ELSEVIER

Acta Psychologica 91 (1996) 121-129

acta psychologica

# Inattention and the perception of visual feature conjunctions

Vito Modigliani *, Richard D. Wright, David S. Loverock

*Department of Psychology, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6*

Received 17 May 1994; revised 22 September 1994; accepted 7 December 1994

**Abstract**

Visual processing of objects in the absence of focused attention appears to be limited. We varied the degree of attention, or visual processing, that observers paid to objects using an instruction set manipulation. In 2 experiments, subjects performed tasks that required superficial or detailed visual analysis of the objects involved. In subsequent recognition tests, information about conjunctions of shape and internal color/texture pattern was limited when only superficial visual analysis was required to encode the object. This implies that the degree of visual processing, during object encoding affects the likelihood that feature conjunctions are incorporated into the visual representation of these objects.

*PsycINFO classification:* 2323, 2346

*Keywords:* Attention; Inattention; Visual perception; Feature conjunctions; Encoding

## 1. Introduction

When looking at two spatially superimposed objects or events, we are able to attend to one while filtering out most of the information about the other. For example, when two spatially superimposed videotaped events were shown and observers attended to one of them, they reported being unaware of several aspects of the other one (Neisser and Becklen, 1975). A similar result was obtained with spatially superimposed static scenes (Goldstein and Fink, 1981) and with overlapping figures (e.g., Butler and McKelvie, 1985; Rock and Gutman, 1981). In the overlapping figures task, observers are shown

* Corresponding author: E-mail: modiglia@sfu.ca, Tel: + 1 604 291-3744.


*SSDI* 0001-6918(94)00058-1

two interleaved figures of different colors (e.g., red and green) and asked to attend to only one (e.g., the red one). In this task, observers do not appear to process the shape of the unattended (green) figure in enough detail to recognize it later. Further, processing of local features of the unattended figure is also quite limited (Wright et al., 1993). Thus, objects and events appear to undergo only limited visual processing under conditions of inattention.

Treisman and Gelade (1980) proposed a model that may account for the role of attention (and inattention) in object perception. In particular, they claimed that feature information (e.g., color and orientation) is processed preattentively and in parallel by lower level perceptual mechanisms before being integrated or conjoined at a particular location to form a perceptual object. Integration was said to occur in one of three ways: (1) by focusing attention at a particular location in space and thereby invoking automatic integration of all features indexed to that location into a perceptual object; (2) in the absence of focused attention but influenced by general knowledge of properties of particular objects (e.g., the sky should be blue, carrots should be orange, apples should be roundish); (3) in the absence of focused attention and general knowledge about the objects, thereby producing incorrect or illusory conjunctions. Thus, when the visual system processes novel stimuli, such that feature integration cannot be guided by prior experience with them, focused attention is necessary to ensure accurate conjunctions of features.

At a more general level, it may be assumed that an object's representation is a function of perceptual analyses (Rock, 1983; Rock and Gutman, 1981; Rock et al., 1972; Sutherland, 1968; Ullman, 1984), and that these depend on the task that is being carried out. Thus, in the overlapping figures task, some visual processing of the to-be-ignored figure must occur in order to filter it out (e.g., its color may be used as the filtering criterion). Yet such analysis is not sufficiently detailed to create a representation that contains enough information about the figure's shape (Rock and Gutman, 1981) and local feature information (Wright et al., 1993) to enable later recognition of it. Equally, the representation does not contain information about the conjunction of features of the unattended figure.

In the line-length task (Rock et al., 1992), observers were required to determine whether the lengths of two bisecting lines, one horizontal and one vertical, were equal. As they did so, they were unable to recognize the familiar shape of a colored pattern presented within the display. In other words, they were unable to process the shape/color conjunction of the unattended stimulus, even though it was presented in roughly the same spatial region as the attended stimulus. Thus, it appears that correct conjunctions of novel object features are unlikely to occur when attention to these objects is limited.

In both the overlapping figures and line-length tasks, the degree of attentional processing was varied by having two or more objects present in the visual field, attention being directed to some of these and not others. Concerns may be raised about the precision of attentional alignment and about spatial scaling. More specifically, in the overlapping figures task, the unattended object is only approximately (spatially) superimposed with the attended object. In the line-length task, the unattended object is presented at the local level within a globally encompassing pair of lines. There is some debate, however, about whether global and local spatial scales can be attended to with equal efficiency at the same time (e.g., Navon, 1977; see Kimchi, 1992, for a review of the global/local issue). Also, in the tasks under discussion, subjects presumably filter out irrelevant object information intentionally (cf. Watanabe, 1988).

In the present paper we are concerned with the effect of different degrees of attentional processing in the analysis of *single* objects. With single stimuli, spatial alignment and global/local differences become irrelevant. Further, degree of attentional processing was manipulated indirectly, so that subjects were not specifically directed to attend or filter out any information intentionally. In the incidental condition, as described below, the task required subjects to detect whether an object was present or not. In this case, the object would receive a superficial analysis, only enough to reach a decision about figure and ground. Such analysis would be unlikely to include feature conjunctions. In the deliberate condition, also described below, objects were to be discriminated from one another, in which case the features of each object, and their conjunctions, would be more likely to be noted and encoded.

## 2. Experiment 1

The purpose of this experiment was to measure retention of conjunctions of an object's attributes as a function of the amount of attention directed to the object. Two presentation conditions, Deliberate and Incidental, were used. In the Deliberate Condition, subjects were instructed to examine each stimulus in order to commit it to memory. In the Incidental Condition, subjects saw multiple copies of each stimulus, but were only required to count them. The stimuli were otherwise identical to those used in the Deliberate Condition. Presentation was followed by a recognition test that was the same for all subjects. It was expected that, if attention must be explicitly directed toward a figure for attribute conjunctions to occur, recognition performance would be poorer in the Incidental Condition than in the Deliberate Condition. On the other hand, if merely noting the presence of an object were sufficient to automatically conjoin its attributes, then recognition performance in the two conditions would be similar.

### 2.1. Method

#### 2.1.1. Subjects

Twenty-four Simon Fraser University undergraduates participated in the experiment, one half randomly assigned to the Deliberate Condition, the other half to the Incidental Condition.

#### 2.1.2. Apparatus and stimuli

Stimulus displays and data collection were carried out by a 286-based microcomputer. Stimuli were displayed on an NEC Multisync color monitor, and were viewed from a distance of 40 cm. Individual stimuli subtended about $3 \times 3°$ and consisted of closed figures that varied on two dimensions – shape and color/texture. We collapsed color and texture into a single property called 'internal aspect'. There were 10 highly discriminable values of each dimension. The 10 shapes were a circle, a square, a

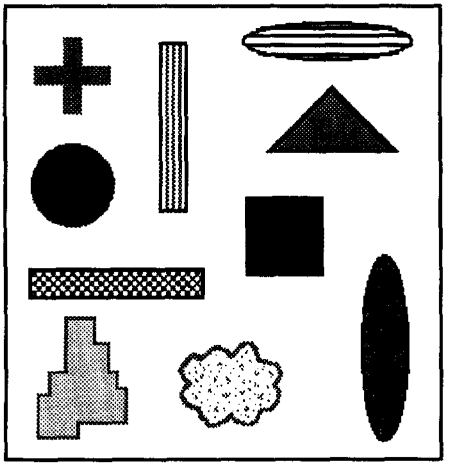

Fig. 1. Examples of the stimuli used in the experiments. The variations of grey stippling represent color differences.

triangle, a cross, a horizontal ellipse, a vertical ellipse, a round-edged figure, a square-edged figure, a horizontal rectangle, and a vertical rectangle. The 10 internal aspect values were solid blue, yellow, green, orange, and purple; and texture patterns with either randomly superimposed dark blue filled circles, randomly superimposed shapes resembling clouds, a striped background, a horizontal/vertical grating background, or randomly superimposed white filled circles (see Fig. 1).

Forty of the 100 possible combinations of shape and internal aspect were selected for use. In particular, four sets of 10 stimuli were generated so that each of the stimuli within a set differed in shape as well as in internal aspect. Therefore, when a set of 10 stimuli was presented, subjects saw each of the 10 shape values and each of the 10 internal aspect values. The four sets of stimuli differed in terms of the combinations or conjunctions of shape and internal aspect. Across the four sets, each shape was paired with a different internal aspect. For instance, if in Set 1 the circular shape's internal aspect was a solid yellow color, then in Set 2 it might have been the horizontal/vertical grating background, in Set 3 it might have been a solid green color, and in Set 4 it might have been the background with randomly superimposed dark blue filled circles.

### *2.1.3. Procedure*

In the Deliberate Condition, each stimulus was shown singly, centered on the screen, for 1700 ms. This presentation time is the same as the average response time in the Incidental Condition (see below). Subjects were instructed to study each stimulus carefully in order to facilitate their performance on a subsequent recognition test. They were then shown the 10 stimuli from one of the sets described previously, one stimulus at a time. The first stimulus was preceded by the word READY and the last followed by the word END, both presented in the center of the screen, also for 1700 ms. After the word END, the experimenter informed the subject that the recognition test was being prepared and that it would be ready in 30 s. The test consisted of 20 stimuli, shown one at a time. The first stimulus was again preceded by the word READY and the last followed by the word END. For each test stimulus, subjects were instructed to make an old/new forced-choice responses by pressing one of two buttons. The stimulus remained on display until the response was made. Ten of the test stimuli were old (had already been seen) and 10 were new. Presentation order of old and new stimuli was random. Each of the four stimulus sets was used equally often as the old and new sets. For example, if a subject saw the stimuli in Set 1 initially, the recognition test included Set 1 and either Set 2, Set 3, or Set 4 stimuli.

In the Incidental Condition, multiple copies of a stimulus, ranging from six to ten in number, were displayed on the screen. Subjects were instructed to count the number of stimuli presented within each of 10 successive displays as quickly as possible. The first display was again preceded by the word READY and the last followed by the word END. There were no instructions to study each stimulus explicitly, and subjects were not informed about the subsequent recognition test. Each display consisted of multiple copies of the same stimuli as used in the Deliberate Condition, located randomly on the screen. For example, if a yellow circle was presented in the Deliberate Condition, the corresponding display in the Incidental Condition might consist of seven identical yellow circles presented simultaneously at different locations. Counting responses were made by pressing one of five buttons labeled 6, 7, 8, 9 and 10. Pilot work had shown that the average response time in this task was 1700 ms. After the word END following presentation, the experimenter informed subjects that we were interested in what they could remember about the stimuli they had just seen, and that a recognition test would therefore follow. Care was taken to engage these subjects in conversation for about the same length of time as subjects in the deliberate condition. The recognition test was prepared and given 30 s following the end of presentation. Each subject in the Incidental Condition was yoked to one in the Deliberate Condition and received exactly the same test. Thus, yoked subjects differed only in the initial task, one of which emphasized deliberate encoding, whereas the other focused on counting, thereby de-emphasizing feature encoding.

### 2.2. Results and discussion

The mean recognition accuracy across subjects was much higher in the Deliberate (83%) than in the Incidental (57%) Condition ($t(22) = 6.41$, $p < 0.001$). In addition, mean accuracy in the Incidental Condition was significantly higher than the 50% (chance) level ($t(11) = 2.48$, $p < 0.05$). Thus, deliberate encoding of the stimulus objects led to significantly more accurate performance on the subsequent recognition test than incidental encoding (counting). Performance in the latter condition was only slightly, although significantly, better than chance.

## 3. Experiment 2

One potentially important difference between the two conditions in Experiment 1 was the presentation of multiple copies of the stimulus in the Incidental Condition but not in the Deliberate Condition. It is possible that encoding the conjunction of shape and internal aspect of stimuli may be affected by whether single or multiple copies are displayed. If true, the difference in performance between the two conditions could be attributable to stimulus number, rather than to deliberate versus incidental encoding. The second experiment was carried out to control for this possibility by examining the effects of deliberate and incidental encoding in single and multiple stimulus presentation conditions.

In a $2 \times 2$ factorial design stimulus number (single vs. multiple copies of a stimulus) was crossed with type of encoding task (deliberate vs. incidental). Thus, the displays consisted either of a single stimulus or of 6–10 copies of a stimulus. In the Single/Deliberate Condition stimuli and procedure were identical to those of the Deliberate Condition in Experiment 1. Thus, single stimuli were presented, and the subjects were given deliberate memory instructions.

In the Single/Incidental Condition, stimulus displays were similar to those of the Single/Deliberate Condition, with one difference. This was the addition of six to ten small ($0.5 \times 0.5°$) white filled squares randomly positioned around the periphery of the stimulus. Subjects were required to count the small squares as quickly as possible and respond by pressing the appropriate button. In doing so, subjects had to look at the object, but we did not expect them to encode the conjunction of shape and internal aspect information to the same extent as subjects in either the Single/Deliberate Condition, or the Multiple/Deliberate Condition described next.

In the Multiple/Deliberate Condition, stimulus displays were identical to those of the Incidental Condition in Experiment 1, i.e., multiple copies of a stimulus. However, unlike the Incidental Condition of Experiment 1, subjects were instructed to concentrate on any one stimulus and disregard the fact that multiple copies of it were presented. As in the Single/Deliberate Condition, instructions emphasized memory coding for the forthcoming recognition test. Finally, the Multiple/Incidental Condition was identical to the Incidental Condition in the first experiment. Subjects counted the stimuli as quickly as possible, and there was no mention of the recognition test. All other aspects of this experiment were identical to those of Experiment 1. Twelve subjects were assigned to each of the 4 cells of the design.

### *3.1. Results and discussion*

Mean accuracy scores are shown in Table 1. A $2 \times 2$ ANOVA was carried out as a function of type of encoding task (Deliberate vs. Incidental) and stimulus number

Table 1
Mean percentages of correct rsponses across subjects as a function of stimulus number and encoding task

| Stimulus number | Encoding task | | |
|---|---|---|---|
| | Deliberate | Incidental | Mean |
| Single | 80 | 56 | 68 |
| Multiple | 71 | 58 | 64 |
| Mean | 76 | 57 | |

(Single vs. Multiple copies). There was a significant main effect of encoding task, mean accuracy in the Deliberate Conditions (76%) being significantly higher than in the Incidental Conditions (57%) ($F(1,44) = 33.85$, $p < 0.0001$). Neither stimulus number ($F(1,44) = 1.48$, $p > 0.05$), nor the interaction ($F(1,44) = 2.98$, $p > 0.05$), were significant. As in Experiment 1, mean accuracy in the Incidental Conditions (57%) was significantly different from the 50% (chance) level ($t(23) = 2.94$, $p < 0.05$). These findings replicate those of the first experiment and demonstrate that the perception of conjunctions of shape and internal aspect information depends on encoding task type, regardless of stimulus number.

## 4. General discussion

The results of both experiments indicate that, in the recognition test, information about the conjunction of shape and color/texture was not readily available when the to-be-remembered object had been processed in only an incidental manner. Conversely, deliberate processing resulted in very good recognition performance. These results indicate that merely noticing the presence of a novel object is *not* sufficient for the accurate integration of its features. As previously suggested, the incidental task can be carried out solely on the basis of a figure/ground discrimination. As such, detailed encoding of the object's features and their conjunctions was not necessary and, as the data show, was not carried out. In the deliberate condition, however, in which subjects explicitly prepared for the recognition test, a relatively higher level of visual analysis was necessary and, again as shown by the data, was carried out.

These results are consistent with accounts according to which visual representations contain the results of perceptual analyses carried out on the stimulus. Sutherland (1968) was one of the first to make a proposal of this kind. Similar proposals were made by Pylyshyn (1973), Rock (1983), Rock and Gutman (1981), Rock et al. (1972), and Ullman (1984), among others. Rock et al. (1972), for example, considered the mere registration of a stimulus insufficient for accurate encoding of most of its properties. Ullman (1984) distinguished between universal visual routines and more specialized ones, the latter being invoked as a function of the goal at hand. These approaches agree in suggesting that visual representations contain the results of perceptual analyses, and that these may differ in the amount of detail they extract from the stimulus. The results of the present experiments indicate that, in the deliberate condition, object representations included information about the conjunction of features, but this was not the case in the incidental condition. We attribute this finding to a more detailed analysis of the stimulus in the former than in the latter condition.

The present results extend conclusions obtained in tasks, such as the overlapping figures and line-length tasks, in which the unattended object was one of two or more objects simultaneously present in the visual field. Neither spatial alignment, not global/local differences are relevant factors in the analysis of single objects. Also, the tasks did not require subjects to explicitly filter out unattended objects. Instead, the deployment of attentional processing was only implicitly under task control. Thus, the results indicate that, when examining *single* objects, people can deploy attention, and thereby vary the level of visual processing, to different degrees, yielding visual representations that contain different amounts of information. Thus, the present results show that attention to single objects, and the degree to which they are examined, can be manipulated, and that such manipulations affect the probability of encoding feature conjunctions.

The discussion so far supposes that, in the type of task used here, the effects of attention and inattention are related to the adequacy or inadequacy of perceptual encoding. However, it is possible that poor recognition performance could result from inadequacies in the retention of originally correctly perceived information. This issue is a very difficult one, as a recognition task is a test of both perception and memory (Bartlett, 1932). Nevertheless, there is some evidence suggesting that in the type of task used here poor recognition may be due to perceptual rather than memory failure (Rock and Gutman, 1981). Rock and Gutman conducted an overlapping figures experiment in which familiar shapes (Christmas tree, house) were presented in two of ten trials. When these shapes were attended to, recognition accuracy was 85%. On the other hand, when these shapes were *not* attended to, recognition accuracy was only 10%. According to the authors it is reasonable to expect that, if such familiar shapes had been initially perceived, they would have also been recognized (cf. Treisman and Gelade's (1980) conjunction of familiar objects' features). Rock and Gutman also conducted a similar experiment in which trials involving the familiar shapes were immediately followed by a blank field. When this blank field was presented, subjects were required to report all that they could remember about both figures on the previous trial. Such an immediate-report technique should eliminate memory effects, yet 89% of subjects did not report seeing the familiar shapes. Moreover, even when subjects were shown these shapes, they said that they had not seen them. Therefore, it appears that the shapes of the unattended figures were not initially perceived. Similarly, in an earlier experiment, Rock et al. (1972, Exp. V) found that recognition remained poor even when the recognition test was given only 200 ms after exposure of the target figure.

Because of such data, we believe that in our experiments differences in recognition performance were likely due to encoding failure rather than memory loss. Regardless of the underlying mechanism, however, the perception/retention of objects under differing degrees of visual processing warrants further study, so that the mechanisms involved and the creation and retention of visual representations may be further elucidated.

## Acknowledgements

This research was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) grants awarded to VM and RDW, and by an NSERC Postgraduate Scholarship awarded to DL. We thank Malcolm Toms and Richard Blackwell for their technical assistance. We also thank Christian Richard and Linda Clarke for their help in running the experiments.

## References


Bartlett, F.C., 1932. Remembering: A study in experimental and social psychology. New York: Macmillan.

Butler, L. and S.J. McKelvie, 1985. Processing of form: Further evidence for the necessity of attention. Perceptual & Motor Skills 61, 215–221.

Goldstein, E.B. and S.I. Fink, 1981. Selective attention in vision: Recognition memory for superimposed line drawings. Journal of Experimental Psychology: Human Perception and Performance 7, 954–967.

Kimchi, R., 1992. Primacy of holistic processing and global/local paradigm: A critical review. Psychological Bulletin 112, 24–38.

Navon, D., 1977. Forest before the trees: The precedence of global features in visual perception. Cognitive Psychology 9, 353–383.

Neisser, U. and R. Becklen, 1975. Selective looking: Attending to visually specified events. Cognitive Psychology 7, 480–494.

Pylyshyn, Z.W., 1973. What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin 80, 1–24.

Rock, I., 1983. The logic of perception. Cambridge, MA: MIT Press.

Rock, I. and D. Gutman, 1981. The effect of inattention on form perception. Journal of Experimental Psychology: Human Perception and Performance 7, 275–285.

Rock, I., F. Halper and T. Clayton, 1972. The perception and recognition of complex figures. Cognitive Psychology 3, 655–673.

Rock, I., C.M. Linnett, P. Grant and A. Mack, 1992. Perception without attention: Results of a new method. Cognitive Psychology 24, 502–534.

Sutherland, N.S., 1968. Outlines of a theory of visual pattern recognition. Proceedings of the Royal Society of London (Series B) 171, 297–341.

Treisman, A. and G. Gelade, 1980. A feature-integration theory of attention. Cognitive Psychology 12, 97–136.

Ullman, S., 1984. Visual routines. Cognition 18, 97–159.

Watanabe, T., 1988. Effect of irrelevant differences as a function of the relationship between relevant and irrelevant dimensions in the same-different task. Journal of Experimental Psychology: Human Perception and Performance 14, 132–142.

Wright, R.D., A.N. Katz and E.A. Hughes, 1993. Inattention and the perception of visual features. Acta Psychologica 83, 225–235.