Journal of Experimental Psychology: Human Learning and Memory
1975, Vol. 104, No. 2, 103–113

Two Types of Information in Picture Memory

Geoffrey R. Loftus and Susan M. Bell
University of Washington

It is assumed that recognition memory for pictures is based on two types of information. The first is information about specific details in a picture. The process of encoding this type of information is identified with what N. H. Mackworth and others have termed looking at "informative areas" in pictures. The second informational component is designated as "general visual information." Two experiments were carried out investigating (a) the extent to which recognition responses to pictures are based on specific detail vs. general visual information, (b) whether the amount of specific detail information may be manipulated by varying the complexity of a target picture, and (c) the rate at which the two types of information are acquired. The results indicate that the rate of encoding specific details varies with the number of potential informative areas in a picture and, given that a detail is encoded, memory performance is not substantially affected by target complexity, exposure time, or presence or absence of a mask.

Consider the task of a subject who is viewing a novel visual scene with the intention of subsequently being able to recognize it. Recognition of a picture essentially involves the ability to *discriminate* a picture from other similar pictures.[1] Thus, a reasonable viewing strategy would involve two steps. First, a decision must be made about which object or area of a picture will best distinguish the picture from other pictures in its class. Second, once this decision has been made, the existence and relation of the candidate object to the rest of the picture must be encoded.

A useful tool for examining viewing strategies is provided by a subject's pattern of eye fixations over a picture. Since the pioneering work of Buswell (1935), it has been known that fixations are concentrated on "general areas of interest" in a picture. More recent studies of viewing behavior have attempted to couch this conclusion in somewhat more quantitative and operational terms. Working within the framework of information theory, Berlyne (1958) displayed pairs of simple visual stimuli, varying the informational content of one member of the pair relative to the other. In two experiments, subjects spent considerably more time looking at the more informative as opposed to the less informative member of the pair. The work of Mackworth (Mackworth & Bruner, 1970; Mackworth & Morandi, 1967) employed a technique in which a picture of a complex, naturalistic scene was divided into an 8 × 8 in. (20 × 20 cm) grid. A group of subjects viewed each of the 64 squares individually and rated how "informative" each square was. An independent group of subjects then viewed the complete picture, and the number of eye fixations on each square was recorded. Two interesting findings emerged from this procedure. First, a high correlation was found between informativeness rating of an area and the number of fixations made on that

This research was supported by a National Science Foundation grant to the first author. We gratefully acknowledge the comments and suggestions of Buz Hunt, Beth Loftus, Colin MacLeod, Tom Nelson, and Steve Poltrock. Tom Nelson created the stimuli used in the present experiment and kindly lent them to us.

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

[1] This notion naturally assumes that the picture being viewed is a member of some known, reasonably well-defined class of pictures. In a typical picture-recognition experiment, the class quickly becomes apparent to a subject via experimental instructions, warm-up pictures, or the first few pictures of the study sequence. Thus, the class of pictures might be naturalistic scenes, faces, common objects, etc.

area. A large proportion of the fixations were on a relatively small number of areas, which Mackworth dubbed "informative areas." Second, subjects tended to fixate on these informative areas very quickly—typically within one or two fixations following the onset of the picture.

Mackworth's definition of an informative area is highly empirical: An area is defined to be informative to the extent that subjects *say* it is informative. Berlyne's (1958) work, on the other hand, suggests that subjects tend to look at those areas of a visual scene that are defined a priori to be informative in an information–theoretic sense. Combining these two results, it seems reasonable to expect that those areas in a picture that subjects call informative are also informative in an information–theoretic sense. More precisely, an area or detail in a picture may be defined as informative to the extent that it has a low a priori probability of being there, given the rest of the picture and the subject's past history. Thus, for example, in a picture depicting a farm scene, a tractor would be an uninformative detail, whereas an octopus would be an informative detail. Conversely, in a picture depicting an underwater shipwreck scene, an octopus would be uninformative, whereas a tractor would be informative.

To the extent that some detatil has a low a priori probability of being in a picture, it will, by definition, provide a useful basis for discriminating the picture from other pictures in its class. Given the above example of a farm scene, a subject will lay a much better groundwork for subsequently recognizing the picture by encoding the fact that there is an octopus in the farm than by encoding the fact that there is a tractor. This is because potential distractor pictures of farms will be more likely to contain tractors than octopuses. A study by Loftus (1972, Experiment 1) provides evidence that subjects do in fact use informative areas of a picture as a basis for subsequent recognition. Loftus's study involved an initial study phase in which subjects were shown a series of target pictures. During this study phase, eye fixations on the pictures were recorded. Following the study phase was a *yes–no* recognition test. On each test picture, the subject was given an opportunity to report whether he was basing his response on some particular detail in a picture or whether he was responding simply on the basis of the picture's "familiarity." Thus, a report based on a detail might be something like "Yes, this is a target picture; I remember the fire hydrant in the lower right-hand corner." Pictures that yielded such reports were then analyzed in terms of the fixation patterns at the time they were originally viewed. The results showed that (a) a remembered detail had invariably been found by the first or second fixation on the picture and (b) a remembered detail had received more than 50% of all the fixations on the picture. Thus, details on which recognition responses were based were operationally equivalent to what Mackworth defined as an informative area.

The preceding discussion has focused on what we shall label a *specific detail* component of picture memory. Subjects are seen as scanning the picture in search of those details that will be maximally efficient in distinguishing the picture from other pictures. It is likely, however, that recognition of a picture is also based on another type of information, which we shall refer to as *general visual* information. Evidence for this contention stems from an experiment by Bahrick and Boucher (1968) in which subjects were shown pictures of common objects and then were tested both for recall of the object names and for recognition of the original picture. In the recognition test, the distractors consisted of objects having the same name (e.g., a target picture of a coffee cup might appear among a series of distractor coffee cups differing in such things as the height of the cup or the size of the handle). Recall and recognition performance were found to be independent, suggesting that they were based on different types of information. We argue that in Bahrich and Boucher's experiment, recall performance was based on specific detail information, whereas recognition performance was based on general visual information.

The present study includes two experiments on picture recognition, carried out for several reasons. The first reason was to test the extent to which recognition decisions are based on specific detail vs. general visual information. Thus, at the time of recognition, subjects were asked whether they were basing their responses on some specific detail or whether they were responding merely on the basis of the picture's general familiarity. Second, if memory for specific, informative details is important in picture recognition, then the more details a picture contains, the higher should be the probability of encoding a detail and, thus, the higher should be the eventual recognition performance. To test this hypothesis, each stimulus picture was varied in terms of the number of potential informative details it contained. Finally, it was of interest to determine the rate at which specific detail and general visual information is acquired. To investigate this question, exposure time of the pictures at the time of original study was varied.

Method

The two experiments were sufficiently similar to each other to be described together.

Stimuli

The stimuli consisted of 120 triads of pictures used in a previous picture recognition experiment (Nelson, Metzler, & Reed, 1974). The three pictures in a triad were constructed in the following way: The first member was a *photograph* of a naturalistic scene; the second member was an *unembellished line drawing* containing only the central information in the photograph; finally, an *embellished line drawing* was made by taking the unembellished line drawing and adding to it some of the details from the original photograph. Each of the three pictures in a triad was made into an individual 35-mm slide, and the 120 picture triads were randomly divided into two sets, Set A and Set B of 60 triads per set.

General Paradigm

Each experiment consisted of a study phase immediately followed by a test phase. In the study phase, 60 target pictures—1 picture from each of the 60 triads of either Set A or Set B—were viewed one at a time by subjects run in groups of five. In the test phase, these same 60 target pictures were randomly permuted and randomly intermingled with 60 distractor pictures—1 member of each of the 60 triads from the other set. The resulting 120 pictures were then shown one by one in a *yes–no* recognition test.

Both experiments were repeated measures designs, varying stimulus type and exposure time during the study phase. Experiment 1 was a 3 (stimulus types: photograph, embellished line drawing, unembellished line drawing) × 5 (exposure times) × 90 (subjects) design. Experiment 2 was a 2 (stimulus types: photograph, unembellished line drawing) × 5 (exposure times) × 100 (subjects) design. In both experiments, each of the 360 pictures (120 triads with 3 pictures per triad) was used equally often as a target and a distractor.

Subjects

All subjects were University of Washington undergraduates who participated for extra course credit in introductory psychology classes.

Apparatus

A Kodak random-access slide projector was used to display the stimuli. Timing was controlled using Gerbrands tachistoscopic shutters with rise and fall times of approximately 5 msec.

Procedure

The pictures were projected on a screen and subtended a visual angle of approximately 10°. Average luminance was 3.29 cd/m² for photographs, 4.35 cd/m² for drawings and .53 cd/m² for the pre- and postexposure fields.

Study phase. During the study phase, subjects saw one member of each of the 60 picture triads in either Set A or Set B. Each of the 60 study trials consisted of the following steps:

1. The experimenter said "ready" and the subjects fixated on a small *x* in the center of the screen.

2. The picture was displayed for one of five exposure times: 60, 100, 250, 350 or 500 msec. At this point, the critical difference between Experiment 1 and Experiment 2 occurred. In Experiment 2, a 1-sec random-noise mask immediately followed the offset of a stimulus picture. The mask, projected via a second projector, appeared in the same spatial location as the stimulus picture that had just vanished. The purpose of the mask was to prevent processing of the picture from iconic store (Neisser, 1967; Sperling, 1960, 1963) following the physical offset of the picture.[2] In Experiment 1, no mask followed the stimulus; rather, each picture was simply followed by the postexposure field.

3. An intertrial interval followed during which the slide projector was changed to the next target picture. The average intertrial interval was about 4 sec and did not depend systematically on any of the conditions.

---

[2] The nature of the mask was such that when the mask and a picture were superimposed, it was virtually impossible to extract any information from the picture.

Test phase. The test phase consisted of 120 test trials. On each test trial, either one of the identical targets from the study phase or a distractor picture was shown. A test trial consisted of the following steps: (a) The experimenter read the number of the test trial, and the subjects located the appropriate space on their response sheets. (b) The test picture remained on for 12 sec while the subjects made their responses.

On each test trial, a subject made three responses. First, he responded "yes" or "no" corresponding to whether or not he thought the test picture had been one of those presented at study. Second, he made a confidence rating ranging from one ("practically certain") to three ("practically guessing"). Finally, the subject was asked to indicate the basis of his response and was given two choices: (a) He was responding on the basis of some *particular detail* in the picture or (b) he was responding simply on the basis of the "familiarity" of the picture. If he indicated that he was responding on the basis of some detail, he was asked to write the name of the detail on his response sheet.

Subjects were told at the beginning of the test phase that for half of the pictures, the correct answer would be "yes," and they were urged to make about half "yes" and half "no" answers. Additionally, they were urged to use the three confidence ratings about equally often.

Designs

Experiment 1. Eighteen groups of five subjects per group were run in a 3 × 5 × 90 factorial design. During the study phase, all three members of a given triad were seen over 3 groups and each of these groups was mirrored by another group that differed only in that the target/distractor set was reversed, resulting in 6 groups. Each stimulus was randomly assigned three presentation times, rather than occurring at each of the five possible presentation times. Thus, each of the 6 groups was mirrored by 2 additional groups in which the presentation time of a given item was varied, thereby producing the total 18 groups. The same study presentation order was maintained over a block of 3 groups that differed only in terms of which member of a given triad was seen. Thus, for example, if the first study item in Group 1 had been the photograph from a particular triad, the first study item in Group 2 would be the embellished line drawing from that triad, and the first Group 3 study item would be the unembellished line drawing from that triad. Within this presentation sequence, the order of occurrence of the three stimulus types was counterbalanced and the order of the five presentation times was block randomized. Each group saw an equal number of slides in each of the 15 combinations of stimulus type and presentation time.

The order of test trials was the same for a block of three groups over which stimulus type was varied and for the three additional groups resulting from reversing target/distractor set. Each test order was a random permutation of 120 pictures—1 picture from each of the 120 triads. A given group always saw at test the identical 60 slides seen at study, randomly intermixed with 60 distractors that included 20 instances of each of the three stimulus types.

Experiment 2. Experiment 2 was similar in design to Experiment 1, differing primarily in the fact that only two stimulus types (photographs and unembellished line drawings) were used and in that each stimulus appeared in all of the 10 possible stimulus type/exposure time combinations. Twenty groups of five subjects per group were run, which also allowed each stimulus to appear equally often as a target and a distractor. Within a group, order of presentation time was block randomized and equal numbers of photographs and drawings were shown in a random order with the restriction that no more than three instances of the same stimulus type would occur successively. Due to a design error, there were not equal numbers of stimuli in each of the 10 stimulus type/presentation time treatments within a given group. However, for each group there corresponded another group that had the same presentation order and differed only in that the stimulus types of each item were reversed. Each such *pair* of groups had the same number of stimuli in each of the 10 conditions.

Five random test orders were prepared. A given test order was used for a group and the corresponding group that had seen the other target/distractor set at study. Two additional groups over which stimulus type varied saw test items in the same order but, of course, photographs were substituted for drawings and vice versa in order that a target be of the same stimulus type at study and at test.

Results

Performance Measure

The data were analyzed within the framework of signal detection theory (Anderson & Bower, 1972; Kintsch, 1969; Egan, Note 1). The method used to compute d' for a given condition is described in the Appendix.

Statistical Analyses

Since the design was within subjects, an individual subject would ordinarily be used as the unit of analysis in statistical tests. However, using the method described in the Appendix to obtain d' scores, not enough data were obtained from each subject to compute a d' for each condition. Therefore, data from two groups of subjects were combined into a single unit of analysis. Accordingly, Experiment 1 had 9 such units and Experiment 2 had 10 such units.

Exposure Time and Stimulus Type

In Experiment 1, the mean d' values for photographs, embellished line drawings, and unembellished line drawings were 2.35, 1.57, and 1.59, respectively. Since performances on the two types of line drawings were virtually equal and since they did not differ significantly ($t < 1$), they were collapsed together for subsequent analysis. Future discussion will distinguish only between photographs and drawings for both experiments.

Figure 1 shows d' as a function of exposure time, the curve parameter being stimulus type. The dashed curves represent data from Experiment 1 (unmasked) pictures, whereas the solid curves represent data from Experiment 2 (masked) pictures. In both Experiment 1 and Experiment 2, there are significant effects of exposure time, $F(4, 8) = 20.6$, $p < .05$, and $F(4, 9) = 84.9$, $p < .05$, respectively; of stimulus type, $F(1, 8) = 21.6$, $p < .05$, and $F(1, 9) = 27.6$, $p < .05$, respectively; and of Exposure Time × Stimulus Type interaction, $F(4, 32) = 2.85$, $p < .05$, and $F(4, 36) = 4.40$, $p < .05$, respectively. Additionally, unmasked pictures showed better performance than masked pictures, $t(17) = 2.89$, $p < .05$.

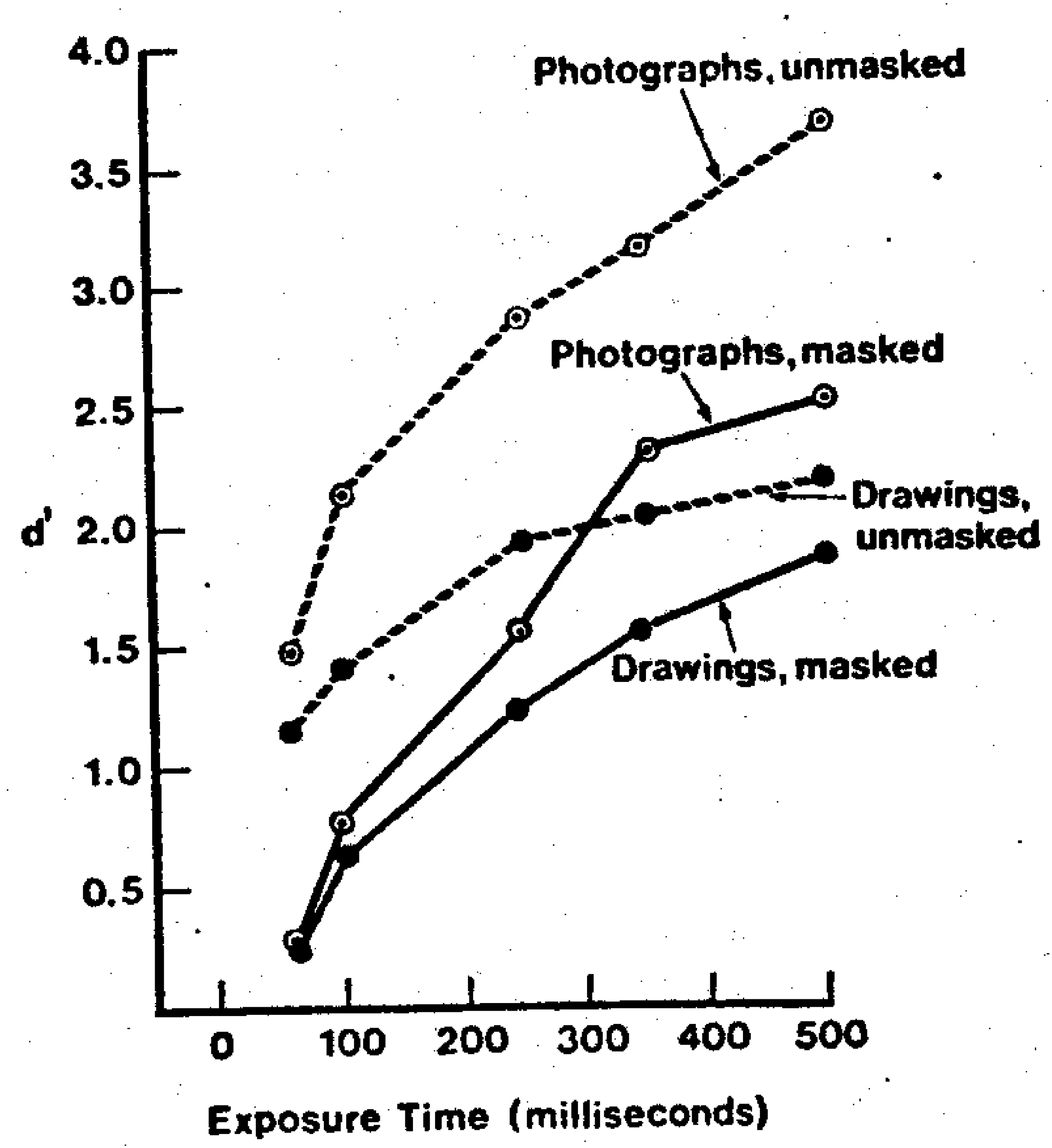

FIGURE 1. Recognition memory performance (d') as a function of exposure time. (Masked and unmasked pictures are represented by the solid and broken lines, respectively, and the curve parameter is stimulus type.)

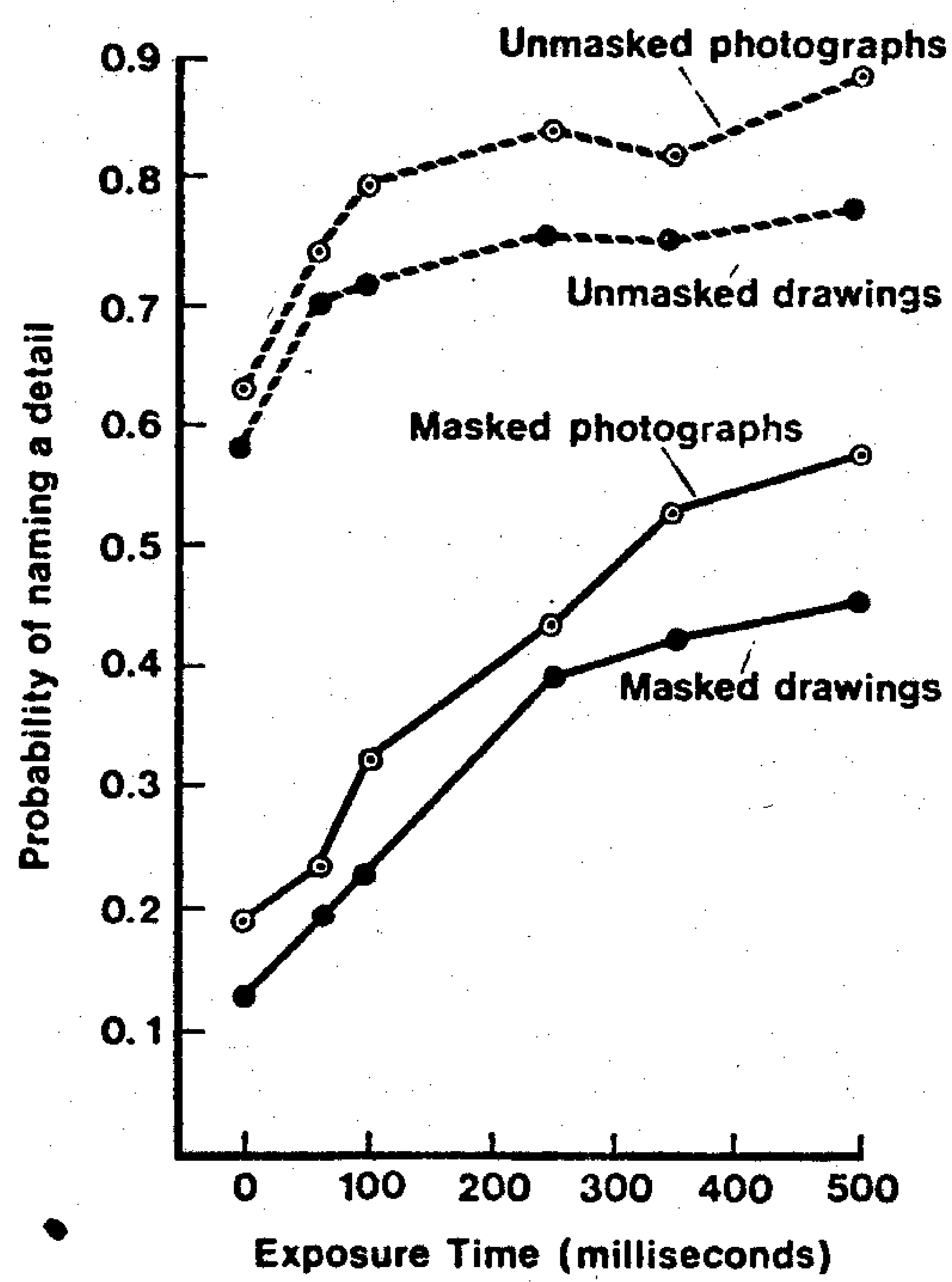

FIGURE 2. Probability of naming a detail as a function of exposure time. (Masked and unmasked pictures are represented by solid and broken lines, respectively, and the curve parameter is stimulus type. The zero-interval points represent the probability of naming a detail to a distractor picture.)

What Are Responses Based On?

Figure 2 is analogous to Figure 1, showing the probability of naming a detatil from a picture as a function of the picture's exposure time. Again, the curve parameter is stimulus type and the dashed and solid lines represent data from unmasked and masked pictures, respectively. For all four curves, the zero-exposure time data point represents the probability of naming a detail from a distractor picture and may be thought of as a bias factor. Let this bias probability be designated $p(B)$.

To interpret these data, we propose the following model of processing at the time of original study. Let a given exposure time t be broken up into a series of units,

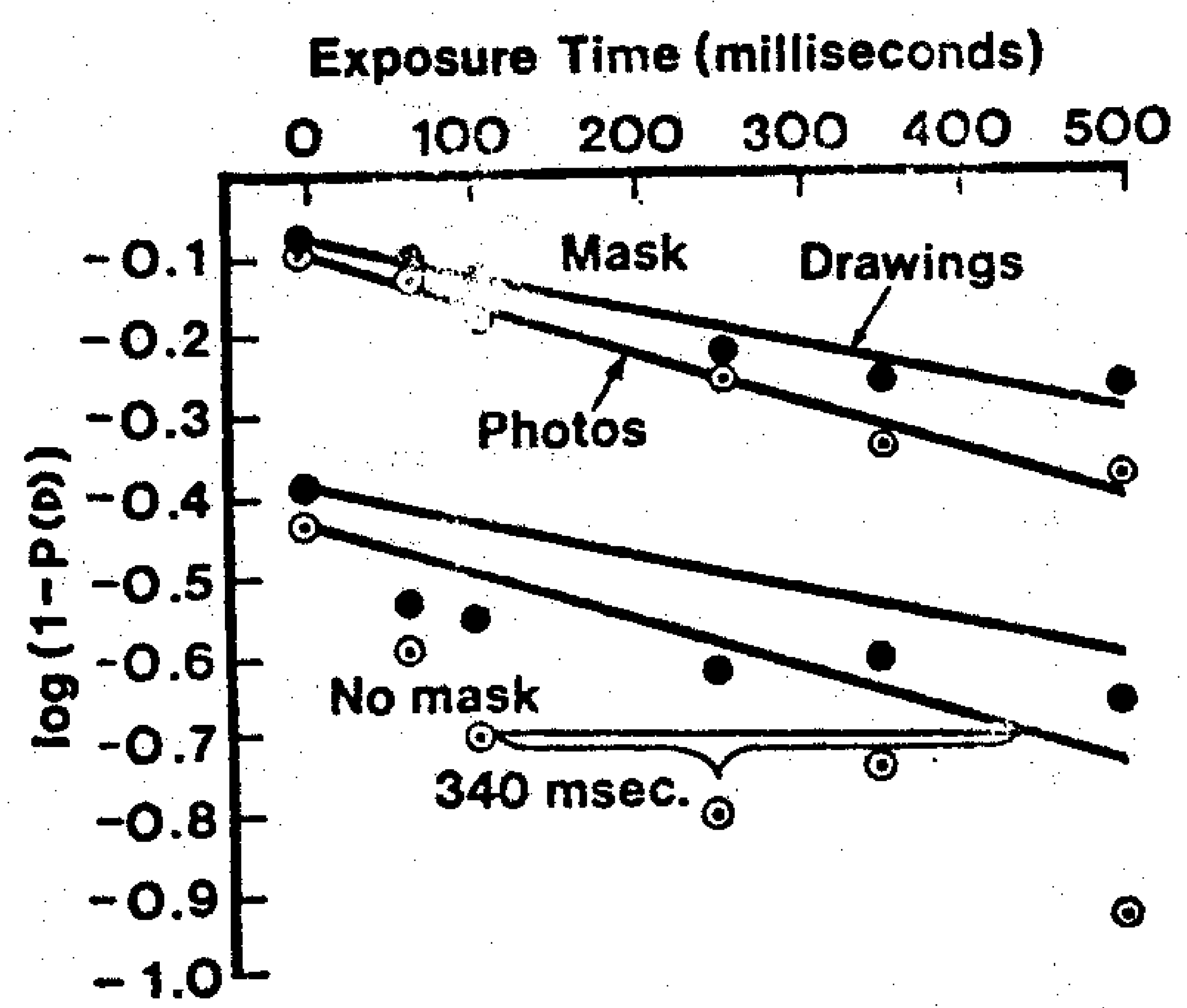

FIGURE 3. Log $[1 - p(\mathrm{D})]$ as a function of exposure time. (The upper and lower data points are from masked and unmasked pictures, respectively, and the curve parameter is stimulus type. The zero-interval data points represent log $[1 - p(\mathrm{B})]$.)

each of length Δt, and assume that during each Δt, there is some probability α that the presence of an informative detail is encoded. If we let $N = t/\Delta t$, then $p(\mathrm{E})$, the probability that at least one detail has been encoded by time t, is

$$p(\mathrm{E}) = 1 - (1 - \alpha)^N.$$

Let the probability that a detail from a target picture is *named at test* be designated $p(\mathrm{D})$. This probability is then equal to $p(\mathrm{E})$, the probability that a detail was encoded, plus the probability that a detail was *not* encoded times the bias probability, $p(\mathrm{B})$. Thus,

$$p(\mathrm{D}) = p(\mathrm{E}) + [1 - p(\mathrm{E})]p(\mathrm{B})$$
$$= 1 - (1 - \alpha)^N + (1 - \alpha)^N p(\mathrm{B}).$$

Rearranging terms and taking the logarithm of both sides of the equation,

$$\log[1 - p(\mathrm{D})] = N \log(1 - \alpha) + \log[1 - p(\mathrm{B})]$$

or

$$\log[1 - p(\mathrm{D})] = \left[\frac{\log(1 - \alpha)}{\Delta t}\right]t + \log[1 - p(\mathrm{B})].$$

That is, log $[1 - p(\mathrm{D})]$ is predicted to be a linear function of exposure time with a slope equal to $[\log(1 - \alpha)]/\Delta t$ and an intercept equal to $\log[1 - p(\mathrm{B})]$. The slope of the function may thus be viewed as a *performance* parameter whereas the intercept is a *bias* parameter.

It may be postulated a priori that even if data from the masked pictures were to show a good fit to this model, data from the unmasked pictures would not because the onset of the mask should effectively cut off processing immediately at the offset of the picture (Sperling, 1963). For example, if a picture were exposed for 250 msec, the subject would have *only* those 250 msec to process the picture and potentially encode a detail. For unmasked pictures, on the other hand, it may be assumed that subjects will have additional time to process information out of iconic store (Neisser, 1967; Sperling, 1960) following the offset of the picture. Thus, if a picture were presented for 250 msec, a subject would actually have somewhat *more* than 250 msec to process information from the picture.

Figure 3 shows $\log[1 - p(\mathrm{D})]$ as a function of t with $\log[1 - p(\mathrm{B})]$ plotted at exposure time zero. The upper data points are from the masked pictures; the lower data points are from the unmasked pictures. Concordant with the above predictions, data from masked pictures yield a fairly linear curve. The linear components account for 95% and 97% of the variance for photographs and drawings, respectively. In neither case is the residual variance significant, $F(4, 36) = 1.93$ and $F(4, 36) = 1.55$. Conversely, data from the unmasked pictures depart markedly from linearity. The linear components account for 79% and 68% of the variance for photographs and drawings, respectively. In both cases, departures from linearity are significant, $F(4, 32) = 2.93$, $p < .05$, and $F(4, 32) = 3.02$, $p < .05$, respectively.

The best fitting regression lines have been drawn through the data points of the masked photographs and drawings (top of Figure 3), producing slopes of $-.000592$ and $-.000434$, respectively. Estimation of α, the detail encoding probability, naturally depends on the value chosen for Δt. An ecologically natural Δt would be the duration of an eye fixation—about 300 msec. With this choice of Δt, the value of α is estimated to be .34 for photographs and .26 for drawings.

Regarding the unmasked pictures, it is not clear whether the amount of extra processing that can be done on the iconic image of a picture is independent of the exposure time of the picture. An independence model would predict $\log[1 - p(D)]$ to be a linear function of t with an intercept somewhat lower than $\log[1 - p(B)]$. If, however, iconic processing capability depends on exposure time, then the relationship between $\log[1 - p(D)]$ and t is predicted to be nonlinear. The data on the independence question are somewhat equivocal (cf. Haber, 1970). Hence, in the absence of a clear-cut model we deal with the problem empirically in the following way. Assuming that processing is the same for the masked and unmasked pictures *during the time the picture is physically present*, we argue that the slopes estimated for the masked pictures represent the way the unmasked pictures "should behave" were it not for the additional processing performed on the iconic image. Since function intercepts are determined solely by response bias, and since bias and encoding probability should be independent, the top curves from Figure 3 have been moved down and (preserving the slopes) have been redrawn through the bias (intercept) points of the data from the unmasked pictures. Now, by comparing these curves with the actual data points, it is possible to estimate the amount of processing done in iconic store at the different exposure times. As demonstrated in Figure 3, for example, an unmasked photograph displayed for 100 msec shows performance equivalent to a masked photograph displayed for 440 msec. Thus there is estimated to be 340 msec of what we term "equivalent processing time" done on the iconic image.

Table 1 shows equivalent iconic processing times from iconic store for pictures and drawings at the five exposure times. An analysis of variance on the data revealed that significantly more information from iconic store was processed from photographs vs. drawings, $F(1, 8) = 5.94$, $p < .05$, there was no effect of exposure time ($F < 1$), and the Stimulus Type × Exposure Time interaction was significant, $F(4, 32) = 6.14$, $p < .05$. The estimates of equivalent processing time shown in Table 1 are roughly equal to the iconic store duration times estimated by others (Averbach & Coriell, 1961; Haber, 1970; Sperling, 1960), although note that the present estimates of equivalent processing time should be underestimates of iconic store *duration*, since as the iconic image fades, progessively less information can be extracted from it. The fact that equivalent processing time is greater for photographs than for drawings indicates that the rate of iconic processing is greater for photographs than for drawings, suggesting that iconic processing may involve the same mechanisms as processing from the actual physical stimulus. The lack of an exposure time effect is, at first glance, consistent with the notion that iconic processing time is independent of exposure time. However, the interaction has no ready interpretation and dictates a good deal of caution in making any conclusions from these data. The problem could lie either in sampling error (data from the unmasked pictures are somewhat noisy, possibly due to lack of complete counterbalancing in the experiment) or in the assumption that processing during the physical exposure time is identical for the two experiments.

TABLE 1

EQUIVALENT ICONIC STORE PROCESSING TIME FOR UNMASKED PHOTOGRAPHS AND DRAWINGS

Exposure time	Photographs	Drawings
60	196	286
100	340	28C
250	361	280
350	163	157
500	372	131

Note. All times are given in milliseconds.

Performance Based on Specific Detail vs. General Visual Information

Figure 4 shows d' as a function of exposure time for masked pictures (solid curves) and unmasked pictures (dashed curves). Here, the curves are double conditionalized—on photographs vs. drawings and

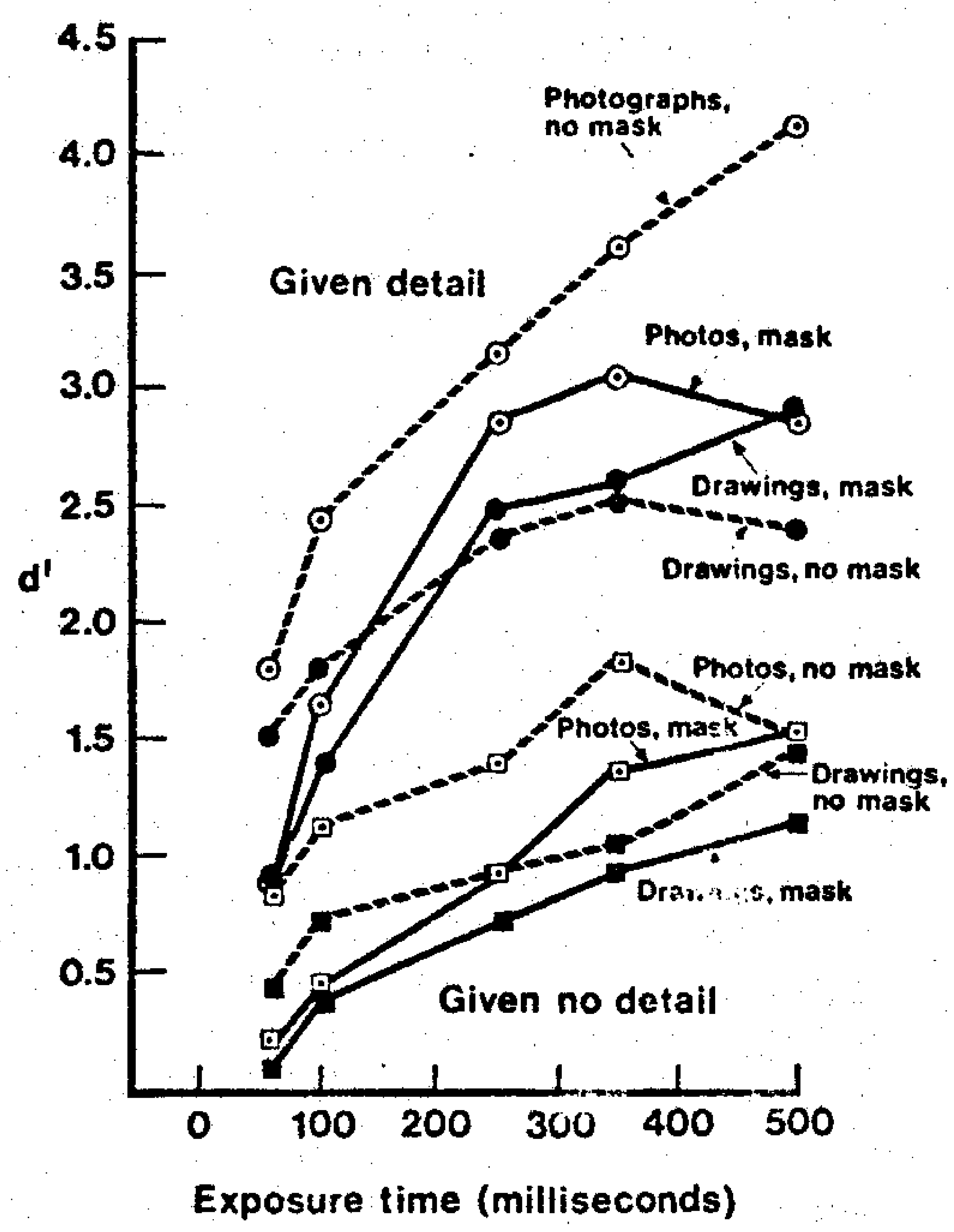

FIGURE 4. Memory performance (d') as a function of exposure time. (The solid and broken lines represent data from masked and unmasked pictures, respectively. The curve parameters are stimulus type and detail/no detail.)

also on whether or not a detail was named.[8] The data from both experiments were analyzed using a 2 × 2 × 5 within-unit analysis of variance. The results are summarized in Table 2. Basically, in each experiment the three main effects are highly significant, but none of the interactions reach significance. When a detail is reported, performance is considerably better than when a detail is not reported. Photographs are recognized better than line drawings, and performance increases monotonically with increasing exposure time. Masked pictures do not differ significantly from unmasked pictures ($t < 1$), which suggests that the effect of the mask on performance (cf. Figure 1) is mediated in large part by its effect on the encoding of specific details (cf. Figure 2).

[8] An implicit assumption in the signal detection framework is that a change in d' represents a change in the mean of the target distribution. However, two of the factors in the present analysis—stimulus type and detail vs. no detail—have separate distractor distributions for the different levels of the variables. A comparison of the false alarm rates for photographs vs. drawings revealed that the two distractor distributions were virtually identical in terms of both mean and variance. A potential problem with the detail vs. no detail factor arose in that the pattern of false-alarm rates over the confidence ratings was different for the two levels of the factor. Such a shift could be due to one of three things: a criterion difference, a difference in the means and/or variances of the two distractor distributions, or some combination of the two. The analyses in this paper are based on the arbitrary assumption of a criterion difference only. However, the analysis was also done making the opposite assumption—that of a distribution difference only. Fortunately, this analysis lead to virtually the same conclusion, both qualitatively and quantitatively, as the previous one.

DISCUSSION

As noted in the introduction to this report, we assume that recognition responses may be made on the basis of one of two types of information: specific detail information or general visual information. Within the framework of signal detection theory, this notion translates into the following: At the time a target picture is initially viewed, general visual information accrues continuously over time, i.e., the familiarity of the target continuously increases. Simultaneously, the subject is engaged in a search for a potentially informative detail and, as discussed above, with each eye fixation there is some constant probability that he will find one. Finding and encoding an informative detail is equivalent to increasing the familiarity value of the picture by a 'quantum jump."

Why are Photographs Remembered Better than Drawings?

The first reason that photographs show better performance than drawings is simply

TABLE 2

ANALYSIS OF VARIANCE SUMMARY FOR FIGURE 4

Source	Experiment 1			Experiment 2		
	df	*F*	*MS*e	*df*	*F*	*MS*e
Exposure time (E)	4, 32	7.62*	1.78	4, 36	11.99*	1.97
Stimulus type (T)	1, 8	9.39*	2.72	1, 9	5.58*	1.20
Object (O)	1, 8	53.2*	2.32	1, 9	74.5*	1.96
E × T	4, 32	1.53	1.44	4, 36	.71	1.36
E × O	4, 32	2.26	.84	4, 36	1.91	1.96
T × O	1, 8	3.69	2.06	1, 9	1.05	1.79
E × T × O	4, 32	.98	.87	4, 36	.53	1.36

* $p < .05$.

that the probability of finding an informative detail is higher for photographs. As noted above, it was estimated that the detail encoding probability was .34 per eye fixation for photographs, but only .26 per eye fixation for drawings. Additionally, when pictures are not followed by a mask, rate of processing from iconic store may continue to be higher for photographs than for drawings. It is evidently the case that the more informative details there are in a picture, the higher is the probability of encoding a detail during a given unit of time.

There appears to be a second reason why performance is superior on photographs than on drawings. The bottom curves of Figure 4 indicate that when one considers responses made only on the basis of visual information (i.e., when no detail is named), photographs still show better performance than drawings. This finding—that photographs are inherently more distinguishable from each other than drawings—was unexpected, and at present we do not offer a theoretical interpretation of it.

The present data are somewhat difficult to reconcile with those of Nelson et al. (1974). In the Nelson et al. study, the same stimuli used in the present experiments were presented for 10 sec apiece in a study phase, and then were given a two forced-choice recognition test, either immediately or after a 7-wk delay. On both tests, recognition performance was not significantly different for photographs vs. drawings. In the immediate test, response probability was about .98; hence the lack of performance difference is easily attributable to a ceiling effect. However, response probability on the delayed test was about 87% for both photographs and drawings. In terms of the present formulation, it is probably the case that during a 10-sec presentation, a subject can always extract sufficient information from a picture to distinguish it from other pictures; however, the Nelson et al. results indicate that the rate of forgetting from a picture does not depend on how many informative details there were in the picture to begin with.

Independence of Codes

Consider now the top curves in Figure 4, which show performance when a detail is named. The *d'* values of these curves are assumed to be the sum of familiarity stemming from general visual information plus familiarity stemming from encoding a detail. To isolate the specific detail component of these curves, it is necessary to subtract the general visual components; essentially, the bottom curves of Figure 4 must be subtracted from the top ones. The fact that the detail vs. no-detail variable does not statistically interact with the other variables in either experiment (cf. Table 2) indicates that the amount of detail information does not significantly depend on either picture type or exposure time. In all cases, encoding a detail boosts *d'* by approximately 1.5.

Specific Details: Verbal Encoding?

A substantial amount of work in the field of visual imagery has resulted in the conclusion that two cognitive processing systems exist: a visual system and a verbal–auditory system that operate independently of each other (Atwood, 1971; Bower, 1970; Brooks, 1968; Paivio, 1969; Seamon & Gazzaniga, 1974; Segal & Fusella, Note 2). This idea is complimented by the results of picture recognition experiments that have attempted to manipulate the degree to which verbal information in pictures may be encoded. Two studies (Freund, 1971; Kurtz & Hovland, 1953) showed that if a subject verbally describes a scene being viewed, subsequent recognition performance is boosted relative to a control condition of normal viewing. The converse finding is that if verbal encoding is *prevented* during viewing (by forcing a subject to count backward by threes, which presumably uses up the capacity of the verbal system), subsequent memory performance is reduced although not to chance (Freund, 1971; Loftus, 1972; Szewczuk, 1970). Finally, many studies have demonstrated that a verbal label given to a nonsense form will increase recognition memory performance for the form (Bostrum, 1971; Clark, 1965; Ellis, 1968). The conclusions of these studies have been that recognition performance for visual material is based on two types of information. A verbal component is inferred because verbalization manipulations substantially affect performance. A nonverbal (or visual) component is inferred because performance is not reduced to chance when verbalization is prevented.

It may be that what we have termed specific detail information is equivalent to what others have described as the verbal component of picture memory, that is, the process of seeking out and encoding specific details may by undertaken by the verbal–auditory cognitive system. At present, however, this notion remains at a speculative level because the only real evidence supporting a verbal–nonverbal distinction in the present study comes from subjects' introspective reports. Subjects claim that the process of encoding information about a detail essentially involves generating a verbal label for that detail. Thus, when a recognition response is based on the detail, what is being remembered is the verbal label. Recognition responses *not* based on details are by definition nonverbal because subjects are unable to verbally describe what they are basing their responses on beyond simply saying that the picture "looks familiar."

REFERENCES NOTES

1. Egan, J. P. *Recognition memory and the operating characteristic* (Tech. Note AFCRC-TN-58-51). Bloomington, Ind.: Indiana University, Hearing and Communication Laboratory, 1968.
2. Segal, S. J., & Fusella, V. *Imagining and perceiving in two separate modalities.* Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, April 1969.

REFERENCES

Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review,* 1972, *79,* 97–123.

Atwood, G. An experimental study of visual imagery and memory. *Cognitive Psychology,* 1971, *2,* 290–299.

Averbach, E., & Coriell, A. S. Short-term memory in vision. *Bell Systems Technical Journal,* 1961, *40,* 309–328.

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.

Bostrum, A. *Memory for form: The effects of memory for verbal labels and label quality.* Unpublished doctoral dissertation, Stanford University, 1971.

Bower, G. H. Analysis of a mnemonic device. *American Scientist,* 1970, *58,* 496–510.

Brooks, L. R. Spatial and verbal components of the act of recall. *Canadian Journal of Psychology,* 1968, *22,* 349–368.

Buswell, G. T. *How people look at pictures.* 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.

Ellis, H. Transfer of stimulus predifferentation to shape recognition and association learning: Role of properties of verbal labels. *Journal of Experimental Psychology,* 1968, *78,* 401–409.

Freund, R. D. *Verbal and nonverbal processes in picture recognition.* Unpublished doctoral dissertation, Stanford University, 1971.

Haber, R. N. How we remember what we see. *Scientific American,* 1970, *222,* 104–112.

Kintsch, W. An experimental analysis of single-stimulus tests and multiple-choice tests of recognition memory. *Journal of Experimental Psychology,* 1969, *76,* 1–6.

Kurtz, K., & Hovlanc, 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. Eye fixations and recognition memory for pictures. *Cognitive Psychology,* 1972, *3,* 525-551.

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

Neisser, U. *Cognitive psychology,* New York: Appleton-Century-Crofts, 1967.

Nelson, T. O., Metzler, J., & Reed, D. A. Role of details in the long-term recognition of pictures and verbal descriptions. *Journal of Experimental Psychology,* 1974, *102,* 184–186.

Paivio, A. Mental imagery in associative learning and memory. *Psychological Review,* 1969, *76,* 241–263.

Seamon, J. G., & Gazzaniga, M. S. Coding strategies and laterality effects. *Cognitive Psychology,* 1973, *5,* 249-256.

Sperling, G. The information available in a brief visual presentation. *Psychological Monographs,* 1960, *74,* 1–29.

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

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

Appendix

The d' scores were estimated as follows. Assume that the variance of the distractor distribution for a given condition is set at unity, and denote the standard deviation of the target distribution by σ_T. For a given criterion level, denote the z scores corresponding to the hit and false-alarm rates by $z(H)$ and $z(FA)$, respectively. It then can be easily shown that, where d' is expressed in units corresponding to the standard deviation of the distractor distribution,

$$d' = \sigma_T \, z(H) - z(FA) \qquad (A1)$$

or,

$$z(H) = d'/\sigma_T + z(FA)/\sigma_T. \qquad (A2)$$

Equation A2 is the function corresponding to a memory operating characteristics (MOC) curve, expressed in z scores. For a given condition, an MOC curve was obtained using the confidence ratings from that condition and the best fitting straight line through the MOC points was computed. The d' score for the condition was then estimated by dividing the intercept of this best fitting straight line by its slope.

Two aspects of the d' computations should be mentioned. First, the MOC curves were fit very well by straight lines—for the 40 MOC curves corresponding to the 40 data points of Figure 4, the median r^2 was .98. Second, the variance of the target distributions was invariably greater than the variance of the distractor distributions.

(Received July 26, 1974; revision received September 19, 1974)