Journal of Experimental Psychology
1969, Vol. 81, No. 1, 10–15

RECOGNITION MEMORY FOR A RAPID SEQUENCE OF PICTURES[1]

MARY C. POTTER[2] AND ELLEN I. LEVY

Massachusetts Institute of Technology

Memory for visual events occurring at and near the rate of eye fixations was examined. In Exp. I, *S*s were shown eight films of 16 unrelated pictures presented at ½, 1, 2, 3, 4, 6, or 8 per second. Later recognition ranged from 93% to 16%. In Exp. II with 32 *S*s, rates were mixed within each sequence to determine whether the probability of recognizing one item is independent of the presentation time of the previous item. The results support the hypothesis that rapidly presented pictures are processed one by one for precisely the time each is in view and are not held with other items in a short-term store as has been reported for verbal material.

When a person examines his visual surroundings, his eyes jump from point to point, resting on each point for only about ⅓ sec. A central problem of perception is to understand how such a rapid, discontinuous sequence of images is processed and acted on or stored. The experiments reported here use a sequence of unrelated pictures to simulate the input from successive fixations in the limiting case when the contents of each fixation are independent.

When a single picture of moderate complexity is shown tachistoscopically, it can be comprehended after a 50-msec. presentation (Vernon, 1954). When a simple item (a letter, number, or geometric figure) is followed by a visually noisy field, it may be masked wholly or in part if the onset to onset interval is 100 msec. or less; masking decreases at longer intervals and is slight beyond 300 msec. (see reviews by Raab, 1963; Kahneman, 1968). If the presentation is complex, e.g., an array of several digits, a mask presented after the stimulus produces a second kind of interference: *S* can report only a part of the array (Sperling, 1960). With a 300-msec. interval between the onsets of target and mask, *S* can report about as much as he can carry in immediate memory, so additional time does not improve his performance unless he has many seconds to "memorize" additional items.

Visual presentation of a sequence of letters or digits one at a time with immediate recall has shown that within the digit span, recall is accurate at rates of 300 msec. or longer per item (Mackworth, 1962). If item recall is scored independently of order recall, five out of six presented digits can be reported at a rate as high as 83 msec. per digit (Scarborough & Sternberg, 1967).

When digits are used as stimuli, the task is to identify and retain a small set of well-known items. Sequences no longer than the digit span are used in order to produce reliable recall and to reduce guessing corrections. It is not known to what extent results obtained in that situation can be generalized to the ecologically more frequent case in which items for recognition and retention are drawn from a large, poorly defined set and are presented in sequences in excess of the immediate memory span.

Shepard (1967) and Nickerson (1965) tested recognition memory for a series of several hundred pictures presented for about 5 sec. each. Nickerson used a yes-no procedure and obtained a hit rate of .87 with a false-alarm rate of .02. In Shepard's experiment, a forced choice between two alternatives gave a mean proportion correct of .97 (words and sentences used in two other

[1] This research was supported by a grant from the Joint Center for Urban Studies, Massachusetts Institute of Technology and Harvard University; by the United States Department of Health, Education, and Welfare, Office of Education (OE6-10-043, Project Z-014); and by a grant from the National Institute of Mental Health (MH-05120-05) to Harvard University, Center for Cognitive Studies.

[2] Requests for reprints should be sent to Mary C. Potter, Room 7-336, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

conditions were each less accurately recognized than were pictures).

In those studies, the rate of presentation was slow compared with the rate of eye fixations. A series of pictures can be retained remarkably well if each is looked at for 5 sec. or so; what happens at higher rates? The question explored in the following experiments is whether and how recognition memory decreases as rate of presentation varies between one picture every 2 sec. and one every $\frac{1}{8}$ sec. Three models of processing and retention are evaluated in the final discussion.

Experiment I

Recognition memory is examined for pictures presented at seven rates between 2 sec. and $\frac{1}{8}$ sec. per picture.

Method

Subjects.—The *S*s were 48 men and women undergraduate and graduate students at Harvard and Radcliffe.

Materials.—Eight short experimental films and one practice film, each composed of 16 still color photographs in sequence, were used in the experiment. The pictures were taken from magazines and included outdoor scenes, interiors, people, animals, food, etc. Highly similar pictures were avoided. All pictures were 4 × 6 in. The pictures were screened tachistoscopically in a preliminary study to make sure that they could be recognized readily after a single 80-msec. presentation preceded and followed by an illuminated gray field.

The 128 experimental pictures were assigned randomly to the eight films, with the restriction that no 2 pictures likely to be confused with each other could appear in the same film. An additional 16 pictures were used to make the practice film. The pictures in each group of 16 were randomized and photographed, two successive frames per picture, on 16-mm. Kodachrome movie film.

For the recognition task which followed the presentation of each film, the 4 × 6 in. originals of the filmed pictures together with an equal number of new pictures were used. All pictures were mounted on cardboard. The new pictures were not screened tachistoscopically, but were otherwise equivalent to those in the films. The 16 pictures used in each film together with 16 new pictures made up a recognition group; again, no obviously confusable pictures were placed in the same group. Each recognition group was randomized, with the restriction that of the first 16, 8 were old (from the film) and 8 were new (distractors).

Apparatus.—An L-W variable-speed 16-mm. movie projector with a 750-w. bulb was used to project the films on a lenticular screen 9½ ft. distant. The *S* sat directly in front of and below the projector. A 2-in. Ektanar lens (f/1.6) was used; the width of the picture subtended 11° of visual angle.

Procedure.—Each *S* viewed the practice film followed by the eight experimental films. He was instructed to pay careful attention to the film. After each film, he was asked to look through the recognition group at his own rate and say, to each picture, "yes" if he was sure he had seen it in the film, "maybe" if he thought so but was not positive, and "no" otherwise. He knew that only some of the pictures had appeared in the film.

The room light was turned off before each film and on within a few seconds after the film for the recognition task. For a few seconds before the film began, a dim greenish area was projected in the position where the pictures would appear, so *S* knew where to look.

Group A (32 *S*s) saw the experimental films at 125, 167, 250, and 333 msec. per picture (projection rates of 16, 12, 8, and 6 frames per second, respectively, with 2 frames per picture). Group B (16 *S*s) was shown the films at 250, 500, 1,000, and 2,000 msec. per picture (projection rates of 8, 4, 2, and 1 frames per second). The practice film was shown to Group A at 250 msec. per picture and to Group B at 500 msec. per picture. In other respects the design and procedure were identical for the two groups.

The order of presentation of the eight films (1–8 or 8–1 with each film reversed), the order of the four rates, and the order of the recognition group of pictures (1–32 or the reverse) were permuted so that each film was shown equally often at each rate, forward and reverse, etc. For any one *S*, each of the four rates occurred once among the first four films and once among the second four films (called first vs. second time in the report of results).

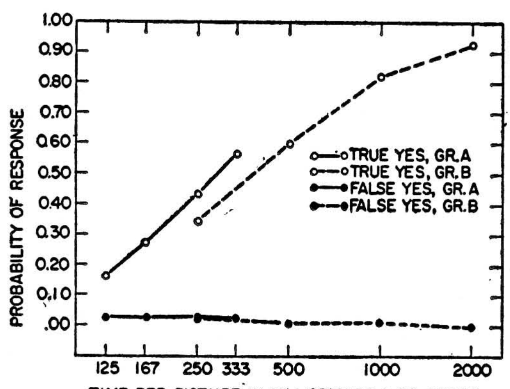

Fig. 1. Probability of "yes" to an old picture and "yes" to a distractor at each rate of presentation in Exp. I. (Each point is based on 1,024 responses in Group A, 512 in Group B.)

Results

Figure 1 shows the proportion of true and false "yeses" at each rate of presentation. In Group A, the overall proportion of true "maybes" was .17 and false "maybes" .13. In Group B, the proportions were .08 and .05.

An analysis of variance was performed for each group (A and B) on the sum of the yes and maybe responses corrected for chance in order to reduce differences between *S*s in response criteria. The analysis compared four film rates by first vs. second time replicated on each *S* (Winer, 1962, p. 213). Within each group, the overall effect of rate was significant ($p < .0001$) and Wilcoxon tests between adjacent rates were all significant beyond .01. The practice effect (first vs. second time) and the Rate × Practice interaction were not significant.

Thus, the effect of rate of presentation is clear-cut: the longer the viewing time of each picture, the greater the overall probability of recognizing it correctly thereafter. At 125 msec. per picture, the probability of true "yeses" is .16; at 2,000 msec. per picture, the probability is .93. The increment in the probability of recognizing falls off rapidly with increased time per picture, as can be seen in Fig. 1, where time is plotted on a log scale. The highest rate of storage occurs at presentation rates of 167, 250, and 333 msec.; at each of those rates, about two pictures are retained for every second of film time.

Group B, who saw films at 250-, 500-, 1,000-, and 2,000-msec. rates, did more poorly than did Group A at 250 msec. ($p <$ .01). For Group B, that rate was a very fast one among three much slower ones. Many *S*s in Group B reported that they tried naming the pictures as a way of remembering them, but at 250 msec. there was not enough time to do so. Perhaps the effort to name interfered with acquisition at that rate.

Serial position.—The most conspicuous serial position effect is the relatively high probability of recognizing the last picture in the film at rates of 333 msec. and faster. For example, at 125 msec. the probability of saying "yes" to the last picture was .66, compared with .13 for the other pictures. The processing of the last picture was only slightly affected by the length of time the picture was on the screen; much more important was the absence of a succeeding picture. The result is consonant with studies in which a single target is followed at various intervals by a mask (see Kahneman, 1968): the length of time that the target is exposed has much less effect on report than does length of time between target onset and mask onset.

A reliable although weaker tendency for the first picture in the film to be *less* well remembered than the others was present at all rates except Group B, 250 msec. The effect did not disappear with practice. Note that since the film order was reversed for half of the *S*s, the same pictures that were well remembered when they occurred last were poorly remembered when they came first. Apart from the first and the last picture, there was no other serial position effect at any of the rates.

Position in the recognition test.—Performance in the first half of the recognition test (eight film pictures and eight distractors) was compared with that in the second half. At the two fastest rates (125 and 167 msec.), performance declined significantly from the first to the second half. The proportion of corrected yes and maybe responses was .23 and .15 at 125 msec., .41 and .33 at 167 msec. At slower rates of presentation, there were no significant differences between halves. Apparently, test interference is greater when the task is more difficult.

Sequential effects.—Contingent probabilities of responding "yes" to a picture given that the previous picture in the film was a yes, maybe, or no were obtained, separately for Groups A and B and for each rate. At no rate of presentation was there a significant difference in the three probabilities. There was no suggestion, in short, that successful processing of one picture is associated with poorer processing of the following picture.

Picture variability.—The responses of the 32 *S*s of Group A to each individual picture were totaled across the four rates, omitting the first and last pictures of each film because of the marked serial position effects described earlier. Among the remaining 112 pictures the range of yes responses per picture was 2–25 (out of a possible 32) and the interquartile range was 7–17. The corresponding binomial distribution has an expected interquartile range of 8.7–13.3. Clearly the pictures were not all of equal difficulty, but neither were they remarkably diverse.

Experiment II

The effect of mixing four rates within a sequence is compared with that of using a single rate per sequence. The question is whether the probability of retaining a given picture depends only on the time it is in view or also on the time that temporally adjacent pictures are in view.

Method

The method was the same as that of Exp. I, except for the following.

Subjects.—Group A and Group B each had 16 *S*s.

Materials.—The same pictures in the same order were rephotographed for Exp. II, but two extra pictures were added to each film, one at the beginning and one at the end. They were not included in the analysis; their purpose was to eliminate from the analysis the serial position effects reported in the previous experiment. In the present experiment, rate was varied by varying the number of frames per picture: one, two, three, or four.

For Group A (the uniform-rates group) all pictures within each film had the same number of frames (i.e., rate). Of the eight experimental films, two were photographed at each of the four rates. Thus each film took between 2 and 8 sec., disregarding the first and last pictures. Two sets of films were made, each with a different order of rates.

For Group B (the mixed-rates group) the number of frames per picture varied within each film. Apart from the first and last pictures, there were four pictures at each of the four rates in each film, and thus all films took 5 sec. The order of rates was randomized with the limitation that two at each rate would appear in each half of the film. The first and last pictures also varied in number of frames. In a second set of films a different order of rates was used within each film.

The recognition pictures were the same, in the same order, as those in Exp. I. The two extra pictures in each film were added at the end of the recognition task in each case.

Procedure.—The films were projected at eight frames per second, so each picture was in view for 125, 250, 375, or 500 msec. In each group, half of the *S*s saw one set of films, half the other.

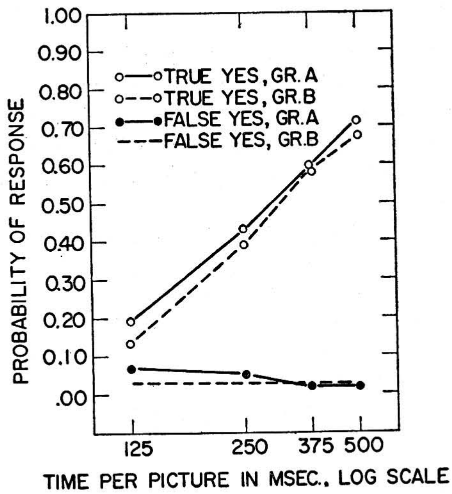

Fig. 2. Probability of "yes" to an old picture and "yes" to a distractor at each rate of presentation in Exp. II for Group A (uniform rate) and Group B (mixed rates). (Each point is based on 512 responses except that the false "yes" probability for Group B is based on 2,048 responses.)

The two directions of presentation of the films and of the recognition pictures were also permuted, as in Exp. I.

Results

Figure 2 shows the proportions of old and new pictures to which the response "yes" was made. An analysis of variance on chance-corrected scores again found the effect of rate highly significant within each group ($p < .001$), and t tests between adjacent rates were all significant beyond .01.

Uniform vs. mixed rates.—As can be seen in Fig. 2, the results of the two conditions were exceedingly close. The proportion of "yeses" was compared at each rate using the Mann-Whitney U test, and no significant differences were found.

Sequential effects.—The relation between presentation time of a given picture in Group B and proportion of "yeses" to the next picture in the film was examined. The proportion of "yeses" averaged over the four presentation times was .46 after a previous picture of 125 msec., .44 after 250 msec., .49 after 375 msec., and .51 after 500 msec. The differences were not significant nor was there any suggestion of an interaction between the presentation time of one picture and the next.

Conclusion.—The probability that a given picture will be retained depends strongly on the time it is in view (as in Exp. I), but is independent both of the time taken by the whole sequence and of the time that the just-previous picture was in view.

Discussion

Temporal units of processing.—A major question for a theory of perception and memory concerns the temporal nature of the units of intake and analysis. One possibility is that intake is interrupted temporarily during some phase of analysis of a given perceptual event: perceiving alternates with analysis at fixed or variable intervals. A prediction of the hypothesis is that a temporal sequence of items will be perceived and missed in alternation at some high rate of presentation. That prediction received no support in the present experiments; if there is a phenomenon of "processing blindness," it is too slight to account for the large losses in retention observed at the faster rates of presentation.

The hypothesis just described should be distinguished from the "psychological moment" hypothesis which proposes that there exists a natural temporal unit of intake of 50–100 msec. which is analyzed as a whole (see Eriksen & Collins, 1967, for a critical discussion with references). The hypothesis does not explain the present results since the times used here were all considerably longer than the "moment."

A second relevant hypothesis is that items enter a temporary store of limited size as they are presented and then are removed to a more permanent store or are lost out of the temporary store probabilistically over time without regard to their order of entry. Such a model was proposed by Waugh and Norman (1965); supporting evidence has come from experiments with words, digits, and other verbal materials. If items enter the temporary store faster than they can be taken into permanent storage, some will begin to be lost as soon as the temporary store is full, but not necessarily in the strict alternating pattern of Hypothesis 1. At least a small negative sequential dependency would be expected, nonetheless, since the probability that Item $n+1$ will be selected for permanent storage will be less if n is selected than if it is not. Many of the present experimental results are consonant with this view since it is possible that a small sequential effect could have been obscured by picture variability. However, the first few and the last few pictures in a sequence did not show an advantage in recognition, in serious contradiction to the hypothesis. The marked advantage found for the last picture only would also require some special assumption.

A third hypothesis is that perception begins when a substantial visual event occurs, and analysis and storage continue only until the next substantial visual change. An event will be retained if the analysis has gone on "long enough" before the next event interrupts it. If long enough is a constant for all *S*s and all kinds of visual material, then the prediction is simple and, as we have seen, false: slower than some particular rate, everything will be retained; faster than that rate, nothing will be retained.

Long enough, however, might vary from item to item and person to person. In that case, the following predictions can be made: (*a*) The faster the rate of presentation, the fewer the items retained—a prediction shared by the three hypotheses; (*b*) the probability of retention of a given item will be independent of the retention or nonretention of the previous item and the time the previous item was in view; (*c*) the last item in a sequence will be well-retained because no substantial visual event follows it; and (*d*) no other serial position effect will be found.

All four predictions of the independence hypothesis were supported by the results of the present experiments, with one exception: The first picture in a series did poorly. That result is not predicted by any of the theories.

In conclusion, the hypothesis that pictures in sequence are processed one by one for exactly the duration of presentation appears to be the most adequate of the three hypotheses considered.[3] The short time required for perception and storage of most pictures (half required 250 msec. or less), despite their complexity and unfamiliarity, is a powerful demonstration of the efficiency of recognition and retention of meaningful visual material.

Eye movements as units.—A general question that arises from the independence hypothesis concerns what constitutes a substantial visual change. In the present experiments, the onset of each picture is undoubtedly such a change, whereas the onset of the uniform greenish field that follows the last picture in a sequence is not. Natural temporal breaks in visual stimulation,

[3] The real time at which processing occurs is, of course, later than the time when the stimulus arrives at the retina, but that constant displacement is irrelevant to the hypotheses under consideration.

such as a blinking light, are the exception in ordinary experience; most events consist of continuous movements in or through the visual field. A substantial visual change sufficient to initiate a new unit of processing (the authors propose) is usually brought about by a new eye fixation. A new fixation is sufficient but not necessary to initiate processing, which can also be initiated by a change in the stimulus during a single fixation.

These considerations raise the question of what happens to an *S*'s eye movements as he views the films in the present experiments. Eye movement records on two individuals were obtained,[4] and in both cases, eye movements were spontaneously and almost completely suppressed during the 2–4 sec. required to present a film at rates between eight and four pictures per second. At slower rates, eye movements were more frequent. The tentative conclusion is that eye movements within the central area of the visual field are only likely to be initiated when the stimulus itself remains unchanged for ½ sec. or longer.

In sum, it is proposed that the eye remains relatively stationary to process stimuli at one locus that are changing at a high rate. With an unchanging or slowly changing visual array, eye movements create the visual change required to initiate a new unit of processing at efficient intervals.

[4] The authors thank Colwyn Trevarthan for obtaining the eye movement records.

REFERENCES

Eriksen, C. W., & Collins, J. F. Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology,* 1967, **74**, 476–484.

Kahneman, D. Method, findings, and theory in studies of visual masking. *Psychological Bulletin,* 1968, **70**, 404–425.

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

Nickerson, R. S. Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology,* 1965, **19**, 155–160.

Raab, D. H. Backward masking. *Psychological Bulletin,* 1963, **60**, 118–129.

Scarborough, D. L., & Sternberg, S. Processing items and their order in sequential visual displays. Paper presented at the meeting of the Eastern Psychological Association, Boston, April 1967.

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

Sperling, G. The information available in brief visual presentations. *Psychological Monographs,* 1960, **74**(11, Whole No. 498).

Vernon, M. D. *Visual perception.* Cambridge: Harvard University Press, 1954.

Waugh, N. C., & Norman, D. A. Primary memory. *Psychological Review,* 1965, **72**, 89–104.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

(Received August 14, 1968)