefore, there is a task dependency in this type of sequential degradation experiment that suggests that at least some component of the underlying mechanisms is not solely dependent on the spatial frequency characteristics (and, therefore, either the information content or some inhibitory masking effect) of the stimuli. If the results had been solely associated with available information (i.e., spatial frequency components), this dependency on task would not have occurred. That our results are qualitatively task dependent raises questions about the spatial frequency model of the underlying processes.

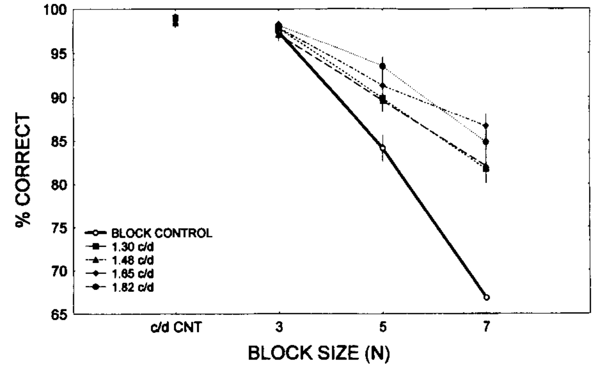

Figure 6. The results of Experiment 3. For this figure and all that follow, the data are arranged in the following manner. The horizontal axis plots the size of the blocking regions. The results for stimuli that were only blocked are the main control and are represented by the solid line. The results for stimuli that were both blocked and spatial frequency filtered are plotted parametrically, with broken lines, as a function of the nominal cutoff limit of the low-pass filter. The other set of controls, plotted with separate, but overlapping, points, includes stimuli that have only been low-pass filtered at the indicated (c/d = cycles per degree) nominal cutoff frequencies. The results shown in this figure are for stimuli that have been blocked and then low-pass filtered and presented without visual interference. This figure confirms the existence of the Harmon and Julesz (1973) phenomenon; stimuli sequentially degraded in this order are recognized better than those that are only blocked for the two larger block sizes.

Since each daily session in Experiment 3 consisted of approximately 325 trials and 8 subjects participated for five daily sessions, each point on this curve represents a mean performance score based on approximately 680 trials.

Experiment 4: The Combined Effect of Blocking Followed by Spatial Frequency Filtering and With Visual Interference

Experiment 3 conclusively confirmed that the Harmon and Julesz phenomenon could be produced by blocking and then spatial frequency filtering aircraft silhouette targets. In their classic 1973 paper, Harmon and Julesz also reported that the addition of random visual interference suppressed recognition of the blocked Lincoln face if the visual interference was "spectrally adjacent to the picture's spectrum." However, the opposite result was obtained by both Morrone et al. (1983) and Durgin and Proffitt (1993). The latter two studies suggest, to the contrary, that the addition of visual interference or grids can be used as an alternative to low-pass filtering in order to enhance recognition of a blocked image. However, in our earlier discrimination study and in Experiments 1 and 2, adding visual interference always suppressed recognition.

None of these studies, however, considered the effect of visual interference on a stimulus that had been blocked and then filtered. The empirical question then arises—what will be the joint effect of blocking, low-pass filtering, and visual interference (in that order) on recognition? The purpose of Experiment 4 was to answer that question. The design of the experiment was identical to that of Experiment 3, with the one exception that the 20% density of visual interference was added to each stimulus presentation. The same control conditions were also used.

The results of Experiment 4 are presented in Figure 7. The data are organized in exactly the same way as those in Figure 6. First, the results for the four control conditions that had neither been blocked nor filtered are plotted as a cluster of isolated points. Second, the results for the three control conditions in which the stimuli were only blocked are plotted with a solid line. Third, the results for the experimental conditions in which the stimuli were sequentially blocked and low-pass filtered are plotted with broken lines. All of these data, however, represent conditions in which 20% visual interference had been added.

The results in this case are very different from those obtained in Experiment 3. First, and not unexpectedly, the magnitudes of the effects are much larger. Adding visual interference made the task more challenging. The effect of each of the variables was magnified (except for the smallest, unfiltered block size), and all performance levels were reduced. More important than this overall quantitative difference is the change in the quality of the effect. Without the visual interference, blocking and then filtering the stimuli had resulted in an enhancement of the recognition scores. When interference was added, however, the same sequence of degradations now resulted in a diminishment in recognition performance in comparison with that for the control stimuli that were only blocked.

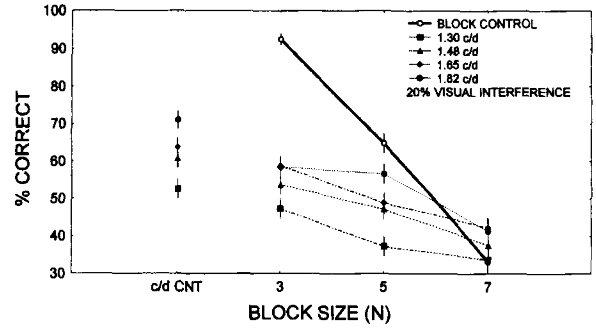

Figure 7. The results of Experiment 4. In this experiment, the stimuli were blocked and then low-pass filtered as in Experiment 3. However, 20% visual interference was added to the stimuli. The Harmon and Julesz (1973) phenomenon not only diminishes under these conditions but is reversed for the two smallest block sizes—the blocked and then filtered stimuli are seen less well than stimuli that have only been blocked in the presence of visual interference.

Since each daily session in Experiment 4 consisted of approximately 325 trials and 6 subjects participated for five daily sessions, each point on this curve represents a mean performance score based on approximately 500 trials.

Experiment 5: The Combined Effect of Low-Pass Spatial Frequency Filtering Followed by Blocking

One of the very surprising results obtained in our earlier paper (Uttal et al., 1995) was the commutativity of the order in which the blocking or the low-pass filtering had been applied. Either order produced the same effect. The purpose of Experiment 5 was to determine the effects of reversing the order of degradation that had been applied in Experiment 3. Other than that change—reversed order— Experiment 5 was identical to Experiment 3. No visual interference was present, and the same values for the size of the three averaging regions (3, 5, and 7 pixels, respectively) and the four nominal cutoff frequencies (1.82, 1.65, 1.48, and 1.30 cycles/deg, respectively) were used. The same set of control conditions was also used.

The results of Experiment 5 are plotted in Figure 8. The same arrangement of the data is used here as in the previous experiments. As expected, in the absence of the challenging visual interference, the range of effects is much smaller than in Experiment 4, but not too dissimilar to that for Experiment 3. Quite unexpected, however, is the *enhancement in performance as a result of blocking low-pass filtered stimuli.* It can be seen in Figure 8 that while there is very little effect of blocking a filtered stimulus when the averaging region is three or five pixels wide,[5] when the averaging region is seven pixels wide there is an enhancement of recognition performance. Thus, the enhancement of recognition performance does not entirely depend on the previously reported reduction in high-frequency edges by blurring—it can also be produced by sequential degradation in which these edges are introduced into a blurred image by subsequent blocking!

It is possible that this is due to the fact that the degraded images themselves become recognizable on the

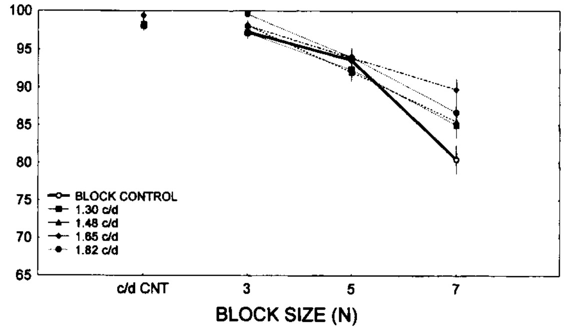

Figure 8. The results of Experiment 5. In this experiment, all of the conditions of Experiment 3 were maintained. However, the blocking and spatial frequency filtered degradations were applied in the opposite order—first low-pass filtering and then blocking. Surprisingly, there is also an improvement in the recognition performance at the largest block size.

basis of secondary cues generated by the transformations. This, however, does not detract from the suggestion that the present results are contrary to the raw predictions of any model based solely on spatial frequency component considerations. The same degradations were used in both the present and the preceding study (Uttal et al., 1995) and the results are quite different in a way that seems to depend solely on the task.

Since each daily session in Experiment 5 consisted of approximately 325 trials and 6 subjects participated for five daily sessions, each point on this curve represents a mean performance score based on approximately 500 trials.

Experiment 6: The Combined Effect of Low-Pass Spatial Frequency Filtering Followed by Blocking and With Visual Interference

Finally, Experiment 6 was carried out to determine the effect of visual interference on the reverse sequence (low-pass filtering and then blocking) used in Experiment 5. Except for the addition of the 20% density of randomly dotted visual interference, all conditions and controls in this experiment were identical to those in Experiment 5.

The results of Experiment 6 are shown in Figure 9. The results of this experiment are now seen to be comparable to those obtained in Experiment 4. That is, the magnitudes of the effects were much greater than in Experiment 5 for all except the smallest, unfiltered block size. Again, however, there is a qualitative change; where there had been a modest enhancement of the recognition scores when low-pass filtered stimuli were blocked in the absence of visual interference, now the result is a universal decrement in performance in the presence of 20% visual interference. In this experiment, as in Experiment 4, the difference between the blocked control stimuli and the experimental stimuli decreases as the nominal cutoff frequency of the filter gets higher. In other words, the difference, as expected, becomes greater as the image is progressively transformed.

Since each daily session in Experiment 6 consisted of approximately 325 trials and 6 subjects participated for six daily sessions, each point on this curve represents a mean performance score based on approximately 600 trials.

DISCUSSION

A Summary of the Empirical Results

In this study, we examined the effects of combining different kinds of image degradations on recognition performance. Our main results can be summarized as follows:

1. The nominal cutoff frequency of the low-pass spatial frequency filter had a relatively weak effect (in comparison with the effect of blocking) on the recognizability of a stimulus (Experiment 1).

2. The size of the averaging block had a relatively strong effect (in comparison with the effect of spatial frequency filtering) on the recognizability of a stimulus (Experiment 2).

(In our earlier discrimination study—Uttal et al., 1995—these relations were reversed. Blocking had less effect than did filtering. Thus, even the qualitative nature of our results depends on the task—among other salient variables.)

3. The superimposition of random dotted visual interference onto either the blocked or the filtered stimuli produced a decline in recognizability that was monotonically related to the density of the interference (Experiments 1 and 2).

4. When stimuli were first blocked and then low-pass filtered, the previously reported (Harmon & Julesz, 1973) paradoxical increase in recognizability occurred (Experiment 3).

5. When random dotted visual interference was superimposed on the same stimuli (degraded by blocking and then by low-pass filtering), the effect was completely reversed. Rather than the paradoxical increase in performance, in the presence of visual interference subsequent low-pass filtering of a blocked stimulus produced a substantially lower recognizability score than did blocking alone (Experiment 4).

6. Quite surprisingly, sequential degradation by low-pass filtering followed by blocking *also* produced an en-

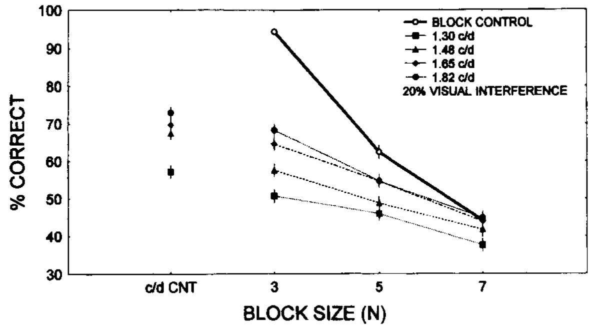

Figure 9. The results of Experiment 6. The experiment is identical in design to that of Experiment 5, with the exception that 20% random dotted visual interference was added. In this case, as in Experiment 4, the addition of the visual interference results in a reversal of the effects observed in Experiment 5—the enhancement in recognition performance observed there has become a substantial decrement except in the case of the largest block size.

hancement in recognition scores under the experimental conditions used in this study (Experiment 5). (This result, coupled with the findings obtained in Experiment 3, suggests that there is a certain degree of commutativity in recognition as well as in discrimination.)

7. Unsurprisingly, given the results of Experiment 4, when random dotted visual interference was superimposed on the stimuli that had been low-pass filtered and then blocked, the unexpected enhanced performance effect observed in Experiment 5 was completely reversed, and the surprising result with reversed order was also eliminated (Experiment 6).

The Issues

The results of this study highlight the true complexity of a visual phenomenon that had seemed at first glance to be straightforward and simple. The task seems to be as influential as the raw spectral characteristics of the stimulus.[6] Our results also illustrate the extreme importance of broad-scale, parametric examination of phenomena first observed in singular demonstrations. Although the data obtained in this study provide confirmation of the existence of the Harmon and Julesz (1973) phenomenon under some conditions, they also show that the effect is not universal (it does not occur in discrimination) and that it can occur in situations which their theory had not predicted—that is, filtering followed by blocking. Furthermore, under challenged viewing conditions (established by the superimposition of the random visual interference), the effect may be not only reduced, but reversed.

Similar caveats warning against premature adherence to a particular theoretical point of view are embodied in the different effects of visual interference reported here and from other laboratories. Morrone et al. (1983) reported that adding visual interference (visual noise, in their terminology) at spatial frequencies adjacent to those composing the stimulus image *enhanced* the recognizability of faces. Harmon and Julesz (1973) reported that "noise" in this same spatial frequency region *inhibited* recognition performance, whereas "noise" outside this range had little effect on perception. In this article and in the preceding one, we report that visual interference always lowers performance in spite of the fact that the punctate interference that we use has a broadband spatial frequency range. Durgin and Proffitt (1993) added lines at the boundaries of the blocks (thus adding high-frequency energy exactly at the critical band and place) and found enhanced recognition. The empirical discrepancies that have emerged between the results of the present study and those of previous efforts emphasize the necessity for more complete examination of the parameters involved in any kind of complex, multivariate study of visual perception.

The most important implication of these differences in empirical results, however, lies in their impact on the previous theoretical approach to the manifold of problems encountered when the perceptual effects of sequential or combined degradations are studied. Even though there may be no disagreement that the units of stimulus specification used in these studies may conveniently be those of the frequency domain space, the impulse toward a theory framed solely in those terms may be inappropriate. That is, spatial frequency theoretical explanations may be too narrow and incomplete even though spatial frequency measures may be useful and adequate.

We also note that the exceedingly difficult problem posed by any kind of recognition cannot be completely characterized by a theory that suggests that any particular spatial frequency band is the only one capable of permitting recognition. We can recognize a face, for example, on the basis of high spatial frequencies (Fiorentini, Maffei, & Sandini, 1983), intermediate bands (Hayes, Morrone, & Burr, 1986; Parker & Costen, 1993), low frequencies (Ginsburg, 1978), or even the few lines of a caricature. Any theory of shape recognition, therefore, must have an eclectic basis and must permit a variety of different perceptual processes to be invoked as the situation requires.

It may be that no universal theory of recognition is possible. Rather, the perceptual system may have to be viewed as a collection of processes, a "bag of tricks" in Ramachandran's (1985, 1990) words, able to either collectively or independently operate by means of a variety of processes on an ensemble of redundant cues, hints, or clues to produce whatever it is that we call "recognition."

Is a change in emphasis from what in another vocabulary would be considered to be a "bottom-up" spatial frequency domain explanation to a "top-down," perceptual organization analysis justified? To us, it seems unlikely (although we cannot yet reject the possibility) that a "bottom-up," spatial frequency based model could explain all of the contradictory results obtained. There are too many contradictions in the data, variations in the results that depend on the conditions of the experiments, and "paradoxes" of one kind or another, to currently conceive of any model based on a single dimension, attribute, or mode.

We offer the conjecture that while the frequency domain characteristics of the stimulus certainly do affect its perception, the spatial frequency spectrum is only one of several salient factors. These other factors include task, stimuli, order, and organization. This last term, from our perspective, is probably the most fundamental. It stresses the idea that the global configuration of the parts of a stimulus is the most salient attribute of an object—not any part or feature, and particularly not any feature from the frequency domain. This is not a novel idea, of course; others have championed the global or configurational precedence point of view. They include the Gestalt psychologists (e.g., Koffka, 1935), Lockhead (1972), Navon (1977), and the other modern students of visual perception who contributed to the seminal volume by Kubovy and Pomerantz (1981) stressing the importance of perceptual organization.

This alternative point of view also suggests that there are many different ways in which the information of the configuration may be communicated to the observer. While features such as edges, corners, "geons," "textons," or spatial frequency components may play a useful role in describing stimuli, the complex and difficult-to-quantify factor to which we refer as arrangement may actually be the most meaningful perceptual parameter. If so, we must once again acknowledge the enormous insight of the classic Gestalt students of perception.

REFERENCES

BACHMANN, T. (1991). Identification of spatially quantized tachistoscopic images of faces: How many pixels does it take to carry identity? *European Journal of Cognitive Psychology*, **3**, 87-103.

BAKER, C. L., & BRADDICK, O. (1985). Temporal properties of the short-range process in apparent motion. *Perception*, **14**, 181-192.

DURGIN, F. H., & PROFFITT, D. R. (1993). *Perceptual response to visual noise and display media* (Final Report, NASA-NAG2-814). University of Virginia, Department of Psychology.

FARRELL, J. E., PAVEL, M., & SPERLING, G. (1990). The visible persistence of stimuli in stroboscopic motion. *Vision Research*, **30**, 921-936.

FIORENTINI, A., MAFFEI, L., & SANDINI, G. (1983). The role of high spatial frequencies in face perception. *Perception*, **12**, 195-201.

GINSBURG, A. P. (1978). *Visual information processing based on spatial filters constrained by biological data* (Report No. AMRL-TR-78-129, Vols. 1–2). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

GRONER, R., GRONER, M. T., MULLER, P., BISCHOF, W. F., & DI LOLLO, V. (1993). On the confounding effects of phosphor persistence in oscilloscopic displays. *Vision Research*, **33**, 913-918.

HARMON, L. D., & JULESZ, B. (1973, June 15). Masking in visual recognition: Effects of two-dimensional filtered noise. *Science*, **180**, 1194-1197.

HAYES, T., MORRONE, M. C., & BURR, D. C. (1986). Recognition of positive and negative bandpass-filtered images. *Perception*, **15**, 595-602.

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

KUBOVY, M., & POMERANTZ, J. R. (1981). *Perceptual organization*. Hillsdale, NJ: Erlbaum.

LOCKHEAD, G. R. (1972). Processing dimensional stimuli: A note. *Psychological Review*, **79**, 410-419.

MORRONE, M. C., BURR, D. C., & ROSS, J. (1983). Added noise restores recognizability of coarse quantized images. *Nature*, **305**, 226-228.

NAVON, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, **9**, 353-383.

PARKER, D. M., & COSTEN, N. P. (1993). Spatial frequency and internal masking in face images. *Perception*, **22** (Suppl.), 48.

PELLI, D. G., & ZHANG, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, **31**, 1337-1350.

RAMACHANDRAN, V. S. (1985). Guest editorial. *Perception*, **14**, 97-103.

RAMACHANDRAN, V. S. (1990). Visual perception in people and machines. In A. Blake & T. Troscianko (Eds.), *AI and the eye* (pp. 21-78). Chichester, England: Wiley.

UTTAL, W. R., BARUCH, T., & ALLEN, L. (1995). The effect of combinations of image degradations in a discrimination task. *Perception & Psychophysics*, **57**, 668-681.

NOTES

1. The 100-msec duration was chosen to avoid any opportunity for eye movements or extended scrutiny of the image. The actual duration of the effective stimulus was, of course, longer because of the impulse function of the eye. We have never obtained any results that suggested that this duration was a critical factor in the kind of experiment described here.

2. Another source of uncertainty is the gamma function of the display. We measured this functional relationship between the digital intensity code and the screen luminance. It is approximately linear, but not exactly. At low coded levels, there is a significant discrepancy. This can be corrected in several ways, as is described in an article by Pelli and Zhang (1991). Our procedure included a veiling light that partially served this function. However, we do not believe that this changed the qualitative nature of our results in any substantial manner.

3. The single degradation experiment was not necessary in this experimental design. These data are available by tabulating the control conditions used in Experiments 3 and 5 and in our earlier paper (Uttal et al., 1995). In brief, all single-variable experiments produced monotonic declines in recognition performance as the degradation (visual interference, low-pass filtering, or blocking) increased. However, the magnitudes of the effects varied from one type of degredation to another.

4. As the number of variables increased from two in Experiments 1 and 2 to three in Experiments 3–6, it was necessary to reduce the number of values of each variable so that an adequate sample size for the remaining conditions could be obtained.

5. These results would be even more remarkable if the block-only control score for the 5 × 5 pixel condition was not so high. The difference between this particular block control condition in Experiment 5 and Experiment 3 is inexplicable, but it may have been influenced by the different subject populations in the two experiments, among other unknown factors.

6. In a subsequent report we shall show that the type of stimulus material is not influential in determining the results for the recognition task.

(Manuscript received June 24, 1994;
revision accepted for publication December 16, 1994.)