

Recognizing and localizing features in brief picture presentations

JAMES R. ANTES

University of North Dakota, Grand Forks, North Dakota 58202

In two experiments, subjects saw a series of brief picture flashes, each followed by a 10-sec presentation of a small picture section. In Experiment 1, subjects were required to decide if the section had been part of the 100-msec picture presentation. Experiment 2 had 100-msec and 500-msec presentations, and the task was to indicate the location of the section in the picture. Both recognition and localization accuracy depended upon the rated informativeness of the probed section and the section's location with respect to point of focus. Performance on the recognition task was superior to that on the localization task, suggesting different rates of acquisition of identity and location information. The results were interpreted in terms of a model of picture perception based upon two kinds of processing: identification of individual objects and a holistic characterization of the scene. The results also have implications regarding the use of information in the guidance of initial eye fixations.

Biederman and colleagues have conducted a series of experiments investigating the role of picture context in the perception of "real-world" scenes (Biederman, 1972; Biederman, Glass, & Stacy, 1973; Biederman, Rabinowitz, Glass, & Stacy, 1974). Picture context was manipulated by cutting photographs into six equal-sized sections and then jumbling the sections. Subjects were required to view jumbled and intact scenes under a number of different task conditions. The results showed that jumbling interfered with identification of and increased search time for objects in the pictures. Biederman et al. (1973) concluded that picture context is an important factor affecting identification of picture objects and that this effect may be mediated by a schema constructed from the information extracted from the context of the scene.

Another approach to the study of the perception of complex scenes was introduced by Mackworth and Morandi (1967). They divided pictures into small sections and required subjects to rate the "informativeness" of each piece. When other subjects viewed the intact pictures, their eye fixations were concentrated on the sections which had been rated highly informative. Pollack and Spence (1968) applied the informativeness rating technique to a visual-search task and found search times to be shortest when the searched-for picture segment possessed a high-informativeness rating.

This research was supported in part by a grant from the University Committee on Research, University of North Dakota. Portions of this research were reported at the 1976 Midwestern Psychological Association meeting. The author wishes to acknowledge the helpful assistance of Daniel Anderson, Wendy Boriskin, Richard Metzger, Carl Schwensohn and Rebecca Wiese. The valuable comments of Dr. Irving Biederman and an anonymous reviewer are also greatly appreciated. Requests for reprints should be addressed to James R. Antes, Department of Psychology, University of North Dakota, Grand Forks, North Dakota 58202.

The first experiment combined the two approaches outlined above to examine the effect of rated informativeness on object recognition in a picture. Subjects were required to determine whether or not a small picture segment was part of a previously flashed picture. The informativeness of the probed picture sections was varied. Since identification of complex stimuli is known to vary with retinal eccentricity (e.g., Edwards & Goolkasian, 1974), the location of the probed section was also systematically varied.

EXPERIMENT 1

Method

Stimuli. The stimuli were selected from a pool of 111 4 in. x 6 in. colored photographs (mainly landscapes and still lifes) taken from popular nature and photography magazines. Each picture was photographed behind a clear plastic overlay which was divided by black lines into eight equal-sized sections, two rows of 2 in. x 1.5 in. rectangles. The pictures were prepared as Ektachrome slides and projected onto a screen one at a time to a group of 22 (10 female) undergraduate student volunteers. The subjects were instructed to rank order the eight sections of each picture on the dimension of "amount of information conveyed."

The rankings of each section of all 111 pictures were summed over the 22 subjects. The pictures were then categorized according to location of section with lowest total rank and highest total rank. Four pictures were then randomly selected from those whose lowest ranking section was the upper left-hand section; similarly, four pictures were randomly selected from those whose upper left-hand section was the highest ranking section of the picture. The procedure was repeated for all eight picture sections, with the restriction that no picture could be represented in a low-ranking group for one picture section and a high-ranking group for another section. Thus, the stimulus set consisted of 64 pictures, with each of the eight section locations being represented by 4 pictures in which that section was "most informative" and 4 pictures in which that section was "least informative."

The original photographs of the 64 pictures were rephotographed, without the plastic overlay, on a black background.

Then the target section (most or least informative) was photographed alone to appear as the same size at the same location as the section in the intact picture.

The design called for an equal number of match and mismatch trials for both informativeness levels of each section. Of the four pictures representing each section at a given level of informativeness, two were selected which were most similar with respect to content and color. The target sections of the pictures were interchanged, and the picture-section pairs were presented on mismatch trials. The remaining two picture-section pairs were presented on match trials.

The two picture sections at the extreme left and the two at the extreme right were designated as peripheral sections, and the four inner sections were designated as central sections. The 64-member stimulus set, then, consisted of 8 match and 8 mismatch picture-section pairs within each factorial combination of location (peripheral and central) and informativeness (high and low).

Subjects. The subjects were 20 (8 female) University of North Dakota undergraduates who participated for course credit from the introductory psychology classes. All reported normal vision without glasses or wore contact lenses.

Apparatus. The stimuli were projected onto a rear-projection screen by means of two Kodak Carousel projectors. Timing of exposures and interstimulus intervals was controlled by Lafayette tachistoscope shutters and external BRS/LVE timers.

Procedure. When the subject reported to the laboratory, he was seated facing the rear-projection screen, his head resting on a chin rest, 60 cm from the screen. On a ready signal from the experimenter, the subject fixated a small dot in the center of the screen. The picture then appeared for 100 msec, followed 400 msec after offset by a 10-sec presentation of the picture section. Subjects indicated on a six-point scale their certainty that the section was contained in the previously flashed picture. The end points of the scale were labeled "1: very certain sections are different" and "6: very certain sections are the same." Subjects were assured that matching sections would always be presented in the same location as they appeared in the intact picture. Subjects were informed that half of the ensuing trials would contain matching picture-section pairs.

Eight practice trials were presented, followed by the 64 experimental trials. All subjects received the trials in the same order, which was randomized with the restrictions that no section could be probed on consecutive trials, that there could be no more than three consecutive trials of matching or mismatching picture-section pairs, and that there could be no longer than three-trial strings where the probed sections were all top, bottom, left, right, center, or peripheral sections.

The intact pictures subtended 20 deg of visual angle horizontally and 13 deg vertically. Thus, each section was 5 deg x 6.5 deg in size. When projected onto the screen, all center sections had one corner in common with the fixation dot and the inner corner of each peripheral section was 5 deg from fixation. The experiment was conducted under photopic viewing conditions, with normal room illumination.

Results

The rating scale was partitioned between responses "3" and "4" in order to obtain a measure of number of errors. The number of errors for each subject was tabulated and summed over subjects and conditions to yield an overall error rate of 13.4%. Table 1 presents the mean number of errors for each condition. The errors were analyzed consistent with the suggestion by Clark (1973) that investigators using language materials (or pictures) as stimuli provide statistical evidence that their findings generalize beyond the specific sample of stimuli chosen for their experiment. Consequently, a 2 (informativeness) by 2 (location) by 20 (subjects) by 16 (pictures within informativeness by location) analysis of variance was computed with subjects and pictures treated as random effects. Quasi F ratios and appropriate degrees of freedom were calculated as described by Clark. Neither the main effects of informativeness [$F(1,72) = 1.78, p > .05, MSe = .75$] and location [$F(1,74) = 1.90, p > .05, MSe = .77$], nor their interaction [$F(2,70) = .63, p > .05, MSe = .74$] were significant sources of variation. There was a tendency, however, toward more errors on peripheral sections and on low-informative sections (see Table 1).

The signal detection theory measures of d' and β were computed for each subject for each of the four trial conditions (Informativeness by Location) and the means are presented in Table 1. Since the computation of these measures required summing over pictures, the data were analyzed by conventional Informativeness by Location by Subjects analyses of variance. Thus, contrary to analyses employing pictures as a random source of variation, any significant results in these analyses are generalizable only to other subjects who may view these pictures and not to other subjects and other pictures. For d' , the main effects of informativeness [$F(1,19) = 14.15, p < .01, MSe = .54$] and location [$F(1,19) = 6.87, p < .05, MSe = .89$] were significant, but the Informativeness by Location interaction did not reach significance [$F(1,19) = .95, p > .05, MSe = .79$]. Subjects showed greater discriminability for central sections and for high-informative sections.

The results of the β analysis revealed only one significant source of variation, a highly significant informativeness main effect [$F(1,19) = 50.58, p < .01$,

Table 1
Mean, Errors, d' and β for Each Condition of Experiment 1

Informativeness	Location						Overall*		
	Central			Peripheral			Errors	d'	β
	Errors	d'	β	Errors	d'	β			
Low	1.85	3.00	1.63	3.45	2.25	.36	5.30	2.62	.99
High	1.40	3.42	7.44	1.90	3.06	5.02	3.30	3.24	6.23
Overall*	3.25	3.21	4.53	5.34	2.65	2.69	8.60	2.93	3.61

*Overall error means were computed by summing errors across conditions within subjects and dividing by the number of subjects. Overall d' and β means were computed by averaging across conditions within subjects and dividing by the number of subjects.

MSe = 10.82]. Examination of the means showed that, relative to low-informative trials, there was a large "nay-saying" bias on trials when a high-informative section was probed.

To further examine the response criterion shift, the errors tabulated for the first analysis were divided into misses and false alarms and subjected to identical Informativeness by Location by Subjects by Pictures analyses of variance. The results showed surprisingly different responses for the two types of errors. Subjects made false alarm errors, which were 55% of total errors, according to the same pattern as total errors, but the differences were much larger. Both informativeness [$F(1,34) = 6.84$, $p < .05$, $MSe = .98$] and location, [$F(1,32) = 4.03$, $p = .05$, $MSe = .95$] were significant, but the Informativeness by Location interaction was not [$F(1,31) = 1.31$, $p > .05$, $MSe = .92$].

However, when misses were analyzed, only one significant effect obtained: informativeness [$F(1,42) = 3.98$, $p = .05$, $MSe = .27$]. The main effect of location and its interaction with informativeness were not significant ($F_s < 1$). Furthermore, the direction of the significant informativeness main effect was opposite to that of previous error analyses. Subjects made significantly *more* misses when high-informative picture sections were probed than when low-informative sections were probed.

Discussion

The ability to recognize small sections from a briefly flashed picture depends not only on the surrounding context, as Biederman (1972) has shown, but also on the subjective informativeness of elements within the section, as the present experiment has shown. This result is not surprising and is entirely consistent with other studies (e.g., Pollack & Spence, 1968) relating rated informativeness to some performance measure.

The pattern of false alarms and misses, however, has important implications regarding the kinds of processing involved on low- vs. high-informative areas. When low-informative picture sections were probed, there were more false alarm errors than misses. However, on high-informative probes, misses predominated. Both results could be related to the dual mode of picture processing, suggested by Biederman et al. (1974) and by Loftus and Bell (1975). Briefly, picture processing according to the dual mode analysis involves (1) identification of individual objects and (2) a holistic characterization of the picture. In the present experiment, when the probe alternative was presented, subjects could have examined it for a specific object identified when the picture was flashed. Or, in the absence of readily identifiable objects, subjects may have ascertained whether the section could have been contained in the flashed picture based upon an overall characterization of the scene. Antes and Stone (1975), factor analyzing subjects' ratings of pictorial informativeness, determined that highly informative sections of a picture are more likely than uninformative

sections to contain a readily identifiable object. Thus, in the present study, when informative sections were probed, subjects could decide whether the information in that section was encoded from the picture flash. Since it is unlikely that specific details from an incorrect probe alternative would be mistaken for specific information in the original picture, misses would be more likely to occur than false alarms. When low-informative sections were probed, which were less likely to contain readily identifiable objects, subjects could have made a decision based upon a holistic characterization of the picture (e.g., "could the section have been part of the picture?") rather than "was that object part of the picture?") Since incorrect alternatives were chosen on the basis of general visual similarity to the flashed picture, subjects would then be more likely to make a false alarm error on low-informative sections. Thus, the processing of high-informative and low-informative elements of a picture may be based, respectively, upon specific object information and a general scene characterization.

If subjects were basing many of their judgments on a general scene characterization, it is not clear whether they were using physical information or semantic information. That is, to what extent were they basing their same-different decisions on color and contour, and, to what extent on the likelihood that the probe fit in with the extracted "meaning" of the scene? To evaluate this, a group of 27 (26 female) students from the same subject pool as described above were shown simultaneously the distractor probe and the original picture section from each of the 32 mismatch trials, and rated their physical and semantic similarity. Subjects were instructed to rate the sections for both similarity of "physical features (color, contour, texture, etc.," and "subject matter." The 10-point scale was anchored at 0 = no similarity and 9 = complete similarity (identity). Mean physical and semantic similarity ratings were obtained for each picture and the values were correlated with the number of errors the subjects of Experiment 1 committed on these pictures. The results are given in Table 2. Although no statistical comparison of the correlations was made, it can clearly be seen that, within each level of location and informativeness, the magnitudes of the correlations of physical similarity and semantic similarity with errors are virtually identical. Thus, the subjects were apparently not basing their decisions more on one type of information (physical vs. semantic) than the other.

The fact that both types of similarity correlated substantially higher with error rate for low-informative trials than for high-informative trials supports the contention made above that the processing of high-informative areas is based upon specific object information, and processing of low-informative areas depends upon a holistic picture characterization. If subjects were examining the probes for an object identified during the picture flash, then the gross physical or

Table 2
Correlations of Rated Physical and Semantic Similarity of Distractor Probes and Picture Sections with False Alarm Errors in Experiment 1

Informativeness	Location					
	Central		Peripheral		Overall	
	Physical Similarity	Semantic Similarity	Physical Similarity	Semantic Similarity	Physical Similarity	Semantic Similarity
High	.26	.12	-.09	.02	-.04	.18
Low	.55	.50	.84**	.87**	.64**	.71**
Overall	.44	.56*	.85*	.59*	.54**	.59**

* $p < .05$

** $p < .01$

semantic similarity of the section and the probe should not be highly related to performance. Only when the probe and section contain nearly identical objects should there be confusion. However, if subjects were making judgments based on the compatibility of the probe with a holistic characterization of the flashed picture, then the physical or semantic similarity of the probe and the section should be highly related to performance. As the correlations in Table 2 show, these are precisely the results for high-informative and low-informative trials, respectively.

Whatever the specific processes involved in the task, it is clear that subjects could process a large amount of information in a very brief time. Overall accuracy was 86.6% with the 100-msec exposure. It may be assumed, however, that the effective processing interval was somewhat longer due to iconic persistence, since a visual mask was not employed. In a similar task with a 100-msec exposure followed by a mask and employing four response alternatives, Biederman et al. (1974) found about 48% accuracy. This figure probably is closer to reflecting the relative amount of information available to a subject in a single-fixation duration exposure. It is apparent that the development of a "schematic map" (Hochberg, 1968) or internal representation of the scene does not require the integration of several successive glances as Hochberg suggested. It is clear that the subject has at his disposal prior to the first eye movement considerable information (central and peripheral) to use in guiding fixation choices. This is supported by the observation of Antes (1974) that many first eye movements are directed to highly informative features several degrees in the periphery.

Experiment 1 of the Biederman et al. (1974) article provides a rough estimate of the speed of information extraction in the development of a schematic map. Subjects viewed brief (from 20- to 4,000-msec) picture presentations, after which they were required to pick from a pair of verbal labels the one that accurately described the scene. Accurate performance of the task when the alternative labels were similar would presumably require that subjects would have developed a rather complete representation of the scene. With a 50-msec exposure, subjects could choose the correct descriptor

about 65% of the time. If this duration is taken as a rough approximation of the time required to form an accurate schema of a picture, then it is comparable to Sperling's (1963) estimate of 10 msec/letter as the speed of forming an internal representation of letter arrays.

It follows from the above discussion that subjects should have available in brief picture exposures not only information to recognize pictured objects, but also information to accurately locate these objects. That is, since (1) Experiment 1 showed that subjects could accurately recognize sections from briefly presented pictures at all locations, (2) saccadic eye movements are ballistic, and (3) the eyes are often guided to important picture features on the first fixation, accurate location information should also be processed in these exposures. The next experiment was designed to test this prediction. Subjects saw a 100-msec picture exposure followed by a single picture section which they had to identify as originating from one of eight locations in the picture flash.

EXPERIMENT 2

Method

Stimuli. The stimulus pictures were the same as those used in Experiment 1. The same probe sections were selected from each picture but were paired with the pictures such that all were correct matches.

Subjects. The subjects were 20 graduate or undergraduate students (13 female) from the University of North Dakota. All reported normal vision without glasses or wore contact lenses.

Procedure. Subjects were seated as before, facing the rear-projection screen. On a ready signal from the experimenter, the subject fixated the focus point and then the picture appeared for 100 msec. Four hundred milliseconds after offset, the picture section appeared for 10 sec, centered 11 deg directly below the lower border of the flashed picture. Subjects responded by marking on a facsimile grid in which of eight locations the section had appeared. Trial order, stimulus size, and room illumination were the same as in Experiment 1.

Results and Discussion

Mean correct responses for each experimental condition are presented in Table 3. Average percent correct over all conditions was 31.4%. An Informativeness by

Table 3
Mean Percent Correct and Percent Central and Adjacent Error Choices in Experiment 2

Condition	Mean Percent Correct		Percent Error Choices					
			Central Errors			Adjacent Errors		
	100 msec	500 msec	100 msec	500 msec	Chance*	100 msec	500 msec	Chance**
High Central	45	58	72	77	43	90	93	71
High Peripheral	23	34	80	75	57	72	78	43
Low Central	35	44	69	71	43	88	90	71
Low Peripheral	23	18	82	78	57	56	66	43

*There were three possible erroneous central choices when a central probe was presented and four when a peripheral probe was presented.

**There were five possible erroneous adjacent choices when a central probe was presented and three when a peripheral probe was presented.

Location by Subjects by Pictures analysis of variance showed performance to be better on central vs. peripheral sections [$F'(1,74) = 8.69, p < .01, MSe = 1.11$]. The informativeness effect was not significant [$F'(1,78) = 1.10, p > .05, MSe = .89$], nor was its interaction with location [$F'(2,79) = .90, p > .05, MSe = .98$].

Examining the pattern of errors (see Table 3) suggests that, first, when subjects erred they tended to choose an adjacent section. Thus, if they could not locate the section precisely, they were able to place it in the correct general area. A more surprising finding was the tendency of subjects to locate the probed section in the center of the picture regardless of its actual location. This "centration" effect was evident in all experimental conditions.

As expected, the location of the probed sections affected subjects' ability to localize them. However, it was also anticipated that performance would be better on the high-informative probes than on the low-informative probes. Furthermore, overall performance was much lower than expected. If initial eye movements are directed to specific elements in a picture, then accurate information regarding the location of the elements should have been extracted from the 100-msec flash. It is, of course, possible that while location information obtained from a 100-msec exposure is relatively incomplete, localization accuracy increases dramatically as exposure duration increases and one or two eye movements are possible. To explore this possibility, the above procedure was repeated with a 500-msec exposure duration.

The subjects were 20 undergraduate students (11 female) from the University of North Dakota. All reported normal vision without glasses or wore contact lenses. The stimuli and procedure were identical to those reported above except that the exposure duration was 500 msec. The probe stimulus occurrence was again timed 400 msec from the offset of the picture.

Overall performance was 38.4%, a slight improvement over that for the 100-msec exposure. As Table 3 indicates, all conditions evidenced improvement except the low-informative peripheral condition in which subjects

actually showed a slight decrease in performance. An Informativeness by Location by Subjects by Pictures analysis of variance was performed on the number correct. Localization was more accurate for high-informative sections [$F'(1,77) = 7.72, p < .01, MSe = .94$] and for central sections [$F'(1,79) = 17.17, p < .01, MSe = 1.08$], and these variables did not interact [$F'(15,78) = .25, p > .05, MSe = .98$]. The data from the two durations were combined to provide an overall analysis, with duration as a between-subjects variable. This resulted in a significant duration effect [$F'(1,86) = 4.52, p < .05, MSe = .72$], substantiating the improved performance with increased duration. Subjects also performed better overall on the high-informative sections [$F'(1,75) = 4.81, p < .05, MSe = .35$] and the central sections [$F'(1,88) = 17.80, p < .01, MSe = 1.53$]. None of the interactions were significant except the Duration by Informativeness interaction, which approached significance [$F'(1,97) = 3.64, p < .06, MSe = .47$]. A Newman-Keuls analysis of this interaction showed performance on high-informative probes from the 500-msec picture presentations to be significantly better ($p < .05$) than that on the high-informative 100-msec presentations and the low-informative probes from both exposure durations. Performance under the latter three conditions did not differ. Thus, subjects' accuracy in localizing picture sections improved with increased exposure time on high-informative probe sections, but not on low-informative probe sections.

Examination of the errors (see Table 3) also suggests an improvement in performance. Erroneous section choices were more likely to be adjacent to the correct section with the 500-msec exposure than with the 100-msec flash for all experimental conditions. Similarly, error choices on probes from the central part of the picture were more likely to be placed in other central sections, and error choices on peripheral probes were more likely to be other peripheral sections. Interestingly, error measures indicated improvement even for the low-informative peripheral probes on which subjects showed a decline in number correct.

Thus, accuracy in localizing picture sections did increase with increased exposure time, but not uni-

formly across experimental conditions, and not as dramatically as was expected. The fact that the greatest improvement occurred on high-informative areas is consistent with the finding that high-informative areas dominate fixation choices early in viewing (Antes, 1974). However, localization accuracy was still rather low on these sections and the fact that location was grossly defined as eighths of a picture suggests that the accurate location of smaller features would be far worse.

The overwhelming tendency with both exposure durations was to mislocate sections toward the center of the picture. This finding was observationally supported 2 years ago in this laboratory in a partial replication of a study by Haber (1970). Subjects were shown 100-msec picture flashes and were asked to sketch and describe what they saw. They could accurately identify many features at various locations, but in most instances they drew the features toward the center of the picture outline which was provided for them. One interpretation of the results is that the subject's representation of the picture is the source of the bias. That is, the schema constructed from information in the brief presentation may be quite imprecise with respect to location of objects. It is also possible that the results only reflect a response tendency, that subjects tend to guess a central response when they are unsure. However, the possibility that the perceptual or memory apparatus contains a centering bias is worth further investigation. Additional research, perhaps with longer exposure times, is certainly warranted.

Taking the pattern of results of the two experiments together, it is clear that location and subjective informativeness are important variables in identifying and localizing pictorial features. The superior performance of central over peripheral presentations has been documented many times elsewhere (e.g., Edwards & Goolkasian, 1974). The effect of informativeness adds to the previously demonstrated relationship of that variable to eye-fixation density (Antes, 1974; Mackworth & Morandi, 1967) and visual-search speed (Pollack & Spence, 1968). However, rated informativeness of arbitrarily defined picture sessions is a rather crude characterization of the complex information in a picture, and better kinds of analyses are clearly needed. One potential method is to employ multidimensional scaling techniques to determine which judgmental factors subjects use in their informativeness ratings (e.g., Antes & Stone, 1975). A new approach to this problem is called "psychopictorics" (Lipkin & Rosenfeld, 1970) which, employing computer technology and psychophysical techniques, "is concerned with investigating where the information in complex pictures resides" (p. vii).

Comparisons of the relative performance on the recognition and the localization tasks are difficult

because of inherent differences in the tasks. The data are suggestive, however, that the information acquired about the identity of objects in brief flashes is more complete than that relating to their location. If we assume that information obtained from successive fixations is integrated into a schematic map (Hochberg, 1968), then this representation, initially at least, is quite accurate with regard to object identity, but approximate with regard to object location. It may be that accurate localization depends upon first having accurate identity information available, or it could be the case that different processes underly identification and localization and the latter process is slower. Trevarthen (1968) supports the second alternative in part with his suggestion that vision in primates involves two parallel and anatomically distinct mechanisms, one concerned with spatial vision and the other with object identity. Another possible explanation of the differences in recognition and localization accuracy may be derived from Gibson's (1950) "visual field" and "visual world" distinction. The visual field of a subject contains information in a single glance and is detailed at the center but progressively less clear toward the periphery. The visual world is made up of solid objects built up from several eye fixations and has no center/periphery differences in clarity. Identification of objects could be made in the visual field, but localization may require the buildup of a representation of objects located relative to other objects in the visual world. Since the present tasks were essentially visual-field tasks, according to this analysis, performance on the recognition task would be expected to be superior.

Given that the early representation of a picture is quite inaccurate, at least with respect to the location of objects, how can initial eye movements be accurately guided? Maybe they cannot, as Buswell (1935) implied several years ago. From his extensive eye movement records, he identified two different types of scanning patterns. According to this analysis, the first few eye fixations are usually idiosyncratic, not guided by information in the picture, but directed to orient the observer to the scene; the second scanning pattern, examination of informative detail, follows. However, this must be reconciled with Antes' (1974) finding that, on the average, the eyes are directed to the areas of highest informativeness by the second fixation, after viewing the picture for about .5 sec. The point of conflict may rest on task differences. In Antes' study, subjects casually viewed a series of 10 pictures for 20 sec each in order to determine a preference. Since the task in the present study was quite different and since it is known that the subjects' task can affect scanning patterns, at least for such tasks as discrimination learning (Schroeder, 1970), visual search (Ford, White, & Lichtenstein, 1959), size estimation (Gardner & Long, 1962), the game of chess (Tikhomirov & Poznyanskaya, 1966-67), and driving

an automobile (Gordon, 1966), it is possible that the results of Antes do not apply here. What is needed is research employing the identification and localization tasks and including the recording of eye fixations. Such research could more directly address the question of information used in the guidance of initial eye fixations.

REFERENCES

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

ANTES, J. R., & STONE, L. A. Multidimensional scaling of pictorial informativeness. *Perceptual and Motor Skills*, 1975, **40**, 887-893.

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

BIEDERMAN, I., GLASS, A. L., & STACY, E. W. Searching for objects in real-world scenes. *Journal of Experimental Psychology*, 1973, **97**, 22-27.

BIEDERMAN, I., RABINOWITZ, J. C., GLASS, A. L., & STACY, E. W. On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, 1974, **103**, 597-600.

BUSWELL, G. T. *How people look at pictures*. Chicago: University of Chicago Press, 1935.

CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 335-359.

EDWARDS, D. C., & GOOLKASIAN, P. A. Peripheral vision location and kinds of complex processing. *Journal of Experimental Psychology*, 1974, **102**, 244-249.

FORD, A., WHITE, C. T., & LICHTENSTEIN, M. Analysis of eye movements during free search. *Journal of the Optical Society of America*, 1959, **49**, 287-292.

GARDNER, R. W., & LONG, R. I. Control, defense and centration effect: A study of scanning behaviour. *British Journal of Psychology*, 1962, **53**, 129-140.

GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.

GORDON, D. A. Experimental isolation of the driver's visual input. *Human Factors*, 1966, **8**, 129-137.

HABER, R. N. How we remember what we see. *Scientific American*, 1970, **222**, 104-115.

HOCHBERG, J. In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception*. New York: Holt, Rinehart, & Winston, 1968.

LIPKIN, B. S., & ROSENFELD, A. (Eds.) *Picture processing and psychopictorics*. New York: Academic Press, 1970.

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., & MORANDI, A. J. The gaze selects informative details within pictures. *Perception & Psychophysics*, 1967, **2**, 547-551.

POLLACK, I., & SPENCE, D. Subjective pictorial information and visual search. *Perception & Psychophysics*, 1968, **3**, 41-44.

SCHROEDER, S. Selective eye movements to simultaneously presented stimuli during discrimination. *Perception & Psychophysics*, 1970, **7**, 121-124.

SPERLING, G. A model for visual memory tasks. *Human Factors*, 1963, **5**, 19-31.

TIKHOMIROV, O. K., & POZNYANSKAYA, E. D. An investigation of visual search as a means of analyzing heuristics. *Soviet Psychology*. Winter, 1966-67, **5**, 3-15.

TREVARTHEN, C. B. Two mechanisms of vision in primates. *Psychologische Forschung*, 1968, **31**, 299-337.

NOTE

1. Because of the limited number of pictures having particular sections ranked as least informative, four pictures were included for which the probed section's overall ranking was next to least. In each case, the section was one of the four center sections.

(Received for publication March 11, 1976; revision received July 28, 1976.)