Journal of Experimental Psychology:
Human Perception and Performance
1978, Vol. 4, No. 4, 565–572

Cognitive Determinants of Fixation Location During Picture Viewing

Geoffrey R. Loftus
University of Washington

Norman H. Mackworth
Stanford University

This experiment involved the question of where human observers look in a picture. The results indicated that observers fixate earlier, more often, and with longer durations on objects that have a low probability of appearing in a scene (e.g., an octopus in a farm scene) than on objects that have a high probability of appearing (e.g., a tractor in a farm scene). These findings (a) imply a role of cognitive factors in peripheral visual processing and (b) suggest a possible relationship between the nature of information initially acquired from a picture and subsequent recognition memory for that picture.

When an observer views a static visual scene such as a picture, his or her visual scanning is a discrete process that is broken into periods of about 300 msec during which the eye is relatively immobile (fixations), separated by quick jumps of the eye from place to place (saccades). Visual information processing is assumed to take place during fixations, whereas vision is essentially suppressed during saccades (Latour, 1962; Volkman, 1976).

The present research concerns factors that determine the sequence of fixation locations during picture viewing by human observers. Early work on this topic has indicated that fixations are not distributed randomly over a picture, but rather that a relatively large proportion of the fixations is allocated to a relatively small portion of the scene (Buswell, 1935). A number of more recent studies have demonstrated that the gaze is attracted to "informative areas" of a picture, where informative is defined in terms of subjective "informativeness ratings" by independent observers (Antes, 1974; Mackworth & Bruner, 1970; Mackworth & Morandi, 1967; Pollack & Spence, 1968).

The notion that observers look at informative areas in a picture is an interesting one. A question that immediately arises, however, concerns what it is about an area of a picture that determines the informativeness of that area; that is, what are the underlying psychological mechanisms that cause one area of a picture both to be rated as more informative and also to be fixated more often than another? A number of clues and speculations have been offered in response to this issue. First, informal observation has suggested that those areas of a picture that are fixated and/or rated as informative tend to be areas of physical discontinuity (Mackworth & Morandi, 1967) or objects whose presence is in some way surprising given the rest of the picture (Yarbus, 1967). Second, Berlyne (1958) has reported that observers tend to allocate the majority of fixations to the less redundant of two simple visual stimuli when redundancy is defined either in physical terms (e.g., a regular vs. an irregular checkerboard) or in cognitive terms (e.g., an

This research was supported by National Science Foundation Grants GB 39615 and BNS76-23367 to Geoffrey Loftus.

We thank V. J. Dark for collecting subjective ratings, D. L. Golding for carrying out the eye movement data reduction, and L. Carter-Saltzman, D. Gentner, E. F. Loftus, C. M. MacLeod, P. Mathews, T. O. Nelson, and S. E. Poltrock for their comments on the manuscript.

Requests for reprints should be sent to Geoffrey Loftus, Department of Psychology, University of Washington, Seattle, Washington 98195.

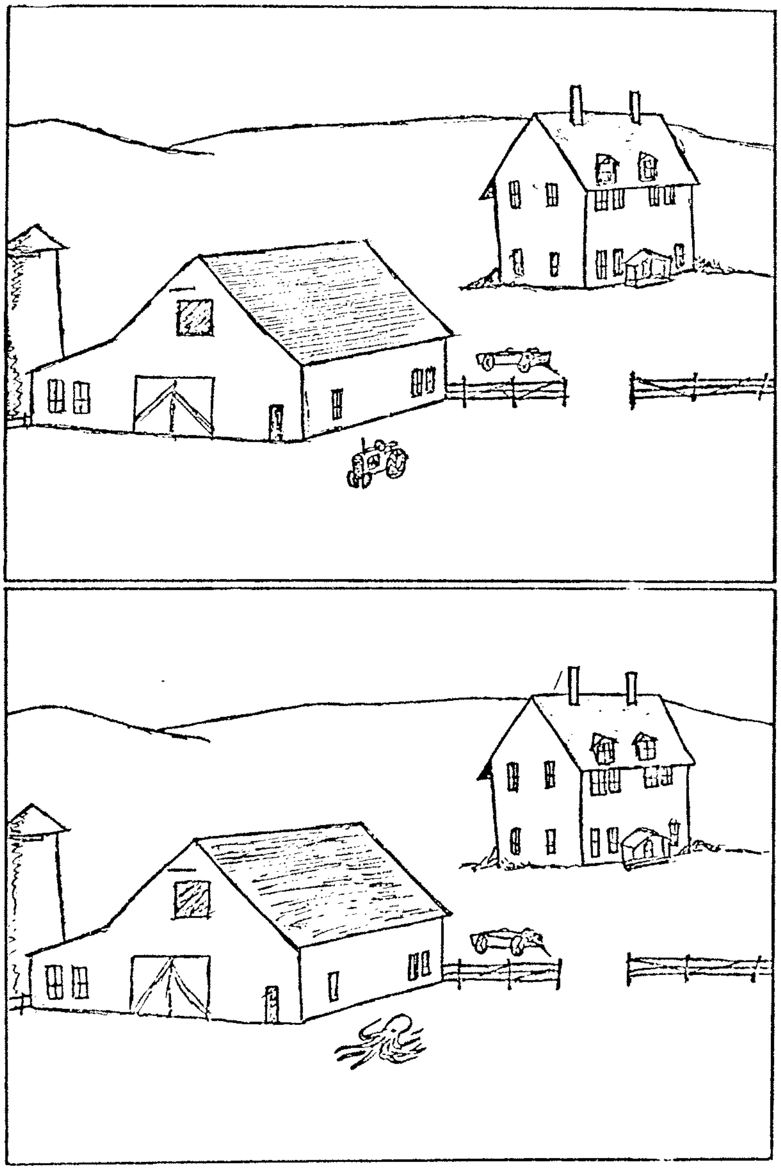

Figure 1. An example of a noninformative object (the tractor in the top panel) and a corresponding informative object (the octopus in the bottom panel).

ordinary elephant vs. a creature that is part elephant, part rabbit, and part bird).

Such evidence suggests that informativeness might appropriately be characterized in terms of intrastimulus redundancy or predictability. Within this broad characterization, it appears that there are at least two distinct kinds of informativeness, one involving physical factors and the other involving cognitive factors. The present experiment concerns informativeness in a cognitive sense, and accordingly the following definition is offered: An object in a picture is informative to the extent that the object has a low a priori probability of being in the picture given the rest of the picture and the observer's past history (Loftus, 1976; Loftus & Bell, 1975). Figure 1 illustrates this definition. The top panel of Figure 1 depicts a farm scene and contains a number of objects that one might reasonably expect to find in a farm—a tractor, a silo, a farmhouse, and so on. The bottom panel of Figure 1 shows the same scene with the one exception that an octopus has been substituted for the tractor. In terms of the above definition, the octopus in the lower panel would constitute an informative object because it is not redundant with or predictive of the rest of the scene—it has a low a priori probability of being in the scene given that the rest of the scene is a farm and given normal observers' past histories with farms. In contrast, the tractor in the top panel would constitute a noninformative object inasmuch as it has a high a priori probability of appearing on a farm.

Two principal questions were addressed in the present experiment: Given the above definition of informativeness, are informative objects in pictures fixated (a) earlier and (b) more often than are corresponding noninformative objects?

Method

Stimuli

Seventy-eight groups of 4 pictures per group were constructed, and each of the 312 pictures was made into a 35-mm slide. Each 4-picture group consisted of two different scenes factorially combined with two different critical objects. The combination of scenes with objects was such that within each group, 2 pictures contained informative objects and the other 2 pictures contained noninformative objects. For example, one group of pictures consisted of the 2 pictures shown in Figure 1 plus 2 other pictures, each depicting an underwater scene. One of the underwater-scene pictures contained the tractor (now informative), whereas the other underwater-scene picture contained the octopus (now noninformative). Note that such an arrangement controls for physical characteristics and means that the relative informativeness of a critical object is defined strictly in terms of the relationship of that object to the rest of the scene.[1]

Subjects

The 12 subjects were residents of the Palo Alto, California area and ranged in age from 18 years to 30 years. They were recruited through a newspaper advertisement and were paid $5 for participating in a session that lasted approximately 1 hr.

Design

Each subject was shown 78 pictures at 4 sec per picture. The 78 pictures consisted of 1 picture from each of the 78 groups with the restriction that half of the pictures contain an informative critical object and the other half contain a noninformative critical object. The 4 pictures within a group were rotated across subjects so that over a set of 4 subjects all 4 (pictures within a group) × 78 (groups) of the pictures were viewed exactly once. Four subjects thereby constituted one complete replication of the experiment, and a total of three replications (12 subjects) were run.

Procedure

Subjects were instructed that we were interested in "how people look at pictures that they will later have to recognize," and they were told to look at the picture as if they were going to have a later recognition test. Pictures were back projected and subtended a visual angle of approximately 30° × 20°. The subject was instructed to look at a fixation point approximately 30° below the viewing area prior to the onset of each picture and to look up at the picture when it appeared.

Eye movements were recorded during picture viewing using a digital, pupillary-reflection camera (Mackworth, 1976). The output of this device consists of an 18-frame/sec movie of the eye. Position of the eye is determined by aligning the center of the pupil relative to an 8 × 5 grid of lights that is reflected in the iris around the pupil.

Results

Thirty-one of the 78 picture groups were selected for analysis. The selection of these 31 picture groups was made on the following basis. Just prior to running the experiment we noticed that despite counterbalancing precautions, many of the informative objects could still be construed as being noticeable on the basis of physical characteristics. For instance, one picture contained a noninformative object consisting of one member of a line of ballet dancers. The corresponding informative object consisted of a football player sub-

[1] The judgment of what objects did or did not belong in a given scene was originally made by the artist entrusted with drawing the pictures. Subsequently, subjective ratings from 24 University of Washington undergraduates were obtained to verify these judgments. For every pair of pictures, the informative object was rated as less likely to appear in the scene than the corresponding noninformative object.

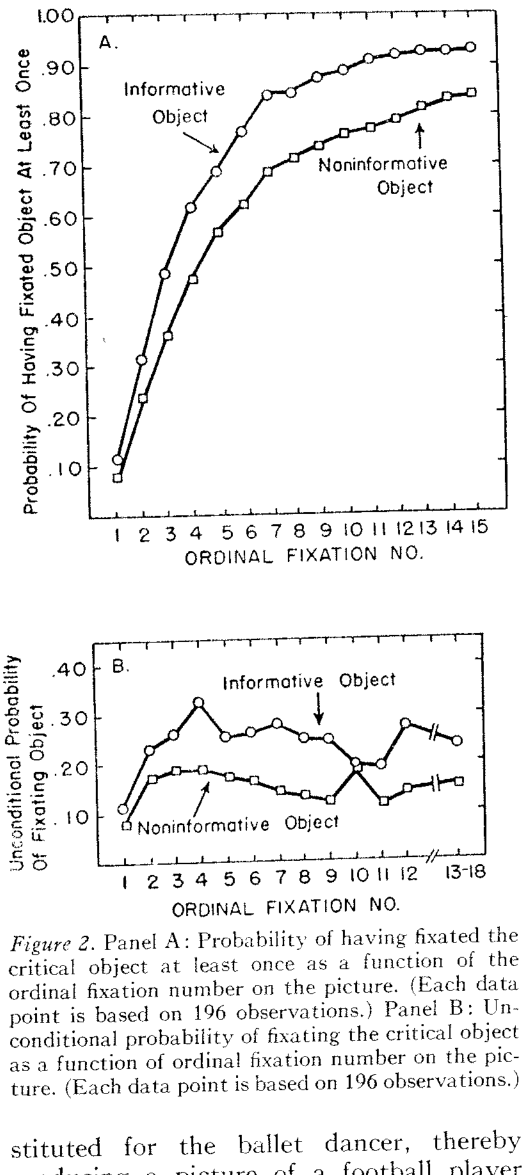

Figure 2. Panel A: Probability of having fixated the critical object at least once as a function of the ordinal fixation number on the picture. (Each data point is based on 196 observations.) Panel B: Unconditional probability of fixating the critical object as a function of ordinal fixation number on the picture. (Each data point is based on 196 observations.)

stituted for the ballet dancer, thereby producing a picture of a football player among a group of ballet dancers. Considering this pair of pictures, the football player might attract more attention than the ballet dancer simply because the football player was the one physically different object with respect to the group of ballet dancers with whom he was juxtaposed. We excluded from analysis all picture groups in which the relative dissimilarity of the critical object to the background could be used as a cue for distinguishing the informative and the noninformative object.

The raw eye movement data were analyzed by a judge who was naive with respect to any of the experimental hypotheses. This judge went through the movie film on a frame-by-frame basis and for each picture viewed by each subject listed the location and duration of each subject's fixation. As noted above, the eye movement recorder was designed such that the location of gaze should be localizable in one $4° \times 4°$ square of an imaginary 8×5 grid superimposed over the picture. However, the judge was actually able to localize the gaze within a particular quadrant of a particular square; hence the accuracy of the device was $\pm 1°$.

Informative Objects Are Found Faster

Figure 2A shows the cumulative probability of having fixated the critical object as a function of the ordinal fixation number of the picture (i.e., as a function of the first fixation on the picture, the second fixation, and so on). Separate curves are shown for pictures containing informative versus noninformative critical objects, and it is evident that informative objects were fixated earlier than were noninformative objects. The statistical reliability of this finding was assessed both across stimuli and across subjects. The analysis across stimuli was executed as follows. As illustrated in Figure 1, the pictures came in pairs. One member of each pair contained an informative object, whereas the other member of the pair, while depicting the same scene, contained a noninformative object. Since each of the 31 picture groups consisted of 2 such pairs, there was a total of 62 pairs. Considering median initial fixation number across the 3 subjects who had viewed each picture, the informative object was fixated earlier than the corresponding noninformative object for 31 of the pairs, whereas the reverse was true for 14 of the pairs. (The other 17 pairs were tied.) This advantage of informative over noninformative objects was significant by a sign test ($z = 2.54$, $p < .01$). With respect to the subject analysis, 8 of the 12 subjects fixated the informative object

Table 1
Mean Distance From the Critical Object (in Degrees of Visual Angle) of the Fixation Prior to the First Fixation on the Object

	Ordinal fixation number of first fixation						
	2		3		4 or higher		
Critical object	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*
Informative	6.4	1.04	7.72	3.48	8.08	2.20	7.40
Noninformative	7.20	2.80	7.20	2.36	7.32	1.64	7.25

earlier (again considering median initial fixation number), whereas 3 subjects fixated the noninformative object earlier. The advantage of informative objects was again significant, $t(11) = 1.80$, $p < .05$.

Of some interest is the probability of initially finding the critical object on the second fixation. These probabilities were .147 for noninformative versus .214 for informative critical objects, a difference that was significant by a sign test for the difference between two proportions ($z = 1.72$).

Spatial analysis. Do the results of Figure 2A imply that informative objects are identified in peripheral vision? This is one possibility, but an alternative interpretation of the results would run as follows: The gaze is drawn about the picture by noncognitive factors, and when it happens to fall relatively close (e.g., 2°–3°) from the critical object, the object is identified not in peripheral but in near-foveal vision. Following such identification, an informative object is fixated with a higher probability than a noninformative object.

To test this conjecture, the following analysis was performed. First, for each subject, pictures were grouped according to whether the first fixation on the critical object was the second, third, or fourth or greater ordinal fixation on the picture.[2] For each of the groupings, the location of the fixation *prior* to the first fixation on the object was identified, and the distance (in degrees of visual angle) from that location to the location of the object was tabulated. Table 1 shows the results of this analysis. The average distance was on the order of 6.5°–8° of visual angle. This finding indicates that, just prior to fixating the critical object, the object was, in general, well outside the foveal region.

Informative Objects Are Looked at More Often

Figure 2B shows the unconditional probability of fixating the critical object as a function of the ordinal fixation number. On any given fixation, the probability of fixating an informative object is greater than the probability of fixating a noninformative object. Again, considering the 62 pairs of pictures, the mean number of fixations on the informative object was greater than the mean number of fixations on the noninformative object for 42 of the pairs, whereas the reverse was true for 10 of the pairs. Again, this difference was significant by a sign test ($z = 4.43$, $p < .01$). All 12 subjects averaged more fixations on informative than on noninformative objects.

Informative Objects Are Looked at With Longer Durations

A final question of interest is, Do average fixation durations differ for informative versus noninformative objects? Although this question is a straightforward one, the proper analysis to answer it is not straightforward because there are a number of

[2] The first fixation on the picture was almost invariably on the center of the picture. Apparently, very little information about any characteristics of the picture could be garnered from the initial fixation point, which, as noted, was about 30° below the picture.

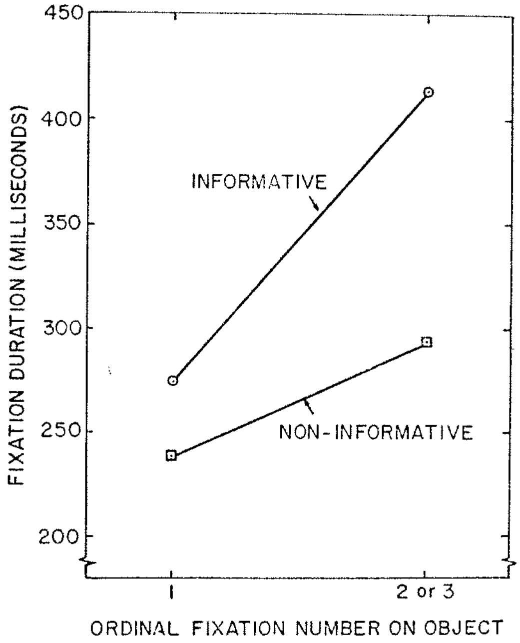

Figure 3. Average fixation duration as a function of ordinal fixation number on the critical object. (The numbers of observations for the four data points are 55, 67, 43, and 30, respectively, starting from the top right of the figure and moving clockwise.)

factors (e.g., number of fixations, ordinal fixation number, etc.) that covary with fixation duration. Trying to control for all these factors results in an inordinate amount of missing data. Ultimately, the following analysis was performed. For each subject, pictures were selected that had received exactly one, two, or three fixations on the critical object. Four data points were then computed across these three types of pictures: duration of the first fixation on the object and average duration of the second and third fixations on the object for both informative and noninformative objects. These data points were subsequently averaged across subjects to produce the results shown in Figure 3. As is evident, durations tended to be longer on informative than on noninformative objects, $F(1, 11) = 8.23$, and duration increased from the first to subsequent fixations, $F(1, 11) = 4.97$. That this increase stemmed primarily from informative objects is demonstrated by a significant interaction, $F(1, 11) = 5.76$.

Discussion

The present results demonstrate that subjects tend to look earlier, more often, and for longer fixation durations at informative as opposed to noninformative objects in pictures. These results are suggestive, both with respect to the nature of peripheral processing during picture viewing and with respect to the relationship between information acquired from a picture and subsequent recognition memory for that picture.

Peripheral Processing

It has generally been assumed that some kind of peripheral processing occurs during a fixation whose output is used to determine the location of subsequent fixations (e.g., Gould, 1976; Note 1). Several properties of such processing have been revealed in past research. For example, fixations can be directed toward the periphery based on gross physical characteristics such as rapid contour change or the onset of a light (Hallett & Lightstone, 1976; Mackworth & Morandi, 1967; Yarbus, 1967). Additionally, somewhat higher level stimulus characteristics such as color and shape may be used as a basis for visual search (Gould & Dill, 1969; Williams, 1967). However, the present finding that informative objects are fixated faster than noninformative objects (Figure 2A) demonstrates that there exists rapid peripheral processing based on cognitive information and that the determination of fixation location is based, at least in part, on the results of this processing. Specifically, at least three events must be occurring in the early stages of picture viewing. First, as suggested by past researchers (Biederman, 1972; Potter, 1975), a quick determination of the gist of the scene must be made. Second, objects in the periphery must be at least partially pattern recognized. Third, a computation must be made of the conditional probabilities that these objects belong in the scene given the gist that has just been revealed. It is apparently the case that fixations are then directed to objects whose conditional probability of being in the scene is assessed to be low.

More Attention to Informative Objects

Figure 2B indicates that more fixations are allotted to informative as opposed to noninformative objects; and similarly, Figure 3 indicates that fixations on informative objects tend to be longer. In short, considerably more processing is being allotted to informative objects.

One reason for allocating more attention to informative objects may relate to a memorization strategy. Subjects in the present experiment were looking at pictures with the intent of subsequently being able to recognize them. Recognition of a picture involves the capability of discriminating that picture from other, similar pictures. Therefore, the most efficient encoding strategy would be to encode the presence of those features that are least likely to appear in potential distractors. Such features are precisely those that are most informative according to the definition offered above.

Loftus (1972) has demonstrated that recognition memory does not depend on fixation duration during initial viewing. Yet fixation durations in the present experiment were longer on informative than on noninformative objects. Why? Several explanations may be rejected. First, it cannot be the case that an informative object "in isolation" demands longer fixations than a noninformative object, since specific objects appear in both an informative and noninformative condition. Second, it is unlikely that the additional time is spent planning subsequent fixations, since informative and noninformative objects appear in the same place in the same scene.

A more reasonable answer may lie in the process of "linking" objects to scenes. Consider the following speculation. A good deal of past research has indicated that when presented with a picture, subjects utilize a pre-formed schema to help organize the information in the picture (Biederman, 1972; Mandler & Parker, 1976). It has been argued above that subjects acquire the gist of the scene very quickly—probably during the first fixation on the scene. Acquiring the gist may be tantamount to activating a schema, and subsequent fixations may have the purpose of verifying the presence of various objects that belong in the schema. An informative object is, by our definition, an object that does *not* belong in the schema. The extra time spent fixating such an object would, by this reasoning, represent the time needed to add it to the schema.

Reference Note

1. Gould, J. D. *Eye movements during visual search* (Rep. No. RC2680). Yorktown Heights, New York: IBM Thomas J. Watson Research Center, 1969.

References

Antes, J. R. The time course of picture viewing. *Journal of Experimental Psychology*, 1974, *103*, 62–70.

Berlyne, D. E. The influence of complexity and novelty in visual figures on orienting responses. *Journal of Experimental Psychology*, 1958, *55*, 289–296.

Biederman, I. Perceiving real-world scenes. *Science*, 1972, *177*, 77–80.

Buswell, G. T. *How people look at pictures: A study of the psychology of perception in art.* Chicago: University of Chicago Press, 1935.

Gould, J. D. Looking at pictures. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes.* Hillsdale, N. J.: Erlbaum, 1976.

Gould, J. D., & Dill, A. B. Eye movement parameters and pattern discrimination. *Perception and Psychophysiology*, 1969, *6*, 311–320.

Hallett, P. E., & Lightstone, A. D. Saccadic eye movements to flashed targets. *Vision Research*, 1976, *16*, 107.

Latour, P. Vision thresholds during eye movements. *Vision Research*, 1962, *2*, 261–262.

Loftus, G. R. Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 1972, *3*, 525–551.

Loftus, G. R. A framework for a theory of picture memory. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes.* Hillsdale, N. J.: Erlbaum, 1976.

Loftus, G. R., & Bell, S. M. Two types of information in picture memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1975, *104*, 103–113.

Mackworth, N. H. Ways of recording line of sight. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes.* Hillsdale, N. J.: Erlbaum, 1976.

Mackworth, N. H., & Bruner, J. S. How adults and children search and recognize pictures. *Human Development*, 1970, *13*, 149–177.

Mackworth, N. H., & Morandi, A. J. The gaze selects informative details within pictures. *Perception and Psychophysiology*, 1967, *2*, 547–552.

Mandler, J. M., & Parker, R. E. Memory for descriptive and spatial information in complex pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, *2*, 38–48.

Pollack, I., & Spence, D. Subjective pictorial information and visual search. *Perception & Psychophysics*, 1968, *3*, 41–44.

Potter, M. C. Meaning in visual search. *Science*, 1975, *187*, 965–966.

Volkman, F. C. Saccadic suppression: A brief review. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes*. Hillsdale, N. J.: Erlbaum, 1976.

Williams, L. The effects of target specification on objects fixated during visual search. *Acta Psychologia*, 1967, *27*, 355–360.

Yarbus, A. L. [*Eye movements and vision*] (B. Haigh, trans.). New York: Plenum Press, 1967.

Received September 27, 1977 ■