

Encoding and Use of Detail Information in Picture Recognition

Geoffrey R. Loftus
University of Washington

Howard J. Kallman
University of Wisconsin

Subjects participated in a yes/no picture recognition experiment in which exposure time varied from 50 to 1,000 msec at the time of initial study. Following each study trial, half of the subjects, the detail at study and test (ST) group, reported whether they had observed a detail in the picture that they thought might help in subsequent recognition. The other half of the subjects, the detail at test only (T) group, did not attempt to name details during study. All of the subjects reported at the time of each test picture whether they were basing their yes/no recognition response on a specific detail in the picture or on the picture's general familiarity. The data provided strong support for a model which assumed that (a) there is a constant probability of encoding a detail during each successive unit of time at study and (b) a detail is named at test either if it was encoded at study or with some bias probability. ST subjects showed superior recognition memory performance relative to T subjects. Within the context of the aforementioned model, this superiority stems from two sources: ST subjects encode details at a faster rate than do T subjects and an encoded detail provided a better discriminative feature for ST subjects.

Human beings are reasonably capable of inspecting a visual scene and deciding whether they have previously encountered it. Two questions that arise from such an observation are: (a) What is the nature of the stored information that permits such a recognition decision to be made? and (b) What are the processes by which this information is initially acquired? In this article, we elaborate on and provide support

for a model of picture encoding proposed by Loftus and Bell (1975). A central claim of this model is that two types of information are acquired during the viewing of a picture, which are termed *general visual information* and *specific detail information*.

Two Types of Information in Picture Memory

The notion that two types of information are acquired from a picture is not new. The idea has been proposed by several investigators who have typically characterized the two types of information as being visual and verbal. Before describing the Loftus and Bell (1975) model, we briefly sketch the evidence that supports the establishment of such a dichotomy.

This research was supported by National Science Foundation Grant BNS76-23367 to G. Loftus.

We thank Gail Goodman, Colin MacLeod, Tom Nelson, and two anonymous reviewers for their comments on the research and the article.

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

Dual codes to enhance memory performance. Numerous experiments have examined memory either for pictures of simple objects or for the verbal labels for these objects and have found that memory performance is better for pictures (e.g., Davies, Milne, & Glennie, 1973; Jenkins, Neale, & Deno, 1967; Lieberman & Culpepper, 1965; Paivio & Csapo, 1969, 1971, 1973). The explanation typically offered for this finding is that spontaneous verbal labeling takes place when pictures are presented, whereas the converse process—generation of images—does not necessarily accompany the presentation of verbal labels. Such an explanation is supported and complemented by a variety of other kinds of results. A myriad of visual-imagery studies have demonstrated that in a verbal-memory task, generation of a visual image produces superior memory performance compared to rote repetition or other verbal activity (cf. Paivio, 1971). Experiments by Bryant (1965) and Kurtz and Hovland (1953) have explicitly manipulated whether or not verbal labels were generated during visual presentation of simple objects. Except in the case of severely retarded children, the label condition produced higher recall than did the no-label condition. In contrast, preventing verbalization during the viewing of pictures dramatically reduced recall (Rowe & Rogers, 1975). In a somewhat different experimental paradigm, Petrino-vich and Hardyck (1970) demonstrated that an instrumental-conditioning response shows more generalization from pictures to words than from words to pictures. This finding is again commensurate with the notion that the degree to which verbal labels are generated to pictures is greater than the degree to which visual codes are generated to verbal labels. The explanation common to all these phenomena incorporates the fundamental postulate that two potential codes can be generated for remembering information: one verbal and one visual.

The experiments cited above have all used simple, easy-to-name objects as stimuli. Analogous experiments have examined memory for complex pictures versus mem-

ory for descriptions of the pictures (Dallett & Wilcox, 1968; Nelson, Metzler, & Reed, 1974) and have produced similar results: Memory for pictorial information is superior to memory for a corresponding verbal description. In other experiments, verbalization of complex pictures has been either forced or prohibited. Freund (1971) compared a condition in which pictures were viewed normally with a condition in which viewing was accompanied by a concurrent verbal description of the picture. The verbal description condition produced better subsequent recognition than did the normal condition. The other side of this coin is the finding of inferior (but above chance) recognition performance following a condition in which verbalization of pictures was prevented relative to a normal-viewing condition (Freund, 1971; Loftus, 1972; Szwedczuk, 1970). Again, the explanation for this finding incorporates the idea that both a verbal and a visual code may potentially be generated for pictures. Pictures are more likely to produce a dual code than are verbal descriptions. The verbalization manipulations are assumed to affect the verbal code for a picture, which in turn affects memory performance.

Independence of codes. Several experiments have compared recognition and recall memory for the same material and have found recall and recognition performance to be independent of one another (Bahrick & Bahrick, 1971; Bahrick & Boucher, 1968; Mandler & Stein, 1974; Tversky, 1974). For instance, the Bahrick and Boucher study used a procedure in which pictures of common objects (e.g., a coffee cup) were presented followed by both a recall test for the names of the objects and a recognition test in which each object had to be selected from an array of objects all bearing the same name but differing in subtle visual ways (e.g., the coffee cup had to be selected from an array of cups differing with respect to the shape of the handle or to the markings on the cup). Presumably, performance on the recall test was based on a verbal code, whereas performance on the recognition test was based on a visual code. Recall and recognition

performance were uncorrelated, in keeping with the hypothesis that recall and recognition are based on two types of information that are stored independently of one another.

Other experiments have demonstrated independence of codes in other ways. Dallett, Wilcox, and d'Andrea (1968) showed subjects complex pictures to which verbal labels had been attached. Presentation or recall of the picture label had no effect on recognition of the picture itself. Paivio and Csapo (1973) reported a series of experiments in which either a picture of a common object and/or its corresponding label were presented. When both a picture and its label were shown, the resulting recall performance was predicted well by a mathematical model, which assumed that the label and the picture provided independent codes for the same concept.

Independence of coding systems. The results discussed above have supported the hypothesis that two types of memorial codes exist. This hypothesis is complemented by several results that suggest the existence of two analogous, independent coding systems in the human (Atwood, 1971; Brooks, 1967, 1968). For example, the Brooks studies have demonstrated that cognitive activity requiring the use of one presumed coding system is interfered with by activity requiring the use of the same system but is not interfered with by activity requiring the use of the other system. Thus, if a subject is visualizing a block letter, the visual activity of pointing to one of two places to indicate various characteristics of the letter is more difficult than the verbal activity of voicing the characteristics.

A provocative start at identifying the physiological locus of such coding systems has been reported by Seamon and Gazzaniga (1973). The paradigm used in Seamon and Gazzaniga's experiment was a memory-scanning task (Sternberg, 1966) in which words were used as stimuli and subjects were instructed to use either a verbal (rote repetition) or a visual (imagery) strategy to maintain the memory set. A test stimulus was then presented either to the subject's left or right hemisphere. The results indi-

cated a strong Strategy \times Hemisphere interaction in terms of reaction time: Responses were faster to left-hemisphere probes when using a verbal strategy but faster to right-hemisphere probes when using a visual strategy. This result provides strong support for the claims of others (e.g., Sperry, 1961, 1968) that the left hemisphere is primarily responsible for verbal processing, whereas the right hemisphere is primarily responsible for spatial (visual) processing.

Probing the Bases of Picture-Recognition Responses

During the test phase of one picture-recognition experiment, Loftus (1972, Experiment 1) asked observers to determine whether recognition responses were being made on the basis of memory for specific details (e.g., "I recognize this picture because I remember that kangaroo in the lower left") or simply on the basis of the picture's general familiarity.¹ Instances of recognition responses that had been made on the basis of a specific remembered object were examined with respect to the original pattern of eye fixations on the picture at the time of study. Figure 1 shows the results of this analysis: First, the remembered object was discovered quickly during the course of viewing—almost invariably within the first few fixations on the picture. Second, the unconditional probability of fixating the remembered object during the course of viewing the picture was very high—above 50% on any given fixation.

¹ The rationale for offering these choices was twofold. First, it was thought that these two responses might correspond to recognition responses based on verbal versus visual information. Second, these two types of responses have intuitive appeal. Introspection and informal discussion suggest that some scenes in everyday life are recognized on the basis of specific details (e.g., "This is the way back down the mountain; there's that unusual tree stump"), whereas other responses are made on the basis of more general, difficult-to-verbalize information (e.g., "I know I've been here before; there's a haunting familiarity," or "I recognize the face, but I can't remember the name").

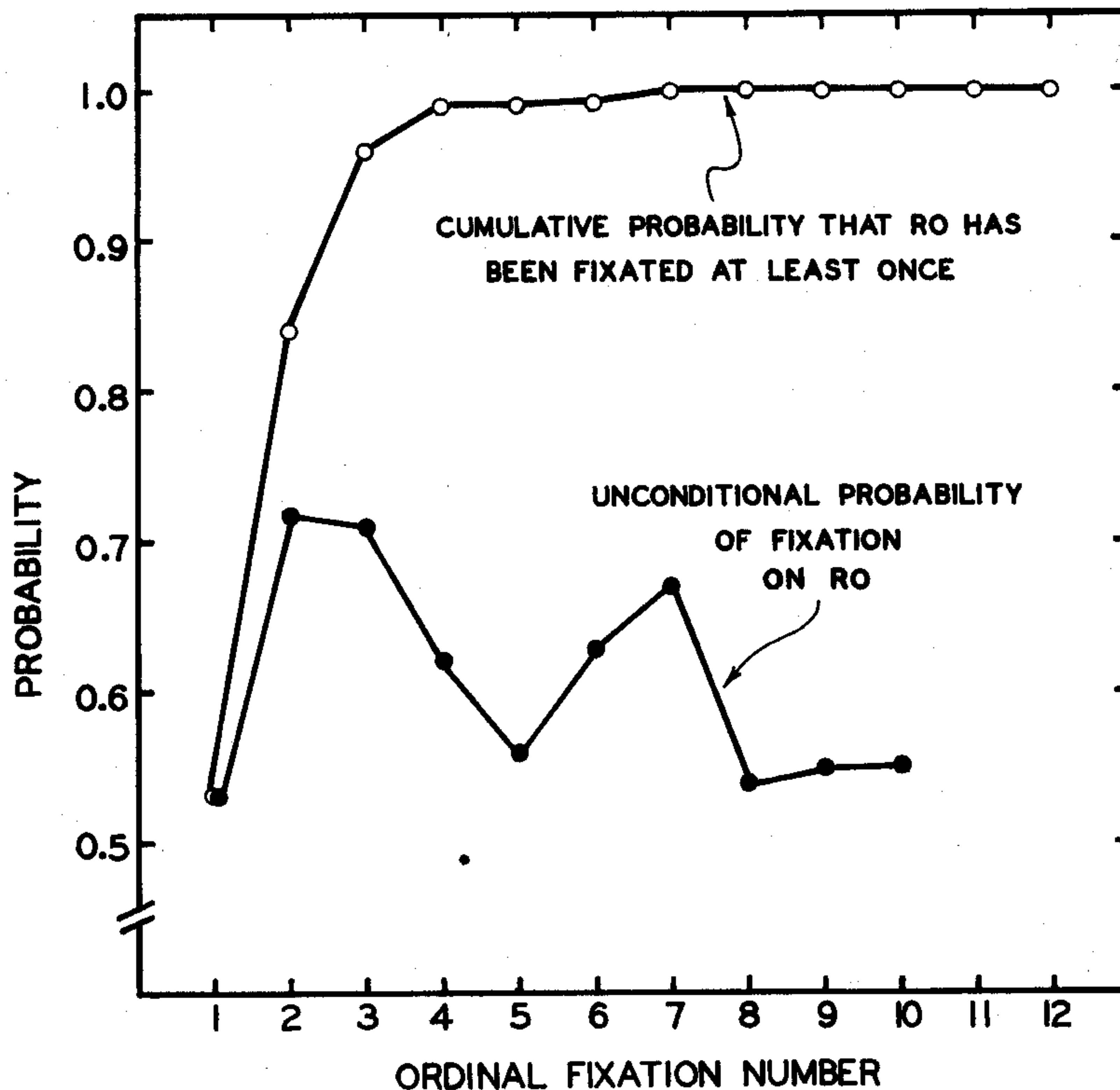


Figure 1. Results from Loftus (1972). (RO is "remembered object"—the object on which a recognition response was ultimately based.)

The data shown in Figure 1 indicate that a reasonable amount of empirical validity results from the technique of asking observers to identify the bases for their recognition responses. A somewhat more systematic exploration of this technique was reported by Loftus and Bell (1975). Their experiment also used a picture-recognition paradigm. The stimuli consisted either of photographs or, for each photograph, a corresponding line drawing constructed so as to capture only the central information in the photograph (cf. Nelson et al., 1974). To investigate the time course of information acquisition, exposure time at study was varied from 60 to 500 msec. A visual noise mask followed the offset of each study picture. At the time of test, subjects were asked to report whether or not their responses were based on a specific detail.

Loftus and Bell (1975) found that when a detail was named at test, recognition-memory performance (d') was considerably higher than when a detail was not named at test. They did not commit themselves to

the claim that picture memory is based on verbal and visual information but instead coined the more neutral terms *specific detail* information and *general visual* information. Based on their data, Loftus and Bell argued that information about specific details was considerably more valuable in terms of subsequent picture memory than was general visual information.

A Model of Detail Encoding

Inasmuch as detail information appeared to be fairly important in terms of recognizing pictures, Loftus and Bell (1975) formulated a specific model for acquisition of detail information. We now describe this model.

Suppose that a to-be-recognized picture is viewed at study for some exposure time (t). Assume that t can be broken up into a series of units, each of length Δt . During each Δt , there is some probability, α , that the viewer encodes the presence of an informative detail (i.e., a detail that will

ultimately allow recognition of the picture). Letting N be the number of Δt s (i.e., $N = t/\Delta t$), then $p(\epsilon)$, the probability that at least one informative detail has been encoded by time t , is

$$p(\epsilon) = 1 - (1 - \alpha)^N. \quad (1)$$

Rearranging terms,

$$1 - p(\epsilon) = (1 - \alpha)^N.$$

Taking the logarithm of both sides of the equation,

$$\log [1 - p(\epsilon)] = N \log (1 - \alpha),$$

And, finally, since $N = t/\Delta t$,

$$\log [1 - p(\epsilon)] = t \frac{\log (1 - \alpha)}{\Delta t}. \quad (2)$$

The model is thus seen to make a strong prediction: The logarithm of $[1 - p(\epsilon)]$ should be a linear function of exposure time with an intercept of zero and a slope that reflects the detail-encoding rate.

The prediction of Equation 2 could not be tested in the Loftus and Bell (1975) study because $p(\epsilon)$ was not observable. What was observed, however, was the probability of naming a detail at test, a probability that we designate as $p(\theta)$. To test the model, Loftus and Bell made the additional assumption that $p(\theta)$ was equal to the sum of $p(\epsilon)$, the probability that a detail was encoded, and $[1 - p(\epsilon)]p(\beta)$, the probability that a detail was not encoded multiplied by a bias probability, $p(\beta)$. Thus,

$$p(\theta) = p(\epsilon) + [1 - p(\epsilon)]p(\beta), \quad (3)$$

or substituting Equation 1 into Equation 3,

$$p(\theta) = 1 - (1 - \alpha)^N + (1 - \alpha)^N p(\beta). \quad (4)$$

Rearranging terms and taking logarithms of both sides of Equation 4,

$$\begin{aligned} \log [1 - p(\theta)] \\ = N \log (1 - \alpha) + \log [1 - p(\beta)]. \end{aligned}$$

Again, substituting in the fact that $N = t/\Delta t$,

$$\begin{aligned} \log [1 - p(\theta)] \\ = t \frac{\log (1 - \alpha)}{\Delta t} + \log [1 - p(\beta)]. \end{aligned} \quad (5)$$

Thus, according to Equation 5, $\log [1 - p(\theta)]$ is a linear function of exposure time with

a slope equal to $\log (1 - \alpha)/\Delta t$ and an intercept equal to $\log [1 - p(\beta)]$. The slope of the function may be viewed as a *performance* parameter and the intercept as a *bias* parameter.

Both $p(\theta)$ and $p(\beta)$ were estimated from the Loftus and Bell (1975) data; the two estimators were referred to as $p(D)$ and $p(B)$, respectively. For any given exposure time, $p(D)$ was defined as the proportion of pictures studied at that exposure time for which a detail was named at test. Likewise, $p(B)$ was defined as the probability that a detail was named at test to a distractor picture. As it turned out, the Loftus and Bell data fit the model quite well. The best-fit straight lines accounted for 95% and 97% of the variance for the photographs and the drawings, respectively. The slope (the encoding parameter) differed for drawings and photographs, indicating that within the context of the model, encoding of details occurred at a faster rate for photographs than for drawings—just as one would expect, since the density of details was greater in photographs than in drawings.

To the degree that the Loftus and Bell (1975) model is valid, it constitutes a useful methodological tool for examining various aspects of picture recognition. This is because its major prediction—Equation 5—is a simple linear equation whose slope and intercepts are meaningful entities. However, the degree of support that the model has received from the Loftus and Bell data is somewhat sparse.

The present experiment had three purposes. First, the experiment was designed to make an important component of the model— $p(\epsilon)$ —definable using observed data. This permitted a test of the Equation 2 prediction as well as a test of the predicted relationships among $p(\epsilon)$, $p(\theta)$, and $p(\beta)$. Second, the experiment included a variable that one would expect a priori to have an effect on α , the encoding rate, given the validity of the model. Finally, aspects of the Loftus and Bell data were replicated with a greater range of exposure times, a different set of slides, and a large number of observations for

each subject. The former two goals were accomplished by including a group of subjects who attempted to name details at the time of *study* as well as at the time of test, in addition to a group that attempted to name details at the time of test only. The probability of naming a detail at the time of study provides an index of $p(\epsilon)$; furthermore, one would expect subjects who must name details at the time of study to encode details at a faster rate than subjects who do not have to name details at the time of study. This expectation may be assessed by comparing the detail-encoding rates (α_s) for the two groups.

Method

Stimuli

The stimuli were 640 color pictures of complex naturalistic scenes—mountains, sunsets, landscapes, and so forth. Pictures containing people were generally avoided, as were pictures involving any signs or other verbal material. The pictures were prepared as 640, 35-mm slides that were randomly divided into two sets (Set A and Set B) of 320 slides per set.

Apparatus

Slides were shown using a Kodak random-access projector equipped with a Gerbrands tachistoscopic shutter. Each slide was followed immediately on its offset by a 1-sec, random-noise masking slide projected from a standard Carousel projector, also equipped with a Gerbrands tachistoscopic shutter. The relationship of the mask and masked slide was such that when the two were superimposed, no information could be acquired from the masked slide. The rise and fall times of the tachistoscopic shutters were approximately 1 msec.

Subjects

Subjects were 20 University of Washington undergraduates who responded to a newspaper advertisement. Each subject was paid a total of \$10 for participating in two experimental sessions of approximately 2 hr. each. None had participated in any picture-memory experiments before.

Design

Subjects were randomly divided into two groups of 10 subjects each. These groups were designated the *detail at test only* (T) group and the *detail at study and test* (ST) group. The ST subjects attempted

to name details at study, whereas the T subjects did not, a difference in procedure that we will describe in more detail below.

A given subject participated in two experimental sessions. In the initial (study) portion of each session, the subject saw 160 target slides. Immediately thereafter, the 160 targets were randomly intermingled with 160 distractors, and the resulting 320 test slides were shown one by one in a yes/no recognition test.

Target slides were shown for one of five exposure times—either 50, 100, 250, 500, or 1,000 msec. For a given subject, order of the five exposure times was random, with the restriction that an equal number of study trials fell into each exposure-time condition. Thus, prior to viewing a given slide, the subject did not know how long that slide would be shown.

Counterbalancing. For ease of discourse, imagine the 10 T subjects to be numbered 1–10 and the 10 ST subjects to be numbered 11–20. Consider now Subjects 1–10. Subjects 1–5 saw Set A pictures as targets and Set B pictures as distractors, whereas Subjects 6–10 saw Set B pictures as targets and Set A pictures as distractors. A given picture always appeared in the same study position and was rotated through the five different exposure times over the 5 subjects for whom that picture appeared as a target. Thus, every picture appeared equally often in all exposure-time conditions and equally often as a target and a distractor. Subjects 11–20 received exactly the same study ordering as did Subjects 1–10. One random test ordering was constructed, and this ordering was identical for all 20 subjects.

Procedure

Study procedure. Subjects were run individually. A subject was initially given full instructions about the nature of a yes/no recognition test and about the different exposure times and was then provided 30 practice target slides—6 each at the five different exposure times. Following a question period, the study session began. The following series of events constituted a study trial: The trial number was read by the experimenter. The slide then appeared, followed immediately by the mask. At this point, the difference between the T and the ST subjects was implemented. The ST subjects used the inter-trial interval (5 sec) to write down the name of any detail in the picture that they thought might subsequently help in terms of recognizing the picture. They were provided with a sheet containing a list of the trial numbers along with a space in which they could write the name of a detail if one had occurred to them. If no detail had occurred, they could leave the space blank. T subjects were also provided with a 5-sec interval following the offset of the mask, but their only task was to check off the trial number of the slide.

Test procedure. As noted, all subjects—both ST and T—were treated identically at the time of test. For each test trial, three responses were requested.

Table 1
Response-Defined Joint Probabilities for the ST Group

Test response	Test	Target study		
		Detail	No detail	Distractor
Yes	Detail	Same	a = .319	k = .120
		Different	b = .062	
No	No detail	Same	c = .076	l = .171
			Different	
	Detail	Same	e = .059	m = .367
No detail	Same	g = .064	n = .342	
				Different
No	Detail	Same	i = .094	m = .367

Note. Distributions of responses are provided separately for the targets and distractors. Data are collapsed over the five exposure times. The labels a–n are arbitrary designations.

First, the picture was classified as old or new. Second, a confidence rating from 1 (very sure) to 3 (just guessing) was solicited. The subject was asked to try to give about an equal number of yes and no responses and to use the three confidence ratings about equally often. Finally, the subject was asked to indicate whether his or her response was based on some specific detail in the picture. The subject was asked to write down either a detail that was used as the basis for a yes response (i.e., a detail that the subject claimed to have remembered from study) or a detail that was used as the basis for a no response (i.e., a detail that the subject claimed would have been remembered if it had been seen at study). If no detail was named, the space was left blank. Subjects were given as much time as they wanted to make each test response. No practice trials were provided.

Results and Discussion

In the following discussion, the claim of a significant effect will denote a *p* value of less than .05. Unless otherwise noted, the claim of a nonsignificant effect will denote a *p* value of greater than .10.

All data points to be presented (probabilities, logarithms of probabilities, and *d'* scores) were obtained by computing an individual data point for each subject and then computing the mean of these individual data points.

Response-Defined Joint Probabilities

Subjects' responses may be jointly classified both in terms of (a) whether a yes or a no response was made at test and (b)

whether an object was named or not named at test. For ST subjects who were responding to target pictures, several additional factors must be considered. First, a target picture may or may not have had a detail named from it at study. Second, if a detail was named from a target picture both at study and at test, the test detail could have been the same or different than the study detail.

In view of these complexities, it seems prudent to systematically characterize all the ensuing joint probabilities. This has been done in Table 1 (for the ST group) and Table 2 (for the T group). These data are collapsed over the five exposure times. For each of the two groups, the tables depict the distributions of target and distractor probabilities into all possible response-defined cells. Tables 1 and 2 serve three purposes. First, they are meant to clarify exactly how responses may be broken down. Second, the labeling of each cell provides a convenient device for defining various measures to be discussed below. (Note that the letters a–w are *labels only*. Some of these letters, e.g., i, j, p, q should not be confused with their often-used role as designations for other quantities.) Finally, the presentation of relevant joint probabilities will allow the reader to calculate various joint and conditional probabilities that are not explicitly presented in the text.

Table 2
Response-Defined Joint Probabilities for
the T Group

Test response	Test	Targets	Distractors
Yes	Detail	p = .304	t = .101
	No detail	q = .261	u = .192
No	Detail	r = .197	v = .349
	No detail	s = .237	w = .357

Note. Distributions of responses are provided separately for the targets and distractors. Data are collapsed over the five exposure times. The labels p-w are arbitrary designations.

Overall Recognition Performance

Figure 2 shows d' as a function of exposure time for the ST and the T groups. These d' scores were calculated from overall hit- and false-alarm rates. Specifically, they are defined as follows. For any probability, x , let $Z(x)$ be the normal deviate corresponding to x . Now consider the cells in Tables 1 and 2 for a single exposure time. For that exposure time,

ST group:

$$d' = Z(a + b + c + g + h) - Z(k + l)$$

T group: $d' = Z(p + q) - Z(t + u).$

As is readily evident from inspection of Figure 2, d' increased as a function of exposure time, $F(4, 72) = 145.42$, $MS_e = .046$. Additionally, the ST group showed higher d' scores than the T group, $F(1, 18) = 5.76$, $MS_e = .37$. Finally, there was a significant Group \times Exposure Time interaction, $F(4, 72) = 5.15$, $MS_e = .046$.

Two major questions arise from these results: (a) What sort of additional information are ST subjects acquiring at study that leads to their superior performance? and (b) What additional information are both groups acquiring that leads to superior performance with increasing exposure time?

Acquisition of Details

As noted earlier, prior research indicates that acquisition of a detail from a picture aids substantially in subsequent recognition of the picture. It seems reasonable to

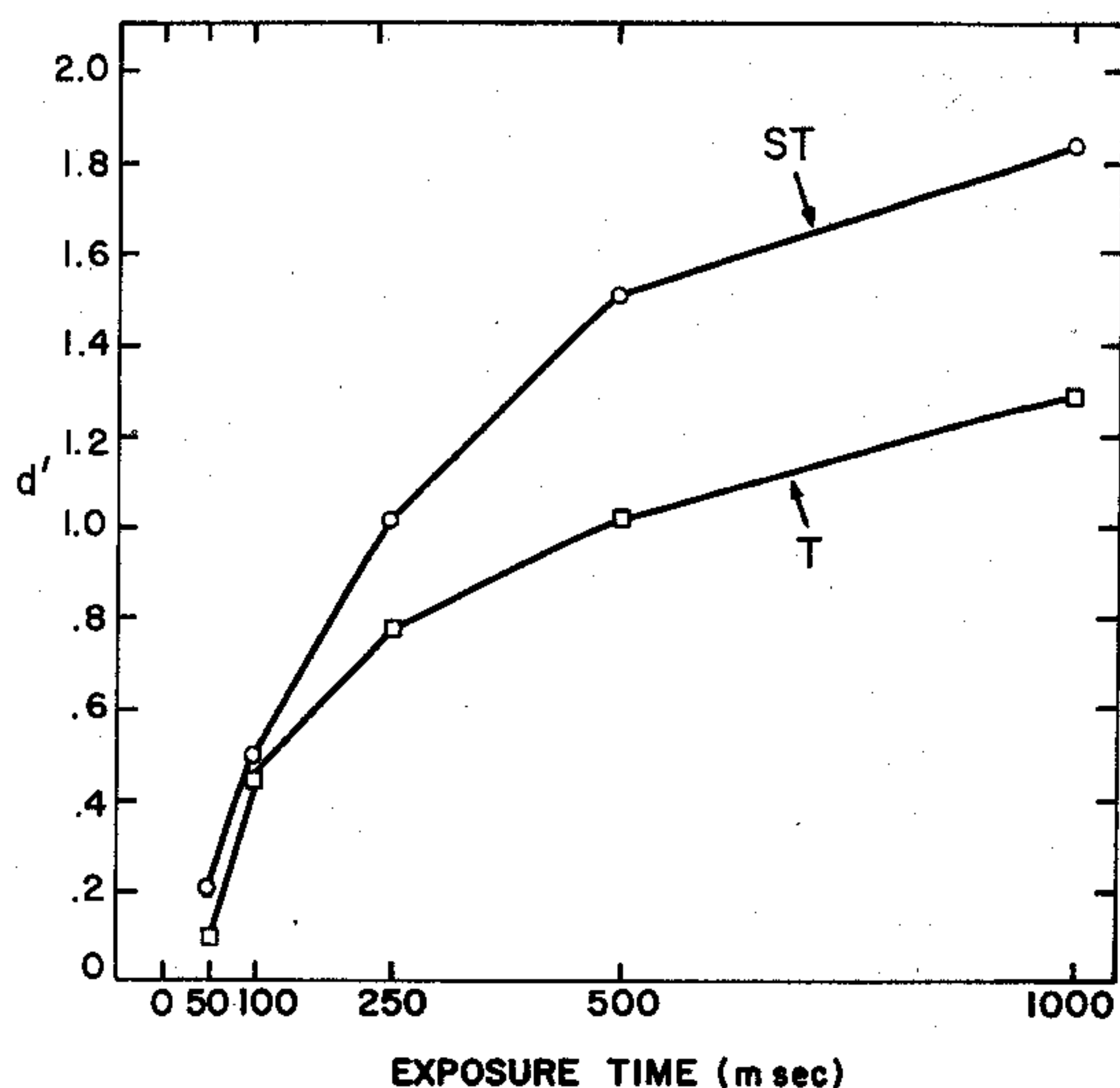


Figure 2. Recognition memory performance (d') as a function of initial exposure time for the detail at study and test (ST) and detail-only (T) groups.

suppose that ST subjects acquired detail information at a faster rate than did T subjects. This may be one reason for the superior recognition-memory performance of the ST subjects.

We compare the T group and the ST group in terms of detail acquisition rate using a two-step process. First, we present data that support the validity of the Loftus and Bell (1975) model. After establishing validity, we estimate α , the encoding rate for the ST group and the T group. The expectation of a faster detail-encoding rate for the ST group translates into the prediction that α should be higher for the ST group than for the T group.

Support for the Loftus and Bell model. To describe the fit of our data to the model, it is convenient to again present the two major equations of the model. First, in terms of the detail-encoding probability,

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

Second, in terms of detail-naming probability,

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

Several fairly strong predictions are evident from an inspection of these two equations. First, $\log [1 - p(\epsilon)]$ and $\log [1 - p(\theta)]$ should both be linear functions of exposure time. Second, the curves relating $\log [1 - p(\epsilon)]$ and $\log [1 - p(\theta)]$ to exposure time should be parallel to one another, differing only by $\log [1 - p(\beta)]$. And third, $\log [1 - p(\epsilon)]$ should have an intercept of zero.

In the present experiment, $p(\epsilon)$, $p(\theta)$, and $p(\beta)$ were estimated from the data. The estimators are referred to as $p(E)$, $p(D)$, and $p(B)$, respectively. The probabilities $p(E)$ and $p(D)$ were defined as follows: For the ST group, $p(E)$ for any given exposure time was defined as the proportion of study trials in which a detail was named at study and that same detail was then named at test as the basis for a yes response. Referring to Table 1, $p(E) = a$ for any given exposure time. For both the ST and the T groups, $p(D)$ was defined as the proportion of study trials for which a detail was named at test as the basis of a yes response. Referring to Tables 1 and 2, for any given exposure time,

$$\text{ST group: } p(D) = a + b + g$$

$$\text{T group: } p(D) = p.$$

For an exposure time of zero (i.e., for distractor pictures), $p(D)$ and $p(B)$ are identical and are defined as the probability that a detail was used as the basis of a yes response to a distractor picture. Thus, for exposure times of 0, $p(D) = p(B) = k$ for the ST group and t for the T group.

Before presenting the data germane to $p(E)$ and $p(D)$, a few remarks should be made about the rationale underlying the definition of $p(E)$. First, $p(E)$ as defined could logically include instances in which a subject names a detail not actually in the picture at both study and test. This actually never happened. As the reader will probably intuit, a subject never named a nonexistent detail at the time of test. Additionally, it turned out that subjects never named nonexistent details at the time of study either.

The second remark is that there are no assumptions in the model relevant to

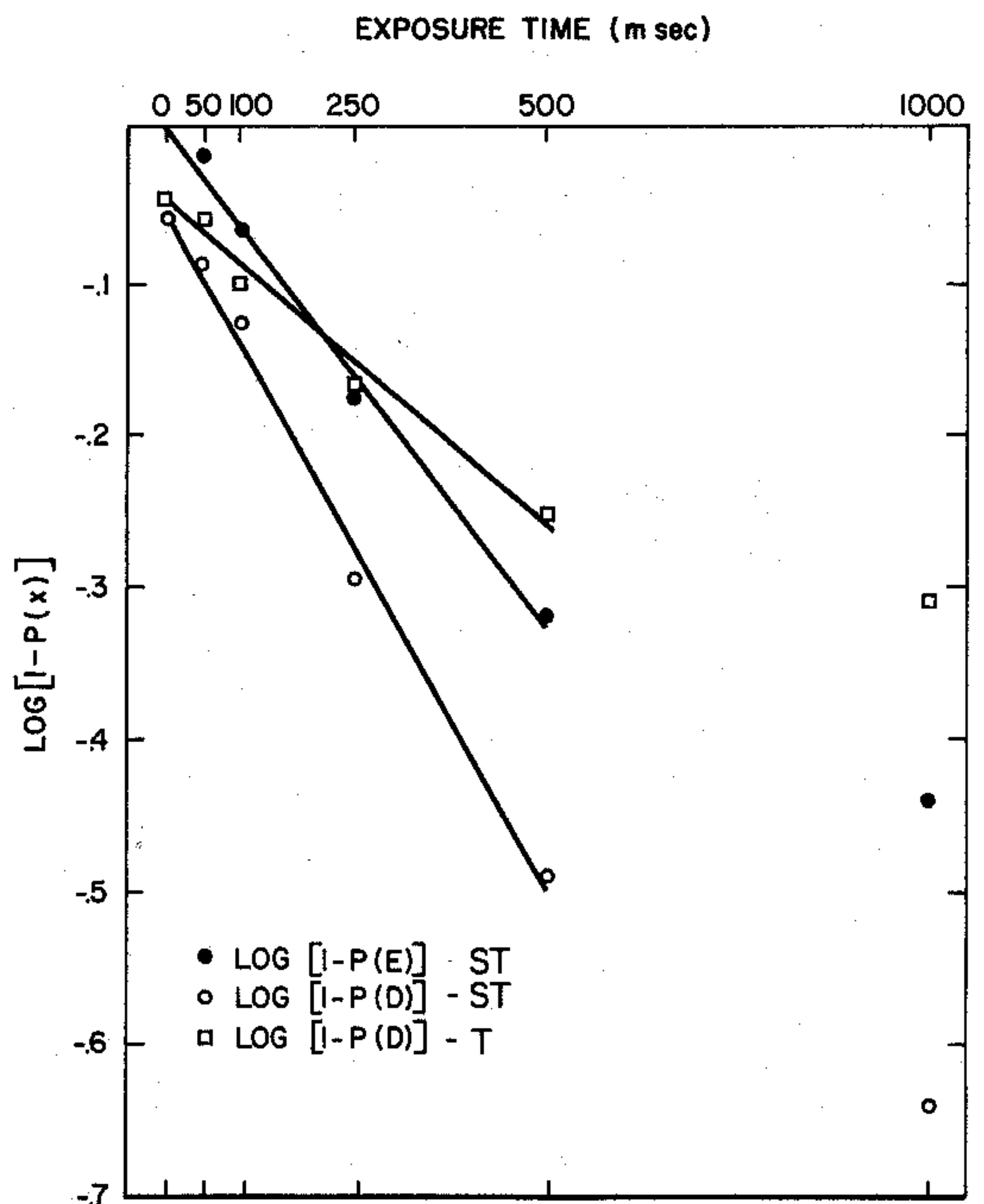


Figure 3. $\log [1 - p(X)]$ as a function of exposure time. [Data points are shown for $X = E$ and $X = D$ for the detail at study and test (ST) and $X = D$ for the detail-only (T) groups. The straight lines are best-fitting regression lines for the data points up to 500 msec.]

possible forgetting of details that were encoded at study. There are several reasons for this omission. First, there is no way that forgetting in the present experiment may be unambiguously measured. As can be computed from Table 1, there are indeed instances in which details named at study are not named at test. This, however, could arise either from forgetting of encoded details or from a subject's assessment at time of test that a detail previously named at study is now not relevant in terms of his/her test decision.

It is, of course, possible to simply *assume* that all details named at study and then not named at test had been forgotten. A forgetting parameter could then be included as part of the model's equations. The model would involve the following assumptions: A detail is encoded with some probability and then forgotten with some probability and a detail is encoded and not forgotten with probability, defined to be $p(\epsilon)$. The rest of the model would then be equivalent to that described earlier.

Table 3

Statistical Analyses of the Three Sets of Data Points in Figure 3

Data point set ^a	<i>F</i> for exposure time	<i>df</i>	<i>F</i> for linear trend	<i>df</i>	% variance accounted for by linear trend	<i>F</i> for residual	<i>df</i>	<i>MS_e</i>
log [1 - <i>p</i> (<i>E</i>)]	34.87*	4, 36	100.2*	1, 36	72	13.11*	3, 36	.0088
log [1 - <i>p</i> (<i>D</i>)] (ST)	31.22*	5, 45	145.8*	1, 45	93	2.57*	4, 45	.0183
log [1 - <i>p</i> (<i>D</i>)] (T)	23.69*	5, 45	106.0*	1, 45	90	3.09*	4, 45	.0048

Note. An individual, one-way, repeated-measures analysis of variance was performed on each of the three sets of data points. The columns depict the *F* values for overall effects of exposure time as well as for linear trend and residual from linear trend.

* *p* ≤ .05 level.

^a Note that there are 5 data points for log [1 - *p*(*E*)] since an exposure time of zero is not included. There are 6 data points for log [1 - *p*(*D*)].

Thus, the addition of forgetting assumptions would do nothing to the fit of the model to the present data—Such assumptions would simply add excess verbiage to the model. For this reason, we ignored forgetting in the present version of the model.

Figure 3 shows log [1 - *p*(*E*)] and log [1 - *p*(*D*)] for the ST group as well as log [1 - *p*(*D*)] for the T group as functions of exposure time. (The significance of the straight lines through the 0- to 500-msec data points will be described later.) At first glance, the results shown in Figure 3 appear to disconfirm the Loftus and Bell (1975) model in two ways. First, none of the three sets of data points is linear; rather, all are concave upward. Statistical evidence for this observation was assessed as follows. A separate, repeated-measures analysis of variance was performed for each of the three sets of data points. The variance due to linear trend and to the residual were then tested. The results of these analyses are shown in Table 3. As can be seen, the overall effects of exposure time as well as the effects of linear trend were all significant. The fact that the *residuals from linearity* were also significant indicates that the hypothesis of linearity is not sufficient to account for all of the variance due to exposure time.

The second apparent disconfirmation of the model stems from the fact that the two

curves from the ST group—log [1 - *p*(*E*)] and log [1 - *p*(*D*)]—were not parallel. We deal with these two apparent disconfirmations in turn.

Bias probabilities. Consider the ST group only. The prediction that the log [1 - *p*(*D*)] and log [1 - *p*(*E*)] curves should be parallel depends on the assumption that bias probabilities, *p*(*β*), do not vary with exposure time. In the Loftus and Bell (1975) data, this assumed constant bias probability was estimated as the proportion of detail responses made to distractor pictures. In the present experiment, however, we were in a position to estimate bias probabilities for each exposure time. In particular, *p*(*B*) was defined as follows: Consider all of the pictures (for a given exposure time) for which a detail was not named at time of study. The proportion of these pictures for which a detail was then named at test as the basis of a yes response was then defined to be the bias probability for that particular exposure time. Thus, referring to Table 1, for any given exposure time,

$$p(B) = \frac{g}{g + h + i + j}$$

The top row of Table 4 shows these bias probabilities. The zero exposure-time bias probability is, of course, the probability of naming a detail as the basis for a yes response to a distractor picture. As can

Table 4
Relationships Among log [1 - p(E)], log [1 - p(D)], and log [1 - p(B)] for ST Group Subjects

Item	Exposure time (msec)					
	0	50	100	250	500	1,000
$p(B)$.1224	.1510	.1272	.2585	.3184	.3292
$\log [1 - p(B)]$	-.0567	-.0711	-.0591	-.1299	-.1665	-.1734
$\log [1 - p(D)]$ minus $\log [1 - p(E)]$		-.0690	-.0614	-.1045	-.1720	-.2006
Error (difference between rows 2 and 3)		-.0021	.0023	-.0254	.0055	.0272

Note. The 4th row entries are the differences between the 2nd and 3rd row entries.

be seen, the bias probabilities did, in general, tend to increase as a function of exposure time. The question of *why* these bias probabilities increased with exposure time is an interesting one and will be taken up again in the Discussion section of this article. With respect to the present arguments, we simply accept this increase as a given.

The second row of Table 2 shows $\log [1 - p(B)]$ as a function of exposure time. Note that according to the model, these values should be equal to the difference between $\log [1 - p(E)]$ and $\log [1 - p(D)]$ for each exposure time. The values of $\{\log [1 - p(D)] - \log [1 - p(E)]\}$ are shown in the third row of Table 4. As can be seen, there was a fairly close correspondence between these values and the $\log [1 - p(B)]$ s. The errors—that is, the difference between rows 2 and 3—are shown in row 4. The mean of the five errors is .0015. An analysis of variance performed

on the errors indicated no effects of exposure time, $F(4, 36) = 1.36$, $MS_e = .0022$. Additionally, the mean error did not differ significantly from zero, $F(1, 9) < 1$, $MS_e = .0030$. In short, the relationships among $p(D)$, $p(E)$, and $p(B)$ for the ST group were precisely what they should have been, thereby providing very strong support for the Loftus and Bell (1975) model.

Refixations. We now deal with the fact that the three curves were not linear as had been predicted by the model. A reasonable explanation for this finding is as follows: It is quite likely that with sufficient exposure time, a subject would refixate previously fixated portions of the picture. The probability of acquiring an informative detail from a refixated area is expected to be less than the probability of acquiring an informative detail from a newly fixated area. Thus, any refixation would entail a lowering of α , resulting in the upward bowing of the curves evident in Figure 3.

Table 5
Statistical Analysis of the Three Sets of Data Points Shown in Figure 3

Date point set ^a	F for exposure time	df	F for linear trend	df	% variance accounted for by linear trend	F for residual	df	MS _e
$\log [1 - p(E)]$	25.56*	3, 27	75.24*	1, 27	98	.77	2, 27	.0072
$\log [1 - p(D)]$ (ST)	25.74*	4, 36	102.3*	1, 36	99	.21	3, 36	.0129
$\log [1 - p(D)]$ (T)	17.64*	4, 36	68.33*	1, 36	97	.74	3, 36	.0042

Note. Organization of this table is identical to that of Table 3. The only difference is that the analyses in this table do not include the 1,000-msec condition.

* $p \leq .05$.

^a Note that there are 4 data points for $\log [1 - p(E)]$, since an exposure time of zero is not included. There are five data points for $\log [1 - p(D)]$.

During 1,000 msec, as many as four or five fixations may be made, thereby providing ample opportunity for refixations. However, 500 msec would rarely be enough time to allow more than two fixations. Since the second fixation on a picture cannot be a refixation, it seems unlikely that many refixations would take place during a 500-msec exposure period. Accordingly, the three curves shown in Figure 3 were reanalyzed using the data from 0 to 500 msec only. The results of this reanalysis are shown in Table 5. As can be seen, the effects of time and linear trend were significant once again for all three curves. However, the residual from linearity was non-significant in all three cases. Considering only the data from 0 to 500 msec, the hypothesis of linearity is sufficient to account for all variance due to exposure time.

Encoding Differences Between the ST and T Groups

The foregoing arguments indicate that the Loftus and Bell (1975) model is a viable instrument. Accordingly, we now use it to examine detail-encoding differences between the ST and the T groups. Recall that the encoding parameter, α , should be reflected by the slopes of the curves shown in Figure 3. A two-way analysis of variance was performed on the $\log [1 - p(D)]$ curves from the ST and the T groups using data points from 0 to 500 msec. This analysis showed a significant Group \times Exposure Time interaction, $F(4, 72) = 5.58$, $MS_e = .0086$. The slope of the ST group curve was significantly greater than the slope of the T group curve, indicating a greater α for the ST group than for the T group. To estimate α , the best-fitting regression lines were computed for the two $\log [1 - p(D)]$ curves. (These regression lines, along with the corresponding line for $\log [1 - p(E)]$ are depicted in Figure 3.) Setting Δt equal to 300 msec—the approximate duration of an eye fixation—the estimates of α were .26 for the T group and .46 for the ST group. (Note that the best-fitting regression line for

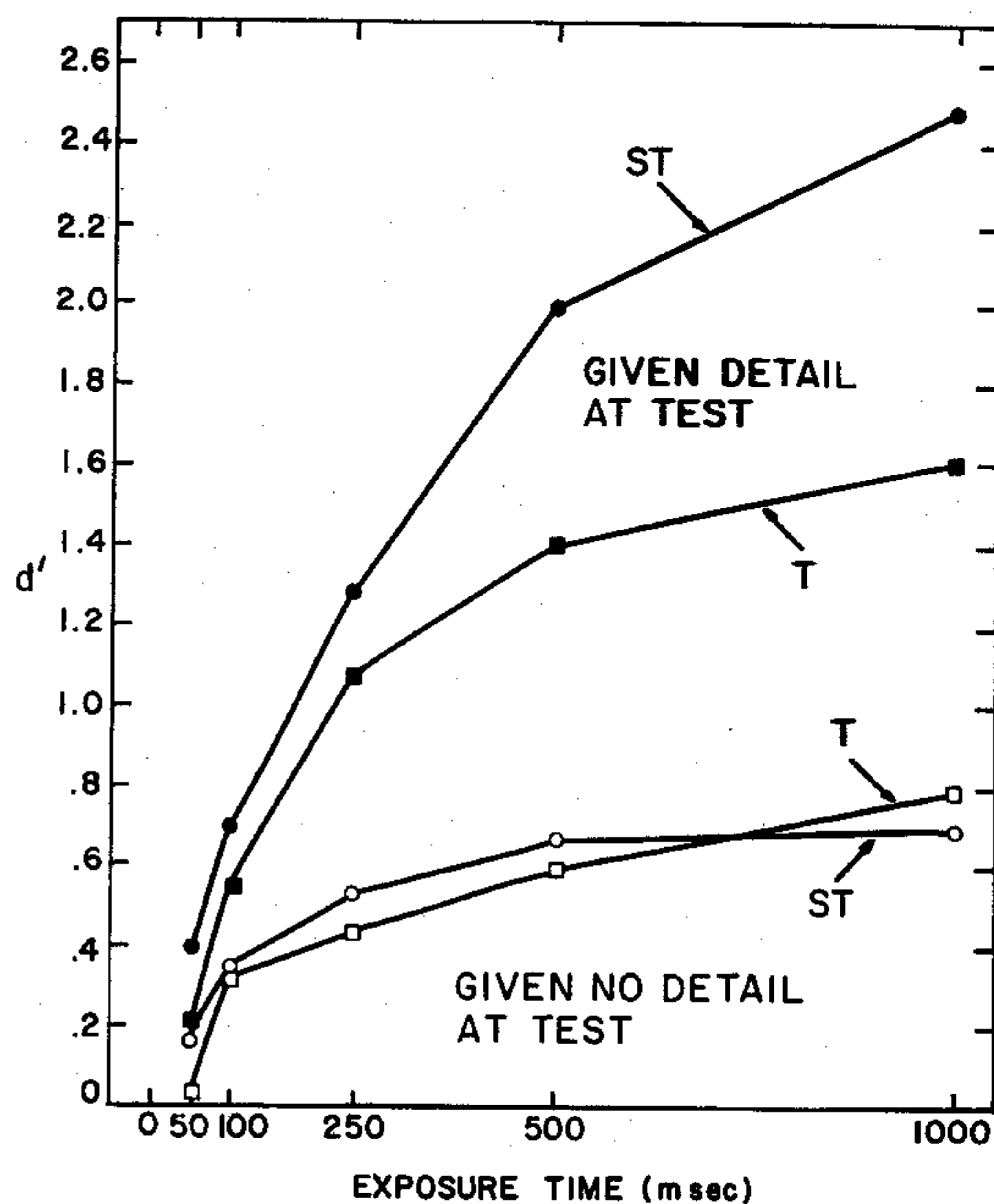


Figure 4. Recognition memory performance as a function of initial exposure time. [The curves are conditionalized on whether or not a detail was named for both the detail at study and test (ST) and the detail-only (T) groups.]

$\log [1 - p(E)]$ resulted in an estimated intercept value of zero, as predicted by the model.)

Recognition Memory Performance Based on Detail Versus General Visual Information

Figure 4 shows recognition memory performance (d') as a function of exposure time. Here, the curves are double conditionalized—separate curves are shown for the ST and the T groups and also separate curves are shown depicting performance when a detail was named at test versus when a detail was not named at test. Referring to Tables 1 and 2, the d' scores shown in Figure 4 were obtained as follows. For a given exposure time,

$$\text{ST—detail: } d' = Z(a + b + g) - Z(k),$$

$$\text{ST—no detail: } d' = Z(c + h) - Z(l),$$

$$\text{T—detail: } d' = Z(p) - Z(t),$$

$$\text{T—no detail: } d' = Z(q) - Z(u).$$

The results of a three-way (Group \times Detail/No Detail \times Exposure Time)

analysis of variance on these data are shown in Table 6. As can be seen, the effect of exposure time and the effect of whether or not a detail was named were both highly significant. Neither the effect of group nor the Group \times Detail interaction was significant. (Although for the latter, $p < .10$.) However, the three-way interaction was significant. When no detail was named at test, the T and ST groups did not differ. This finding is consistent with the notion that no-detail responses are based on general visual information that would not be expected to differ for the two groups. However, a rather substantial performance difference appeared when a detail was named, as reflected by the significant three-way interaction and the marginally significant Group \times Detail interaction. It might have been expected a priori that there would be no difference at all between the ST group and the T group when the data were conditionalized on detail versus no detail. This expectation would be based on the hypothesis that the only reason that the ST group shows better performance than the T group is that the ST group has a propensity to encode details at a greater rate than does the T group. However, as is evident from Figure 4, this was not the case. A detail named at test was more valuable in terms of distinguishing target from distractor pictures for the ST group relative to the T group subjects. This difference may stem from the fact that the ST group subjects *wrote down* details at the time of study, thereby increasing the probability of a verbal code for the detail.

General Discussion

The unconditional recognition results (Figure 2) provide additional evidence for the general notion that multiple codes may underlie memory for pictures. Furthermore, the detail-naming results (Figure 3) provide a good deal of support for the Loftus and Bell (1975) model of detail acquisition. And finally, the results analyzed within the context of the Loftus and Bell model provide fairly precise suggestions as to the mechanism by which detail

Table 6
Analysis of Variance Results for Data Shown in Figure 4

Source	df	F	MS _e
Group (A)	1, 18	2.42	.874
Detail/no detail (B)	1, 18	65.36**	.380
Exposure time (C)	4, 72	104.38**	.096
A \times B	1, 18	3.66*	.380
A \times C	4, 72	1.96	.096
B \times C	4, 72	31.50**	.070
A \times B \times C	4, 72	5.75**	.070

* $.10 < p < .05$.

** $p \leq .05$.

information enters into the picture-recognition process.

Multiple Codes in Picture Memory

As noted in the Introduction, a large body of evidence supports the notion that memory for visual material may be based on multiple codes. The fact that the ST group showed considerably better performance than did the T group (Figure 2) provides additional evidence in support of this notion. In keeping with the findings of Bryant (1965), Kurtz and Hovland (1953), and Freund (1971), verbalization of details during initial study was seen to increase performance in the present experiment. This is consistent with the notion that strengthening of a verbal code during initial viewing augments the memorial representation of a picture.

Support for the Loftus and Bell (1975) Model

Of somewhat more interest is the degree to which the present data provide support for the Loftus and Bell model of detail acquisition. To briefly reiterate, the model involves two major propositions. First, detail-encoding probability is assumed to be constant during each successive unit of time during initial viewing. Second, a detail is *named at test* either if it has been encoded or with some bias probability. When detail-encoding probability, detail-naming probability, and bias probability

are all estimated, they relate to one another exactly as they should according to the model. That is, for any given exposure time, it is the case that $\log [1 - p(D)] = \log [1 - p(E)] + \log [1 - p(B)]$. Additionally, within certain boundary conditions, $\log [1 - p(D)]$ and $\log [1 - p(E)]$ are linear functions of initial exposure time, as predicted by the model.

The principle mystery in the data concerns the increase in bias probability over exposure time as indicated in Table 4. There are (at least) two plausible explanations for this increase. The first explanation would involve the postulation that bias may stem from two sources. Define *positive bias* as the propensity to name details at test that were not encoded. This is the sort of bias originally envisioned by Loftus and Bell, and it would be expected to remain constant over original exposure time. But suppose that there also exists a *negative bias*, which is a propensity to not name details at study even if such details were encoded. It is reasonable to suppose that this negative bias could increase over exposure time as more details were being encoded and were thereby becoming candidates for negative bias.

The second explanation of the bias probability increase is not totally independent of the first, but it would involve the postulation of an alternative to the Loftus and Bell (1975) model. Suppose that longer exposure times as well as instructions to name details at study lead to "deeper encoding" of the picture. Deeper encoding would then lead to superior memory performance, which in turn might increase the probability of detail elicitation, both at study and at test. (Note that this alternative would reverse causality relative to the Loftus and Bell model vis à vis memory performance and detail naming. The plausibility of this alternative illustrates the principle weakness of the present paradigm—detail/no detail is a subject-defined variable, and some of the interesting data are therefore correlational.) In any case, under this hypothesis, pictures would be more deeply encoded over exposure time even if a detail is not named,

and this deeper encoding would subsequently lead to an increase in detail naming at test.

The present data do not permit an assessment of these possibilities; such an assessment must remain the task of the proverbial future research. However, in view of the reasonable, quantitative account of the results offered by the Loftus and Bell (1975) model, it seems reasonable to preserve the model until arrival of data that demand its exit. To the degree that the Loftus and Bell model is a viable one, two goals are accomplished. First, the model provides a substantial building block for a general model of picture recognition (cf. Loftus, 1976). Second, the model may be used as an instrument for examining the effects of various independent variables on detail-encoding rate because α , which can be easily estimated, is a measure of that rate. As examples, Loftus and Bell demonstrated that α is greater for photographs than for corresponding line drawings. Mathews (Note 1) has demonstrated that details are encoded at a slower rate when subjects are forced to carry out a distracting task while viewing pictures relative to a control condition involving normal viewing. And, as will be discussed later, the present experiment has demonstrated that subjects who must name details at study encode them at a faster rate than do subjects who do not have to name details.

Effects of Naming Details on Picture Recognition

As described earlier, the ST group in the present study showed superior performance relative to the T group. There are several possible reasons for this result. First, estimation of α for the two groups indicates that ST subjects encode details from pictures at a faster rate than do T subjects. This result makes sense: A subject who is trying to write down details at study would be quite naturally inclined to focus attention on the encoding of details during viewing. What the ST subjects may not realize is that encoding

details constitutes an excellent strategy in terms of ultimately being able to recognize the picture. Second, a detail named at test is a more valuable distinguishing feature for the ST relative to the T group. It is not entirely clear why this is so. A reasonable hypothesis, however, is that writing down a detail at study provides an additional (perhaps verbal) memory code for that detail.

Reference Note

1. Mathews, P. Unpublished data, University of Washington, 1974.

References

- Atwood, G. An experimental study of visual imagery and memory. *Cognitive Psychology*, 1971, 2, 290-299.
- Bahrick, H. P., & Bahrick, P. Independence of verbal and visual codes of the same stimuli. *Journal of Experimental Psychology*, 1971, 91, 344-346.
- Bahrick, H. P., & Boucher, B. Retention of visual and verbal codes of the same stimuli. *Journal of Experimental Psychology*, 1968, 78, 417-422.
- Brooks, L. R. The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology*, 1967, 19, 289-299.
- Brooks, L. R. Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, 1968, 22, 349-368.
- Bryant, P. E. The effects of verbal labelling on recognition of pictures and names in severely subnormal and normal children. *Journal of Mental Deficiency Research*, 1965, 9, 237-244.
- Dallett, K., & Wilcox, S. G. Remembering pictures vs. remembering descriptions. *Psychonomic Science*, 1968, 11, 139-140.
- Dallett, K., Wilcox, S. G., & d'Andrea, L. Picture memory experiments. *Journal of Experimental Psychology*, 1968, 76, 312-320.
- Davies, G., Milne, J. E., & Glennie, B. J. On the significance of "double encoding" for superior recall of pictures to names. *Quarterly Journal of Experimental Psychology*, 1973, 25, 413-423.
- Freund, R. D. *Verbal and non-verbal processes in picture recognition*. Unpublished doctoral dissertation, Stanford University, 1971.
- Jenkins, J. R., Neale, D. C., & Deno, S. L. Differential memory for picture and word stimuli. *Journal of Educational Psychology*, 1967, 58, 303-307.
- Kurtz, K., & Hovland, C. The effect of verbalization during observation of stimulus objects upon accuracy of recognition and recall. *Journal of Experimental Psychology*, 1953, 45, 157-164.
- Lieberman, L. R., & Culpepper, J. T. Words vs. objects: Comparison of free verbal recall. *Psychological Reports*, 1965, 17, 983-988.
- Loftus, G. R. Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 1972, 3, 525-551.
- Loftus, G. R. A framework for a theory of picture memory. In J. W. Senders & R. Monty (Eds.), *Eye movements and physiological processes*. Hillsdale, N. J.: Erlbaum, 1976.
- Loftus, G. R., & Bell, S. M. Two types of information in picture memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1975, 1, 103-113.
- Mandler, J. M., & Stein, N. L. Recall and recognition of pictures by children as a function of organization and distractor similarity. *Journal of Experimental Psychology*, 1974, 102, 657-669.
- 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. *Imagery and verbal processes*. New York: Holt, Rinehart & Winston, 1971.
- Paivio, A., & Csapo, K. Concrete image and verbal memory codes. *Journal of Experimental Psychology*, 1969, 80, 279-285.
- Paivio, A., & Csapo, K. Short-term sequential memory for pictures and words. *Psychonomic Science*, 1971, 24, 50-51.
- Paivio, A., & Csapo, K. Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, 1973, 5, 176-206.
- Petrinovich, L. F., & Hardyck, C. D. Generalization of an instrumental response between words and pictures. *Psychonomic Science*, 1970, 18, 239-241.
- Rowe, E. J., & Rogers, T. B. Effects of concurrent auditory shadowing on free recall and recognition of pictures and words. *Journal of Experimental Psychology: Human Learning and Memory*, 1975, 1, 415-422.
- Seamon, J. G., & Gazzaniga, M. S. Coding strategy and cerebral laterality effects. *Cognitive Psychology*, 1973, 5, 249-256.
- Sperry, R. W. Cerebral organization and behavior. *Science*, 1961, 133, 1749-1756.
- Sperry, R. W. Hemisphere disconnection and unity in conscious experience. *American Psychologist*, 1968, 23, 723-733.
- Sternberg, S. High-speed scanning in human memory. *Science*, 1966, 153, 652-654.
- Szewczuk, W. The role of verbalization in the retention of nonverbal material. *Polish Psychological Bulletin*, 1970, 1, 87-97.
- Tversky, B. Eye fixations in prediction of recognition and recall. *Memory & Cognition*, 1974, 2, 275-278.

Received July 17, 1978

Revision received November 3, 1978 ■