COGNITIVE PSYCHOLOGY 3, 525–551 (1972)

Eye Fixations and Recognition Memory for Pictures[1]

GEOFFREY R. LOFTUS[2]
Stanford University

Three experiments were performed investigating the extent to which recognition memory for pictures can be predicted by eye-movement patterns on the picture at the time of study. In each experiment, 180 pictures were viewed followed by a yes–no recognition test on all the pictures. Eye movements were recorded at the time of study. Experiment I investigated payoff structure: It was found that higher-valued pictures both received more fixations and were remembered better than low-valued pictures, but when number of fixations was held constant, memory performance was independent of value. Experiment II showed that (a) when pictures are viewed for a fixed amount of time, memory performance is a positive function of number of fixations on the picture, (b) with number of fixations held constant, performance is independent of exposure time, and (c) there is no memory for pictures which were originally viewed only peripherally. In Expt. III, pictures were viewed either normally or while a distracting task (counting backward by threes) was being performed concurrently. The distracting task was found to reduce both number of fixations and memory performance for a picture. When number of fixations was held constant, performance was still better for normally viewed pictures, suggesting that the distracting task was doing more to inhibit encoding besides simply reducing the fixation rate.

A number of studies have concluded that long-term recognition memory for pictures is remarkably good. Shepard (1967) presented subjects (Ss) with 680 pictures; median performance on a later two forced-choice recognition test for these pictures was over 98%. An experiment by Nickerson (1964) yielded similar results: Performance on a yes–no recognition test for pictures was 90% even when 200 trials had intervened between the time the picture was tested and the time it was originally

---

[1] This paper is based on a PhD dissertation submitted to Stanford University. The author expresses thanks to Dr. Richard C. Atkinson, the chairman of his dissertation committee and also to Drs. Gordon H. Bower, Norman H. Mackworth, Roger N. Shepard and Edward E. Smith, the other members of his committee, for their constant advice, encouragement and friendship.

[2] Present address: Department of Psychology, University of Washington, Seattle, Washington 98105.

viewed. Standing, Conezio and Haber (1970) showed Ss 2560 pictures; again, subsequent recognition accuracy exceeded 90%. Haber (1970) in summarizing a series of experiments on picture recognition, remarked that, "These experiments . . . suggest that recognition memory for pictures is essentially perfect. . . . Recognition is based on some kind of representation in memory that is maintained without labels, words, or the need for rehearsal."

This may be overstating the case. In particular, the results of two recent studies suggest that conditions at the time pictures are studied can be manipulated such that recognition memory is affected rather dramatically. An experiment by Freund (1971) indicated that when Ss were forced to do a distracting task (counting backward by threes) when looking at pictures, performance was considerably reduced compared to a control condition in which the pictures were viewed normally. Potter and Levy (1969) manipulated the amount of time Ss had to study pictures and found that recognition accuracy ranged from about 15% when 125 msec was allowed for study to over 90% when the pictures were seen for 2 sec. Thus it appears that under some circumstances, recognition memory for pictures is far from perfect.

This paper attempts to specify some of the major variables affecting memory for pictures in terms of how these variables regulate the encoding processes carried out by a person at the time he originally views the picture. A convenient framework in which to view such processes is provided by a theory of memory proposed by Atkinson and Shiffrin (1968) which places a strong emphasis on *rehearsal* as a mechanism for transferring information to long-term store (LTS). Empirical support for this notion was provided by Rundus (1971), who employed a free-recall paradigm, and required Ss to rehearse out loud at the time of study. Rundus' results indicated that a number of free-recall phenomena (e.g., the serial position effect, the Von Restorff effect) could be parsimoniously accounted for in terms of the number of rehearsals accorded individual items. Thus the process of rehearsal which, in previous work, had to be inferred was made observable and turned out to be an excellent predictor of memory performance both for recall (Rundus & Atkinson, 1970; Rundus, 1971) and for recognition (Rundus, Loftus, & Atkinson, 1970).

How might this technique of operationalizing encoding processes be utilized in examining memory for pictures? For a variety of reasons it was decided in the present research to study the pattern of eye movements at the time of study and relate these patterns to subsequent recognition memory. The rationale for this strategy is discussed briefly below.

When a static visual scene is being viewed, eye movements over the scene take the form of discrete periods of relative immobility (fixations) separated by quick jumps (saccades). Pictures are generally scanned at the rate of about 3 fixations/sec (Buswell, 1935; Yarbus, 1967) and it is assumed that visual information is taken in and processed during periods of fixation, while vision is essentially suppressed during saccades (Woodworth & Schlosberg, 1954). Furthermore, it is generally the case that the part of the visual field falling on the fovea during a fixation corresponds to that area from which the observer is currently abstracting information, or "attending to." A considerable amount of research has been concerned with investigating fixation patterns on pictures, although none of this work has been directly concerned with memory for the pictures. Buswell (1935) conducted the first such systematic work, obtaining 1877 eye movement records of an individual scanning a picture. Buswell's results were of a qualitative and general nature, noting such things as the existence of "general areas of interest" in a picture, individual differences in scan patterns, and changes in the scan pattern with continued viewing of a picture. More recently, a good deal of research has been aimed at couching such observations in more quantitative terms. A number of studies have examined the pattern of fixation *vis à vis* "informative areas" in a picture. (Berlyne, 1958; Mackworth & Morandi, 1967; Mackworth & Bruner, 1970; Pollack & Spence, 1968). As an example, Mackworth's general technique was to divide a picture into an 8 × 8 in. grid. A group of Ss then rated how "informative" was each of the 64 squares where informativeness was defined in terms of ease of future recognizability. Subsequently, an *independent* group of Ss viewed the pictures and their fixation patterns were recorded. The results showed a high correlation between the informativeness rating of a particular square and the number of fixations on the square.

The work of Gould (Gould & Schaffer, 1965, 1967; Gould, 1967; Gould & Dill, 1969; Gould & Peeples, 1970) views the eye-fixation pattern as a direct reflection of cognitive processing. In these experiments, a 3 × 3 array of patterns was presented to S whose task was to count the number of outside patterns which were identical to the center (target) pattern. It was found that: (a) The average *number* of fixations as well as the average fixation *duration* was greater for targets than for nontargets and (b) For nontargets, fixation duration increased as a function of the number of features held in common by the target and nontarget. These results are directly analogous to those of reaction-time experiments which show that when S must decide whether two patterns are the same or different, reaction time is greater for "yes" than for "no" responses, and for "no" responses, reaction time increases as a function of target–nontarget similarity (e.g., Smith & Nielson, 1970; Tversky, 1969).

This research suggests that eye-fixation patterns may be profitably regarded as an observable basis for inferring internal processing of visual information. As indicated, whereas a good deal of work has examined various aspects of eye-movement patterns in some detail, none has systematically related these patterns to subsequent memory for the viewed material. The present experiments attempted to examine just such a relationship.

Although the nature of eye-movement patterns may be specified in terms of a number of parameters (cf. Mackworth & Bruner, 1970), the present research focuses mainly on two of them: *Number* of fixations (NF) on a picture and total fixation *time* (TFT) which is simply the total amount of time the gaze fell on the picture. These two parameters were selected because a good deal of theoretical, as well as empirical, work suggests that one or both of them may act as a predictor of recognition memory performance.

It has long been known that the time allowed for encoding material is an important determiner of memory for the material. This is the case in both short-term tasks (Pollack, Johnson, & Knaff, 1959; Pollack & Johnson, 1963; Mackworth, 1962; Norman, 1966) and long-term tasks (Bulgelski, 1962; Murdock, 1960; Johnson, 1964). Studies of picture recognition which manipulate study time have typically found that recognition accuracy increases with increasing study time (Potter & Levy, 1969; Freund, 1971). On the other hand, there is ample reason to assume that when pictures are viewed, NF may, independently of total time, play a strong role in building up memory for the pictures. Atkinson and Shiffrin (1968) assume long-term memory strength for an item to be a monotonically increasing function of the number of rehearsals accorded the item. Whereas it would be stretching matters to assume that an eye fixation is the visual analogue of a verbal rehearsal, it nonetheless might be assumed that information about a picture is transferred to LTS in discrete chunks, each chunk corresponding to a fixation. A recent model of visual processing (Gaarder, 1968) explicitly incorporates this assumption, and evoked potential evidence (Gaarder, Krauskopf, Graf, Kropfl, & Armington, 1964) is cited in support of it. Potter and Levy (1969) also suggest NF as being an appropriate predictor of memory strength: Although the variable actually manipulated in their experiment was exposure time, the authors draw the conclusion that "perception (of a picture) begins when a substantial visual event occurs and analysis and storage continue only until the next substantial visual change." Memory strength for a picture is assumed to be a positive function of the number of "substantial visual events" on the picture and "a substantial visual change sufficient to initiate a new unit of processing is usually brought about by a new eye fixation."

In this paper, three experiments on picture recognition are presented. In each experiment a particular variable was manipulated at the time the pictures were studied, and the variable was examined in terms of its effect on (1) recognition memory for the pictures and (2) the pattern of eye fixations at the time of study. Experiment I investigated payoff structure. Pictures were studied in pairs and each member of the pair was assigned a *value* which corresponded to the amount of money S would receive for correctly recognizing the picture. Ss were thus able to differentially attend to high vs low-value pictures. Experiment II had two independent variables: (a) exposure time was varied from 0.3 to 5.0 sec; and (b) for each exposure time, pictures were either fixated or viewed only peripherally. The latter manipulation was directed to the question of memory for nonattended pictures: Is there any memory for such pictures, and if so, does it increase with amount of exposure time? In Expt III, pictures were viewed either normally or while a concurrent distracting task (counting backward by threes) was being performed. In all three experiments eye movements were recorded at the time of study, noting in particular, the NF and TFT on each picture, and the results of major interest are the correlations of NF and TFT with subsequent recognition memory for the pictures. The main question is: To what extent can it be inferred that the effect of the independent variable is being mediated by the observable eye-movement parameters?

GENERAL METHOD

The general method common to all three experiments was the following:

Subjects. All Ss were female graduate and undergraduate students at Stanford who were paid for participating. None wore glasses or contact lenses and all had at least 20/20 vision. A total of 38 Ss was used, none of whom served in more than one experiment.

Stimulus materials. Three hundred and sixty naturalistic color photographs in the form of 35 mm slides were used as stimuli. The scenes were classifiable into the following types: 27% houses, 22% rural and forest, 17% seascapes, lakes and rivers, 10% buildings, 9% mountains, 7% city scenes and 8% miscellaneous. Pictures with easily memorizable parts (e.g., containing signs or people) were eliminated. The 360 pictures were randomly divided into two sets (A and B) of 180 pictures/set.

Design. Each experiment was divided into a study phase and a test phase. In the study phase, the 180 pictures in either Set A or Set B were presented as targets. The test phase was a yes–no recognition test of all 360 pictures. The test pictures were arranged randomly in six slide trays and the order of trays was random for each S. For each of the 360 pictures seen during the test, S was asked to make two (oral) responses. (1) A "yes" or a "no" recognition response was required. Before the test phase, Ss were reminded that the correct answer for half the slides was "yes" and were urged to give about half yes's and half no's over the course of the test. (2) Secondly, Ss were asked to indicate how they made their decision to say yes or no. They did this by saying "a" or "b" corresponding to one of the following two alternatives:

If S responded "yes" then the response "a" meant that there was some *particular* object in or attribute of the picture which she remembered. The response "b" meant that there was nothing in particular about the picture which S remembered, but it just "looked familiar."

If S responded "no" then the response "a" meant that there was some particular object in or attribute of the slide which S felt that she *would have* remembered had she seen the picture. The response "b" meant that there was nothing in particular that S felt that she would have remembered, but the picture just "looked unfamiliar."

If S responded "a" she was asked to name the item. Thus, typical responses might be "yes, a, I remember the fire hydrant," or "no, a, there's a red truck in this picture that I would have remembered."

Memory performance was measured in terms of (1) the probability of a "hit," $p(\mathrm{H})$, which was the probability of correctly saying "yes" given that a target picture was presented during study and (2) the probability of a "false alarm," $p(\mathrm{FA})$, which was the probability of incorrectly responding "yes" when a distractor picture was presented. In some cases, a single measure of memory strength d' from the theory of signal detectability was computed from $p(\mathrm{H})$ and $p(\mathrm{FA})$ using the tables from Elliott (1964).

Eye movement recording. During the study phase of the experiment (but not during the test phase) S's eye movements were recorded using a modified Mackworth stand camera (Mackworth, 1967). This camera utilized a corneal reflection technique and the output to a visual recording device consisted of (a) the same visual scene being viewed by S and (b) superimposed on this scene, a spot of light (fixation spot) whose location corresponded to the area on the scene which S was fixating (accurate to about ±0.5° of visual angle). In the present research, the recording device used was a closed-circuit television camera. During the time S's eye movements were being recorded, the scene and the fixation spot were visible to E via a TV monitor, and were simultaneously recorded on videotape. A chinrest, a forehead rest and biteboard were used to prevent head movements.

Statistical analyses. Most of the results to be reported in this paper involve functions relating some measure of performance y to number of fixations. The question of interest is: Can NF be used to predict y? Since NF was never directly controlled, Ss contributed unequally to the various abscissa points; normal statistical procedures were inappropriate because increases in the function could be spuriously obtained via subject-selection effects. Therefore, the following technique was used: (1) For each S, a modified Vincintizing procedure (Hilgard, 1938) was used to divide the abscissa into five points with approximately equal numbers of observations at each point. (2) The y-value for each point was computed. (3) A score was obtained by summing the cross products of the five y-values and the numbers, -4, -1, 0, 1, 4; this score is positive to the extent that the y-values increase monotonically (cf. Abelson and Tukey, 1970 for the rationale behind this). Thus a single score is obtained for each S and a sign test or a t test can be made of these scores against zero to test the monotonicity of the overall function. In the remainder of this paper, the assertion that "NF is a predictor of y", or that "y increases as a function of NF" means, unless otherwise stated, that a sign test for monotonic trend was significant beyond the 0.05 level.

EXPERIMENT I

In previous work using verbal paired-associates, it was found that assigning monetary incentive to items at the time of study produced large effects in memory performance (Loftus and Wickens, 1970). These investigators interpreted their results in terms of S's control processes at the time of study, concluding that more of S's limited processing capacity was devoted to high than to low-value items. In the present experiment, pictures were assigned different values, and Ss were permitted to differentially attend to high and low-value pictures. It was expected that high-value pictures would be both remembered better and looked at more than low-value pictures; the major results of interest are the functions relating memory performance to NF and TFT.

Method. Ten Ss were used. The basic methodological strategy was to present the 180 target pictures as 90 *pairs* at the time of study, each pair being shown for 3 sec. Thus, amount of time in the field of view was the same for all pictures; what varied was the NF and TFT accorded a particular picture. In an extreme case, for example, S could spend the 3 sec devoting 10 or so fixations to one member of the pair, and *no* time (or fixations) to the other member.

Before a pair was presented, each member of the pair was assigned a value. The left member of the pair was assigned one, five or nine points, and independently, the right member was assigned one, five or nine points. The value of a picture corresponded to the amount of money S would eventually receive for correctly identifying the picture, and Ss were paid according to the following scheme: When a target picture was presented for test, the original value of the picture was gained if S said "yes" and lost if S said "no". The 180 distractor items were worth five points apiece; thus, when a distractor was presented, five points were gained for saying "no" and lost for saying "yes." A quarter cent per point was then paid over and above a base sum of $3.00. Ss were informed of the payoff scheme at the start of the experiment.

Each of the 180 target pictures fell into one of nine study conditions defined by factorially combining three values of the picture itself with three values of the other member of the pair. On each trial, the values of the two members of the pair were selected randomly with the restriction that over an 18 trial block, exactly four pictures be in each of the nine conditions. Thus of the 180 target pictures, 20 were in each study condition.

The two pictures were presented side-by-side and were back-projected on a screen 28 in. from S's eyes. The pictures were separated by 1 in. and the entire width of the two pictures was 12 in., subtending a visual angle of approximately 24°. In the space between the two pictures was a 0.5 × 0.5 in. black rectangle which served as a fixation point.

The sequence of events on each of the 90 study trials was as follows:

1. S fixated the fixation point on an otherwise blank screen.

2. *E* read a pair of numbers. The first number corresponded to the value of the left member of the to-be-presented pair, and the second number corresponded to the value of the right picture.

3. After *E* read the numbers, S was permitted to look anywhere she liked, i.e., to where the left or right picture was about to appear or at the fixation point.

4. S pressed a button initiating presentation of the pictures. Two carousel projectors, each equipped with a Lafayette tachistoscopic shutter were used. Both shutters opened simultaneously and closed after 3 sec.

5. After the shutters closed, S looked immediately back to the fixation point. In order to control as well as possible the time spent encoding, Ss were instructed not to think about or rehearse pictures during the intertrial interval but to concentrate instead on staring at the fixation point. Ss reported that they were able to do this with no difficulty.

6. *E* advanced the slide trays, made any necessary calibrations in the eye camera, and the next trial began. (Calibrations were necessary

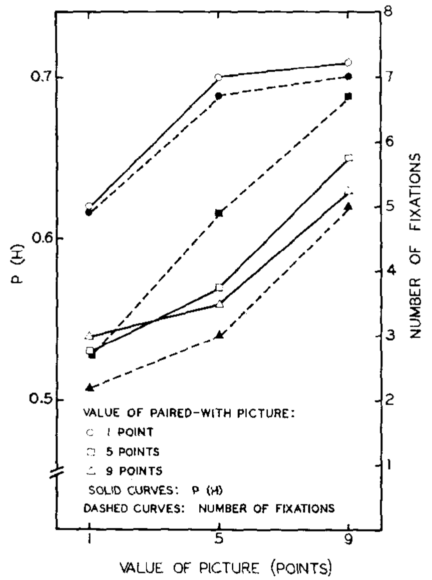

FIG. 1. Hit probability and mean number of fixations for pictures worth 1, 5 and 9 points. Separate curves are plotted for each of the three values of the paired-with picture. Each data point is the average of 300 observations.

about once every 5 to 10 trials due to slight head movements by *S*.) Average intertrial interval was about 3 or 4 sec.

Results and discussion. Figure 1 shows the hit probability (solid curves) and NF[3] (broken curves) for a picture as functions of the picture's value; the curve parameter is the value of the other member of the study pair. Relative value of the two pictures in the pair has similar effects on the two variables: The hypothesis that $p(\mathrm{H})$ *increases* monotonically with value, but *decreases* monotonically with the value of the paired-with picture is significant, $F(1,72) = 18.7$, $p < .01$, and accounts for 92% of the variance between the nine conditions. This same hypothesis applied to NF is again significant, $F(1,72) = 337.9$, $p < .01$, and accounts for 95% of the between condition variance. The value manipulation thus appears to be highly effective in controlling the dis-

---

[3] To analyze the eye-movement data, the videotape was played back at approximately 1/15 normal speed, and *E* sat at a computer console reading the sequence of fixations into the computer using a real-time program. This read-in was accomplished by pressing one of three keys corresponding to "left picture," "right picture," or "anywhere else" each time a saccade occurred. The technique of slowing down the videotape meant that the error in determining fixation durations (due to *E*'s reaction time in pressing the keys) could be kept reasonably small relative to the fixation duration itself.

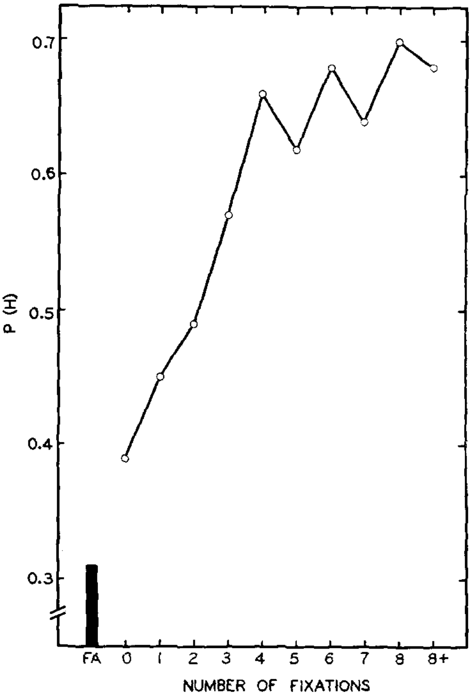

FIG. 2. Hit probability as a function of number of fixations. The bar at the left represents the false-alarm rate.

tribution of attention (fixations) over the pair of pictures, as well as response probabilities for the members of the pair.

Figure 2 shows $p(\mathrm{H})$ as a function of NF; for comparison $p(\mathrm{FA})$ is also shown by the bar at the left. A number of aspects of Fig. 2 should be noted: First, as expected, NF on a picture is a strong predictor of recognition memory performance on that picture. With zero fixations, $p(\mathrm{H})$ is 0.39; with eight or more fixations, $p(\mathrm{H})$ is 0.69. Second, the probability of saying "yes" to a picture which was seen but received *no* fixations (i.e., pictures seen only peripherally) exceeded by 0.08, the probability of saying "yes" to a picture which *in fact* was never seen (sign test, $p < .05$). This finding suggests that there is some memory for pictures seen only in the periphery.

It appears to be the case that the effect of value is mediated at least in part by NF, that is, higher-valued pictures are both fixated more and remembered better. The question now arises: Are Ss doing anything to remember the high-value items besides simply fixating them more? To answer this, the function relating $p(\mathrm{H})$ to NF was plotted separately

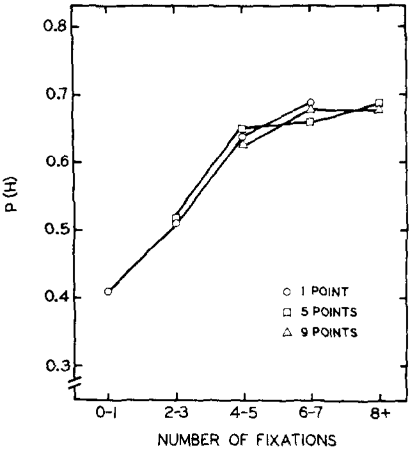

FIG. 3. Hit probability as a function of number of fixations. A separate curve is plotted for pictures worth 1, 5 and 9 points.

for pictures worth each of the three values. If NF is the only way in which effect of value is mediated, then for a given NF, $p(\mathrm{H})$ should be independent of value. Figure 3 shows that this is the case: The three curves corresponding to the three different values are virtually identical. The implications of this finding is discussed in more detail below.

Since Ss were free to look back and forth between the pictures, NF and TFT are correlated. In order to clarify the issue of which variable is a better predictor of memory performance, the following analysis was performed: Consider pictures which received exactly i fixations, where i ranges from 3 to 7. The median total time spent making the i fixations was computed, and pictures were divided into those which were viewed for longer and for shorter times than the median; $p(\mathrm{H})$ was then computed for these two subsets of pictures. This procedure yields a set of two-point functions which relate $p(\mathrm{H})$ to total fixation time with NF held constant. Table 1 shows these functions; as shown, when NF is held constant, $p(\mathrm{H})$ is essentially independent of time (for no value of i was performance for high viewing times significantly different from performance for low viewing times). The last column labeled "average" represents the mean of the five functions, weighted by the number of observations which went into each one. The two numbers in this column are almost identical. The implication of this analysis is that viewing time by itself is not a predictor of $p(\mathrm{H})$, but rather it correlates to the extent that it covaries with NF.

The next results concern fixation *durations*. A typical scanning rate for pictures is around 3 or 4 fixations/sec; how is this rate affected by

TABLE 1
Hit Probability for High and Low Viewing Times with Number of Fixations Held Constant[a]

	No. of fixations					
	3	4	5	6	7	Av
Viewing time						
High	.62 (120)	.66 (130)	.60 (128)	.68 (111)	.63 (108)	.643
Low	.53 (117)	.65 (132)	.61 (126)	.70 (137)	.63 (112)	.614

[a] The numbers in parentheses are the sample sizes for each cell.

the value of a picture? Table 2 shows average fixation duration on a picture as a function of value for pictures which received exactly i fixations where i ranges from 3 to 7. Neither the main effect of i, nor the $i \times$ value interaction is significant, $F(4,36) = 1.90$ and $F(8,72) = 2.01$. However, the main effect of value is significant, $F(2,18) = 3.92$, $p < .05$. This latter result is of some importance. It demonstrates that for a given number of fixations, the more time per fixation was spent the higher the value of the picture. But since for a given number of fixations, memory performance is *independent* of value (Fig. 3) it must be assumed that the extra time spent viewing the higher-valued pictures did not add anything to memory strength. This finding is consistent with the notion that NF as opposed to TFT is the appropriate predictor of memory strength.

The final data to be discussed from Expt I concern the probability of

TABLE 2
Average Fixation Duration (sec) as a Function of Values for Pictures Which Received Exactly i Fixations ($3 \leq i \leq 7$)[a]

Value (points)	No. of fixations				
	3	4	5	6	7
1	.292 (117)	.292 (111)	.290 (69)	.279 (49)	.300 (26)
5	.325 (85)	.311 (92)	.312 (93)	.304 (88)	.300 (69)
9	.350 (35)	.369 (59)	.336 (92)	.308 (111)	.311 (125)

[a] The numbers in parentheses are the sample sizes for each cell.

an "a" response, $p(\mathrm{a})$. As discussed above an "a" response means that there is some *particular* object in the picture which S remembers (given a "yes"). When a correct response is made, NF predicts $p(\mathrm{a})$ which varies from 0.47 for pictures with zero fixations to 0.69 for pictures with eight or more fixations. Two hypotheses suggest themselves to account for this finding. The first is that with more fixations on a picture, the probability increases that S will find some object that is easily memorizable. The second is that when looking at a picture, S finds an easily memorizable object or attribute rather quickly and subsequent fixations are spent committing the remembered object (RO) to memory. To the extent that there are more fixations on the RO, a better job is done committing it to memory.

To examine this problem in more depth, the following analysis was carried out: Out of the 3600 total responses, 250 satisfied the following criteria: (1) The test picture had been seen during study, (2) the picture had been given an "a" response, (3) the RO was a small, easily identifiable part of the picture (e.g., "the fire hydrant" or "the red flower"), and (4) a correct response ("yes") had been made. The pattern of fixations on these 250 pictures was examined noting which fixations had been on the RO and which had not. The probability that the RO had been fixated at least once by the *i*th fixation was 0.53, 0.86 and 0.95 for the first, second and third fixations on the picture, respectively. This probability reached 1.0 by the seventh fixation indicating that there was never a case when the RO was not fixated at all. This result is consistent with the findings of Mackworth and Morandi (1967) who concluded picture which is *in fact* recognized is almost always fixated within about 2 sec. These investigators define an "informative area" as one that Ss *think* is recognizable. The present results indicate that the area of a picture which is *in fact* recognized is almost always fixated within about 1 sec. The *unconditional* probability that the second and third fixations on the picture are on the RO is quite high: about 0.71. On subsequent fixations, this probability drops, but remains above 0.5 for all fixations.

Thus, for correctly recognized pictures, the RO is found very quickly, and a high proportion of subsequent fixations is spent on the RO. What is not clear is the nature of the cognitive activity being carried out during the fixations on the RO; many possibilities exist. Fixations may function in a way analogous to verbal rehearsals, corresponding to covert verbal repetitions of a label for the object. It would seem, however, that such repetitions could just as easily be made during one long fixation. Another possibility is that subsequent fixations may correspond to the abstraction of more features of the RO. The present data do not reject either of these possibilities.

EXPERIMENT II

The results of Expt I suggest that NF on a picture is an important predictor of recognition memory performance, and that TFT is a predictor only to the extent that it covaries with NF. In Expt I, however, TFT was controlled by S, since S was permitted to look back and forth between two pictures at will. In Expt II, TFT is experimentally controlled, which provides data bearing on the following questions: (a) When exposure time is held constant, how will $p(H)$ vary as a function of NF? (b) When different exposure times happen to yield the same NF, will $p(H)$ differ as a function of exposure time?

An additional purpose of Expt II was to gather more data concerning memory for pictures which reside in the field of view for some time but which are not fixated. Experiment I suggested that at least some memory strength accrues for pictures seen only peripherally. In Expt I, however, S decided which pictures she would and would not fixate. In Expt II, this process is experimentally controlled.

Method. Twenty Ss were used. During the study phase, the 180 pictures were again presented as 90 pairs and the two major independent variables were:

1. Exposure duration for a pair. Exposure times were 0.3, 0.7, 1.0, 3.0, and 5.0 sec.

2. Look vs peripheral. Before a pair appeared, S was instructed to look *only* at the left or the right member of the pair. Thus, of the 180 pictures, S fixated 90, and viewed the other 90 only peripherally. To insure that S would neither fixate, nor attend the peripheral member of the pair, she was led to believe that she would only be tested on the 90 looked-at pictures. This was accomplished by a cover story to the effect that the purpose of the experiment was to examine memory for pictures viewed with another, distracting, picture in the periphery. Postexperiment questioning indicated that all Ss believed the cover story.

During the study phase, the 90 pairs were presented in 15 blocks of 6 pairs/block. Within each block, the exposure condition remained the same, and Ss were informed at the beginning of each block what the exposure time would be for that block. The 15 blocks were divided into three groups, corresponding to the first, middle and last five blocks. The five blocks of each group consisted of a random permutation of the five exposure conditions. Before each pair was exposed, S was told whether to look at the left or right member of the pair. Left vs right was random with the restriction that for each S, half the pairs of each exposure condition be "left" and the other half be "right."

The sequence of events on a trial was as follows:

1. S fixated the fixation point.
2. E said "left" or "right."
3. S looked to where the appropriate picture was about to appear.
4. S initiated the trial, and the pair of pictures was exposed for a duration specified by the exposure condition.
5. When the pair disappeared, S immediately looked back at the fixation point. As in Expt I, Ss were instructed not to think about or rehearse the pictures during the intertrial interval.
6. E made any necessary calibrations advanced the slide trays, and the next trial began.

The test portion of the experiment was identical to that of Expt I.

Results and discussion. Of the 1800 pairs of pictures viewed by the 20 Ss, 36 pairs were viewed with at least one fixation on the member which had been designated as "peripheral." The following analyses excluded these 36 pairs.

Figure 4 shows NF[3] (broken curves) and *d′* (solid curves) as functions of exposure time for looked-at and peripheral pictures. Both these functions increase approximately linearly for pictures which were

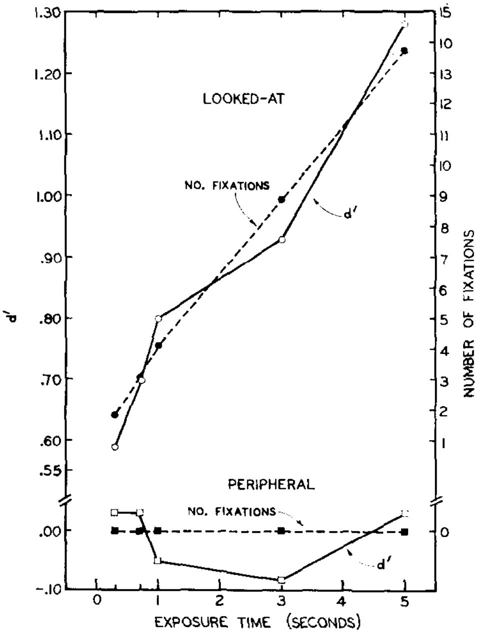

FIG. 4. Mean number of fixations and *d′* as functions of exposure time for looked-at and peripheral pictures. Each data point is the average of approximately 360 observations.

fixated; for peripheral pictures, NF is by definition, always zero, and d' is approximately zero for all exposure values. The hypothesis that d' increases linearly as a function of exposure time for looked-at pictures and is zero for all peripheral pictures is significant, $F(1,171) = 261.1$, $p < .01$, and accounts for 79% of the variance between the 10 conditions.

It will be recalled that in Expt I, $p(\text{H})$ was significantly above $p(\text{FA})$ for nonfixated pictures. In the present experiment, however, the finding that d' is equal to zero for peripheral pictures indicates that there is no memory whatsoever for such pictures. The reason for this discrepancy is not entirely clear. A possible explanation lies in the fact that in Expt I, Ss knew that they would be tested on all pictures, looked at or not. If the center of fixation does not necessarily correspond with the focus of attention then it is possible that Ss could if they wished, attend to peripheral material. In the present experiment, however, Ss believed that they would not be tested on the peripheral pictures; hence there was no reason to attend to them. In fact, many Ss spontaneously reported that they completely "blocked out" the peripheral member of the pair or "forgot it was there." Inasmuch as it may be safely assumed that no memory accrues for the peripheral pictures, subsequent discussion of Expt II concerns looked-at pictures only.

The solid curves of Fig. 5 present $p(\text{H})$ as a function of NF for each of the five exposure conditions. (For each of the five graphs, there are approximately equal numbers of observations at each abscissa point.) In each exposure condition, $p(\text{H})$ increases as a function of NF. For example, pictures viewed for 0.3 sec are recognized 51% of the time when one fixation is made on them but are recognized 62% of the time when

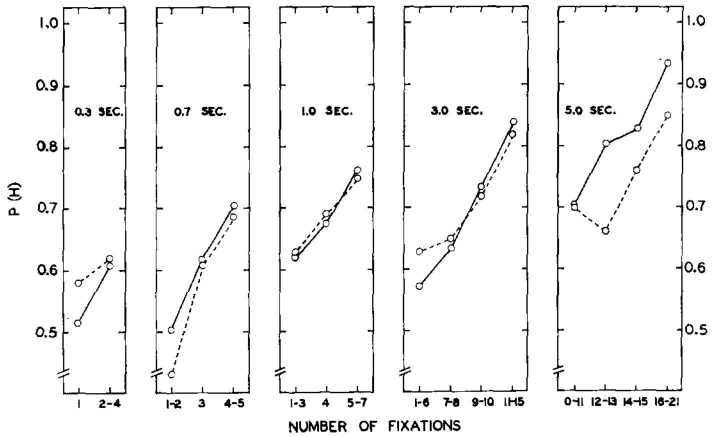

FIG. 5. Hit probabilities as functions of number of fixations for five exposure times.

more than one fixation is made. Analogously, for pictures viewed for 5 sec, $p(\mathrm{H})$ is 0.71 when fewer than 11 fixations are made; this probability rises to 0.93 when more than 16 fixations are made. It is of some interest to see whether the sharp increases of these five functions are due in part to subject selection effects, i.e., it is possible that Ss with good memories also have high fixation rates. For this reason, the functions were plotted in a different way: Each of the five curves was calculated for each S, and the curves were averaged over Ss. A subject selection hypothesis would predict that the resulting five curves would be flatter than the five unconditional curves. The broken curves of Fig. 5 show the result of this procedure: These are almost identical to the solid curves, and it is probably safe to assume that the increases are not due to subject selection.

Figure 5 demonstrates that for a fixed exposure time, performance increases with NF. To investigate the converse question—whether for a given NF, $p(\mathrm{H})$ should depend on exposure time—performance was examined where different exposure times happened to yield the same NF. These data are shown in Table 3 and the following should be noted: With 1–3 fixations, $p(\mathrm{H})$ is about the same whether those 1–3 fixations were made during 0.3, 0.7 or 1.0 sec. The same holds true when 10–12 fixations are made; $p(\mathrm{H})$ is about the same for the 3 and 5 sec curves. For 4–6 fixations, $p(\mathrm{H})$ is about the same for 0.7 and 1.0 sec; however, the point for the 3.0 sec curve is about 15% lower. The reason for this anomaly is unclear, but the result is nonsignificant (sign test) and is in

TABLE 3
$p(\mathrm{H})$ as a Function of NF for Cases When the Same NF Was Yielded by Different Exposure Times[a]

Exposure time (sec)	No. of fixations		
	1–3	4–6	10–12
0.3	0.59 (352)	—	—
0.7	0.58 (230)	0.70 (126)	—
1.0	0.62 (82)	0.71 (274)	—
3.0	—	0.56 (47)	0.77 (134)
5.0	—	—	0.76 (68)

[a] The numbers in parentheses are the sample sizes for each cell. No values are shown when the sample size was less than 40.

the opposite direction from what would be predicted by the notion that exposure time is the appropriate predictor of memory performance.

The results of Expt II thus provide support for the hypothesis that the effect of exposure time on recognition memory performance is mediated by NF.

EXPERIMENT III

Several experiments have been performed comparing recognition memory for pictures viewed either normally or during a concurrent distracting task (Freund, 1971; Szewczuk, 1970). In Freund's experiment, pictures were studied for 7 sec either normally, or while counting backward by threes. It was found that in a two forced-choice recognition test, response probability was 0.89 for normally viewed pictures and 0.65 for pictures viewed while counting backward.

What is the distracting task doing to reduce encoding? A partial answer to this question was suggested by some pilot data collected by the present author. Ss were asked either to count backward while viewing pictures or to view the pictures normally, and eye fixations were recorded. It was found that in the count backward task, there was about a 40% reduction in the fixation rate. Presumably, this effect obtained in Freund's experiment and since Expts I and II of the present research have demonstrated that NF is a predictor of memory strength, a performance decrement in the count-backward condition would be expected simply due to a reduction in NF.

The purpose of Expt III was to compare memory performance per fixation for pictures which have been viewed either normally or during a count-backward task.

Method. Eight Ss were run in Expt III. In the study phase of the experiment, each picture was viewed in one of four conditions:

1. Normal viewing for 2 sec (2N condition).
2. Normal viewing for 3 sec (3N condition).
3. Ss are required to count backwards while viewing for 3 sec (3C condition).
4. Count backward for 5 sec (5C condition).

It was expected that these conditions would yield distributions of NF per picture which would be approximately equal for normal and count-backward viewing.

At the beginning of an experimental session, Ss were given 30 practice pictures on which they learned to count backward by threes while simultaneously trying to memorize a picture and biting on a biteboard. The task was generally mastered without undue difficulty.

The 180 pictures presented during study were shown individually in

12 blocks of 15 trials/block. Within a block, the 15 pictures were all viewed in the same condition; Ss were informed at the beginning of each block what the condition was for that block. The 12 blocks were divided into three groups consisting of the first, middle, and last four blocks of the study phase. Each group was a random permutation of the four study conditions.

The configuration of the screen was changed somewhat from Expts I and II. The size of the picture remained the same (5.5 in. wide), but since in this experiment, pictures appeared individually, they were centered on the screen, and the fixation point was placed 0.5 in. above the center of the picture. The sequence of events on a trial was as follows:

1. S fixated the fixation point on an otherwise blank screen.
2. *E* read a three-digit number in the count conditions (said "ready" in the normal conditions).
3. S looked down to where the picture was about to appear and began to count backward by threes in the count conditions (looked down in the normal conditions).
4. *E* initiated the trial, the picture appeared for a length of time defined by the study condition, and then disappeared.
5. S looked back at the fixation point, and, in the count conditions, ceased to count backward. Again, S was requested not to think about or rehearse the pictures during the intertrial interval.

Results and discussion. Table 4 shows the hit probabilities, fixation rates and average NF[4] for the four conditions. For the 3N condition,

TABLE 4
Hit Probabilities, Fixation Rates and Average NF per Picture for the Study Conditions[a]

	Condition			
	2N	3N	3C	5C
p(H)	.67	.76	.52	.55
Fixation rate (fixations/sec)	2.88	2.67	1.70	1.35
Av NF	5.76	8.01	5.10	6.75

[a] Each cell contains 360 data points.

---

[4] Eye fixation data in Expt III was analyzed by slowing the videotape to approximately 1/10 normal speed and simply counting the NF for each trial. Small head movements during the count-backward task precluded an analysis of *where* in the picture S was looking; however, these head movements were easily distinguishable from saccades, and determination of NF was quite accurate.

$p(H)$ is 0.76, which is about the same as the corresponding hit rates for Expts I and II. For the 3C condition, however, the hit rate is 0.52; this difference is significant (sign test, $p < .01$). The results of Freund (1971) are thus replicated; a distracting task at the time of study impairs performance considerably. Table 4 also indicates that when Ss are required to count backward, the fixation rate is reduced by about one-third. For every S, the fixation rates for the 3C and 5C conditions were below those of *both* the 2N and 3N conditions.

Figure 6 shows $p(H)$ as a function of NF for each of the four conditions. A number of aspects of Fig. 6 should be noted. First, the results of Expts I and II are replicated. Performance in both the 2N and 3N conditions increases with increasing fixations, and for a given NF, $p(H)$ is approximately independent of exposure time. The same results generally obtain for the count conditions: For the 3C and 5C conditions, performance increases as a function of NF. For 4–6 fixations and for 7–9 fixations, performance is independent of exposure time. When 1–3 fixations are made, performance is better in the 3C than in the 5C condition; again, there is no ready explanation for this, but the difference is nonsignificant (sign test) and the hypothesis that performance is dependent on exposure time would predict a difference in the opposite direction.

Of considerable interest is the fact that performance per fixation is lower in the count than in the normal conditions; the difference between the two curves ranges from about 0.1 at 1–3 fixations to about 0.18 at

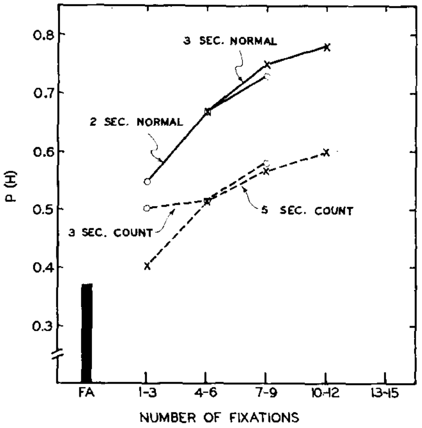

FIG. 6. Hit probability as functions of number of fixations with one curve for each of the four experimental conditions. The bar at the left represents the false-alarm rate.

10–12 fixations. Consider the curves for the 3N and 3C conditions: The average difference between the two curves is about 0.16. Whereas the unconditional difference of 0.24 between 3C and the 3N conditions is attenuated, it is by no means eliminated when NF is controlled for.

It thus appears that the counting task has other effects besides that of simply reducing the fixation rate. Several possible factors may be affected: First, many studies have underlined the importance of verbal encoding for remembering visual material (Kurtz & Hovland, 1953; Glanzer & Clark, 1964; Clark, 1965, 1968; Bahrick & Boucher, 1968; Dallet, Wilcox, & D'Andrea, 1968; Kaplan, Kaplan, & Samson, 1968; Petrinovich & Hardyck, 1970; Freund, 1971). The count-backward task used here is generally assumed to prevent verbal encoding or rehearsal; thus the difference between the normal and the count curves could represent the extent to which a verbal code comprises memory for pictures.

Another possibility is that the counting task affects *which parts* of the picture S fixates. It is typically the case that a small number of "informative areas" in a picture receive the great majority of fixations, and it is believed that peripheral mechanisms act to direct the gaze to such informative areas (Williams, 1966; Gould & Dill, 1969; Gould, 1969; Mackworth & Morandi, 1967; Mackworth & Bruner, 1970). Quite possibly, such peripheral directing mechanisms are inhibited by the counting task; it would follow that the fixations occurring in the presence of a distracting task would fall on relatively noninformative areas and would thus yield less information about the picture as a whole than fixations during normal viewing.

If either or both of these possibilities were true, then it would be expected that when a picture was tested, quite different information would be used as a basis for responding depending on whether the picture had been viewed in the normal or the count condition. If verbal encoding is prevented, then responses should be made on the basis of primarily nonverbal information. Similarly, if fewer noninformative areas have been fixated, then it might be expected that fewer responses would be made on the basis of small, definite (informative) objects in the picture. Data bearing on these predictions are shown in Table 5. Here, $p(\mathrm{a})$ given a correct response is shown as a function of NF for those pictures on which a correct response was made. The function is plotted separately for pictures viewed in the count or normal conditions. Whereas NF is a predictor of $p(\mathrm{a})$ in both cases, the curve for the count conditions is considerably lower; that is, when a correct response is made, S is much less likely to remember some particular object in or attribute of a picture than if it had originally been viewed during a dis-

TABLE 5
Probability of an "a" Response as a Function of NF for Pictures Which Were Correctly Recognized[a]

Conditions	No. of fixations 1–5	6–7	8–9	10–12
Normal	.51 (165)	.59 (269)	.60 (192)	.62 (94)
Count	.20 (307)	.32 (180)	.30 (131)	.32 (102)

[a] The function is shown separately for pictures in the normal and in the count conditions. The numbers in parentheses are the sample sizes for each cell.

tracting task. Support is thus given for both the possibilities outlined above. Ideally, the scan patterns should be compared for the two types of viewing; it could then be seen whether fixations during the counting task are, in general, on less informative areas. Unfortunately, unavoidable head movements by Ss during the counting made such an analysis technically unfeasible.

CONCLUSIONS

At the outset of this report it was noted that recognition memory performance for pictures has typically been very high. The present research does not attempt to solve the question of *why* such recognition memory can be so good; what it does do is to demonstrate that (a) a number of variables affecting traditional verbal memory affect memory for pictures in similar ways and (b) the effects of these variables may in some cases be parsimoniously explained in terms of their effects on observable eye movement processes—in particular, number of fixations—taking place at the time the picture is originally viewed.

The results of Expts I and II suggest that the effects of value and of exposure time are mediated by NF. In both of these experiments, the independent variable was shown to have a strong effect upon memory performance; however, when NF was held constant, the effect of the variable vanished. Apparently, Ss improve their performance on higher-valued pictures simply by fixating more times on these pictures, and, in general, utilize longer exposure times to make more fixations. Experiment II substantiates the suggestion of Expt I that exposure time in and of itself does not predict memory performance when NF is partialled out.

In Expt III, counting backward was shown to reduce the fixation rate; however this reduction was not sufficient to explain the performance decrement caused by the distracting task. The data of this experiment suggest that viewing while counting backward yields somewhat different information than normal viewing. In particular, verbal encoding may be reduced or eliminated, and the spatial scan pattern over the picture may be changed. These results point to the conclusion that other factors besides NF are important for remembering pictures, in agreement with the suggestions of other investigators (Noton, 1970; Mackworth & Bruner, 1970; Freund, 1971).

Number of fixations as a predictor of memory performance. Implicit in the above remarks is the notion that number of fixations on a picture is an important predictor of memory performance for that picture. A strong correlation between performance and NF was found in all three experiments, and this function was quantitatively as well as qualitatively very similar across three sets of experimental conditions, indicating that it is fairly robust. When only one or two fixations are given a picture, $p(\mathrm{H})$ is about 0.15 above $p(\mathrm{FA})$, whereas with 16 or more fixations, $p(\mathrm{H})$ exceeds $p(\mathrm{FA})$ by about 0.52.

The overall probability of a correct response is about 0.65 which is quite low compared to performance obtained in previous research. It appears, though, that this low response rate is simply due to the relatively low numbers of fixations accorded the pictures. For pictures which received high numbers of fixations performance was quite comparable to that obtained by Nickerson (1964) and Shepard (1967). Note that in these latter experiments, study time was quite long: Nickerson allowed 5 sec for viewing whereas in Shepard's experiment, study was self-paced and averaged about 6 sec/picture. The present results suggest that with less viewing time, NF and performance would have been considerably reduced.

What happens during a fixation? There is little doubt that many processes take place during an eye fixation. For example, memory codes are probably being formed in multiple (e.g., visual and verbal) modalities (Wallach & Averbach, 1955; Freund, 1971) and part of the fixation is probably spent making a decision about where to fixate next (Gould, 1969). Assume for the moment, however, that these processes may be viewed as yielding a single measure of memory strength (e.g., d'), and consider a hypothetical function relating this measure to fixation duration for a given fixation. The data of the present research provide some clues as to what this function might look like, viz, it should rise in some manner and then asymptote rather quickly. Such a function would account for the following findings: (a) In Expt I, higher-valued pictures received longer fixations than lower-valued pictures but, with NF held constant, were remembered no better. (b) For a fixed exposure time, performance increases as a function of NF and conversely, for a given NF, performance is independent of exposure time. If the assumption is made that all fixations are at least as long as the asymptotic duration value, then both these data follow.

A process which would predict such a function is the following: Assume that at the start of a fixation, S begins to abstract features from the area of the picture being fixated. If there were a finite number of features to be abstracted, and if sampling took place with replacement, then the function relating number of abstracted features to fixation duration would rise in a negatively accelerated fashion to some asymptote. A variation of this model is to assume that only *one* feature is abstracted during a fixation and that the amount of time necessary to abstract it is normally distributed. In this case, the function would represent *probability* that the feature is abstracted by time t, and the rising portion of the curve would be an ojive rather than negatively accelerated. With this latter model, the data described above are predicted independently of the function parameters; it need simply be assumed that a fixation does not terminate before the feature is abstracted.

It should be emphasized that the above discussion is highly speculative; the present data support *only* the notion that the function relating memory strength to fixation duration should asymptote rather quickly. Other types of research are necessary to determine the nature of processes occurring during a fixation and the functions that such processes would imply.

Correlational results. The major results reported in this paper have been of a correlational nature. There are a number of problems with such results stemming from the fact that it is difficult to assign causality to any one factor. In particular, it might be argued that the results in this paper are due to subject selection or item selection effects.

The types of statistical analyses used tend to rule out subject selection effects since the assertion that "y increased as a function of x" required that such an increase be shown by a significant proportion of Ss. Furthermore, the results shown in Fig. 5 suggests that increases in such functions are not even attenuated when variance due to subjects is removed.

Item selection effects are somewhat harder to rule out. In its extreme, an item selection hypothesis would predict that some pictures were both fixated more and remembered better than others, both effects being due to some third cause (e.g., "intrinsic interest" of a picture). Since there were many fewer Ss than items, it was not feasible to obtain, for example, a function relating $p(\mathrm{H})$ to NF for each individual picture. The best evidence against an item selection hypothesis comes from the data of Expt I: Here, NF on a picture was controlled to some extent via the value manipulation and the function relating performance to NF still increased quite sharply.

Obviously, however, the only way to completely eliminate problems due to correlational results is to experimentally control the variable of interest; controlling NF, for example could be done with an on-line computer system which would allow S some given NF on a picture before removing the picture. The prediction would then be that performance would increase as a function of NF but would be independent of TFT. This result combined with the results of Expt II of the present research would constitute substantial experimental evidence of a causal relationship between number of fixations accorded a picture and memory performance on that picture.

REFERENCES

Abelson, R. P., & Tukey, J. W. Efficient conversion of non-metric information into metric information. In E. R. Tufte (Ed.) *The quantitative analysis of social problems.* Reading, MA: Addison-Wesley, 1970.

Atkinson, R. C., & Shiffrin, R. M. Human memory: a proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.) *The Psychology of Learning and Motivation: Advances in Research and Theory.* New York: Academic Press, 1968. Vol. 2.

Bahrick, H. P., & Boucher, B. Retention of visual and verbal codes of the same stimuli. *Journal of Experimental Psychology,* 1968, **78**, 417–422.

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

Bugelski, B. R. Presentation time, total time and mediation in paired-associate learning. *Journal of Experimental Psychology,* 1962, **63**, 409–412.

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

Clark, H. J. Recognition memory for random shapes as a function of complexity, association value and delay. *Journal of Experimental Psychology,* 1965, **69**, 590–595.

Clark, H. J. Random shape recognition of brief exposure durations. *Psychonomic Science,* 1968, **12**, 245–246.

Dallett, K., Wilsox, S. G., & D'Andrea, L. Picture memory experiments. *Journal of Experimental Psychology,* 1968, **76**, 312–320.

Elliott, P. B. Tables of *d'*. In J. A. Swets (Ed.) *Signal detection and recognition by human observers.* New York: Wiley, 1964.

Freund, R. D. Verbal and non-verbal processes in picture recognition. Unpublished doctoral dissertation, Stanford University, 1971.

Gaarder, K., Krauskopf, J., Graf, V., Kropfl, W., & Armington, J. C. Averaged brain activity following saccadic eye movement. *Science,* 1964, **146**, 1481–1483.

Gaarder, K. Interpretive study of evoked responses elicited by gross saccodic eye movements. *Perceptual and Motor Skills,* 1968, **27**, 683–703.

Glanzer, M., & Clark, W. H. The verbal loop hypothesis: Conventional figures. *American Journal of Psychology,* 1964, **77**, 621–626.

GOULD, J. D. Pattern recognition and eye-movement parameters. *Perception and Psychophysics,* 1967, **2**, 399–407.

GOULD, J. D. Eye movements during visual search. Research Report No. RC2680. Yorktown Heights: IBM Thomas J. Watson Research Center, 1969.

GOULD, J. D., & DILL, A. B. Eye movement parameters and pattern discrimination. *Perception and Psychophysics,* 1969, **6**, 311–320.

GOULD, J. D., & PEEPLES, D. R. Eye movements during visual search and discrimination of meaningless, symbol and object patterns. *Journal of Experimental Psychology,* 1970, **85**, 51–55.

GOULD, J. D., & SCHAFFER, A. Eye-movement patterns in scanning numeric displays. *Perceptual and Motor Skills,* 1965, **20**, 521–535.

GOULD, J. D., & SCHAFFER, A. Eye-movement parameters in pattern recognition. *Journal of Experimental Psychology,* 1967, **74**, 225–229.

HABER, R. N. How we remember what we see. *Scientific American,* May, 1970, 104–112.

HILGARD, E. R. A summary and evaluation of alternative procedures for the construction of Vincent curves. *Psychological Bulletin,* 1938, **35**, 282–297.

JOHNSON, N. F. The functional relationship between amount learned and frequency vs rate vs total time of exposure of verbal materials. *Journal of Verbal Learning and Verbal Behavior,* 1964, **3**, 502–504.

KAPLAN, S., KAPLAN, R., & SAMPSON, J. Encoding and arousal factors in free recall of verbal and visual material. *Psychonomic Science,* 1968, **12**, 73–74.

KURTZ, K., & HOVLAND, C. I. The effect of verbalization during observation of stimulus objects upon accuracy of recognition and recall. *Journal of Experimental Psychology,* 1953, **45**, 157–164.

LOFTUS, G. R., & WICKENS, T. D. Effect of incentive on storage and retrieval processes. *Journal of Experimental Psychology,* 1970, **85**, 141–147.

MACKWORTH, J. F. Presentation rate and immediate memory. *Canadian Journal of Psychology,* 1962, **16**, 42–47.

MACKWORTH, N. H. A Stand camera for line-of-sight recording. *Perception and Psychophysics,* 1967, **2**, 119–127.

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 Psychophysics,* 1967, **2**, 547–552.

MURDOCK, B. B., JR. The immediate retention of unrelated words. *Journal of Experimental Psychology,* 1960, **60**, 222–234.

NICKERSON, R. S. Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology,* 1964, **19**, 155–160.

NORMAN, D. A. Acquisition and retention in short-term memory. *Journal of Experimental Psychology,* 1966, **72**, 369–381.

NOTON, D. A theory of visual pattern perception. *IEEE Transactions on Systems Science and Cybernetics,* 1970, **4**, 349–357.

PETRINOVICH, L. F., & HARDYCK, C. D. Generalization of an instrumental response between words and pictures. *Psychonomic Science,* 1970, **18**, 239–241.

POLLACK, I., & JOHNSON, L. B. Continuing memory span for digits. *Perceptual and Motor Skills,* 1963, **17**, 731–734.

POLLACK, I., JOHNSON, L. B., & KNAFF, P. R. Running memory span. *Journal of Experimental Psychology,* 1959, **57**, 137–146.

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

POTTER, M. C., & LEVY, E. I. Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology,* 1969, **81**, 10–15.

RUNDUS, D. J. An analysis of rehearsal processes in free recall. *Journal of Experimental Psychology,* 1971, **89**, 63–77.

RUNDUS, D., & ATKINSON, R. C. Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9**, 99–105.

RUNDUS, D., LOFTUS, G. R., & ATKINSON, R. C. Immediate free recall and three-week delayed recognition. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9**, 684–688.

SHEPARD, R. N. Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6**, 156–163.

SMITH, E. E., & NIELSON, G. D. Representations and retrieval processes in short-term memory: Recognition and recall of faces. *Journal of Experimental Psychology,* 1970, **85**, 397–405.

STANDING, L., CONEZIO, J., & HABER, R. N. Perception and memory for pictures: Single trial learning of 2500 visual stimuli. *Psychonomic Science,* 1970, **19**, 73–75.

SZEWCZUK, W. The role of verbalization in the retention of nonverbal material. *Polish Psychological Bulletin,* 1970, **1**, 87–97.

TVERSKY, B. Pictorial and verbal encoding in a short-term memory task. *Perception and Psychophysics,* 1969, **6**, 225–233.

WALLACH, H., & AVERBACH, E. On memory modalities. *American Journal of Psychology,* 1955, **68**, 249–257.

WILLIAMS, L. G. The effect of target specification on objects fixated during visual search. *Perception and Psychophysics,* 1966, **1**, 315–318.

WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental Psychology.* New York: Holt, Rinehart & Winston, 1954.

YARBUS, A. L. *Eye Movements and Vision.* (Translated by Basil Haigh). New York: Plenum, 1967.

(Accepted December 1, 1971)