Journal of Experimental Psychology:
Learning, Memory, and Cognition
1987, Vol. 13, No. 1, 27–35

0278-7393/87/$00.75

Remembering Left–Right Orientation of Pictures

James C. Bartlett
Program in Human Development and Communication Sciences
University of Texas at Dallas

Morton Ann Gernsbacher
University of Oregon

Robert E. Till
University of North Dakota

In a study of recognition memory for pictures, we observed an asymmetry in classifying test items as "same" versus "different" in left–right orientation: Identical copies of previously viewed items were classified more accurately than left–right reversals of those items. Response bias could not explain this asymmetry, and, moreover, correct "same" and "different" classifications were independently manipulable: Whereas repetition of input pictures (one vs. two presentations) affected primarily correct "same" classifications, retention interval (3 hr vs. 1 week) affected primarily correct "different" classifications. In addition, repetition but not retention interval affected judgments that previously seen pictures (both identical and reversed) were "old." These and additional findings supported a dual-process hypothesis that links "same" classifications to high familiarity, and "different" classifications to conscious sampling of images of previously viewed pictures.

The confusability of visual stimuli with their left–right reversals is a well-known fact of perception and memory (Corballis & Beale, 1976). Children have difficulty distinguishing *b* from *d,* and even adults are slower to judge the left–right orientation of a stimulus than its up–down orientation (Farrell, 1979). Adults are also prone to left–right confusions when drawing previously viewed pictures from memory (Bartlett, 1932), even when these images are as frequently experienced as the profile on a penny (Nickerson & Adams, 1979). Furthermore, in tests of adult recognition memory for pictures, laterally reversed copies of input items are frequently recognized as "old," but incorrectly classified as "same orientation" (e.g., Intraub, 1980; Standing, Conezio, & Haber, 1970).

Given that recognition of pictures is often quite accurate (Shepard, 1967), left–right confusions in such recognition are provocative for theories of visual long-term memory. These left–right confusions would seem to carry implications for how pictorial information is represented internally. It is tempting to speculate about the type of memory trace that would allow for accurate old–new judgments, and yet also produce frequent misclassifications of "same orientation" for reversed test items. However, speculation about memory traces causing left–right confusions must be based on some notion of how picture orientation is retrieved from long-term memory. Thus, we view workable hypotheses for orientation retrieval as critical for theories of memory traces for pictures. Developing one such hypothesis was our goal in this study.

Fortunately, although orientation retrieval is relatively unexplored (but see Jolicoeur, 1985), two observations offer clues to its nature. One observation comes from tasks in which subjects attempt to judge both identical copies and left–right reversals of previously viewed pictures as "old." The hit rate in such old–new tasks is usually higher for identical items than it is for reversed items, but the difference is small (Bartlett, Till, Gernsbacher, & Gorman, 1983; Bartlett, Till, & Levy, 1980; McKelvie, 1983), sometimes unreliable (Dallett, Wilcox, & D'Andrea, 1968; Intraub, 1980; Klatzky & Forrest, 1984), and sometimes apparently absent altogether (Standing et al., 1970).

A second observation comes from tasks in which subjects must classify "old" pictures as "same" versus "different" in left–right orientation. Many such tasks require old–new judgments along with same–different classifications of all items judged "old" (e.g., Klatzky & Forrest, 1984). Others require only same–different classifications (e.g., Madigan, 1983). However, regardless of whether old–new judgments are made, correct "same" classifications of identical-copy items are much more probable than correct "different" classifications of left–right-reversed items. Put differently, accuracy of orientation classification is markedly asymmetric (Gernsbacher, 1985; Klatzky & Forrest, 1984; Madigan, 1983; McKelvie, 1983). Although the asymmetry in orientation classification has been attributed to response bias (Klatzky & Forrest, 1984), in pilot studies we have found that when entirely new lures are falsely judged to be "old," they are no more likely to be classified "same" than "different" in orientation. In one such study using 52 subjects and three different encoding tasks, entirely new lures were classified "new," "same," and "different" at rates of .89, .05, and .06, respectively. Although this pattern suggested no general bias favoring "same" over "different" responses, the

We thank Rosemary Frederickson for conducting the experiment and W. Jay Dowling, Valerie Reyna, Henry L. Roediger III, Douglas Nelson, and Margaret Intons-Peterson for valuable comments on drafts of the article.

Correspondence concerning this article should be addressed to James C. Bartlett, Program in Human Development and Communication Sciences, University of Texas at Dallas, Box 688, Richardson, Texas 75080.

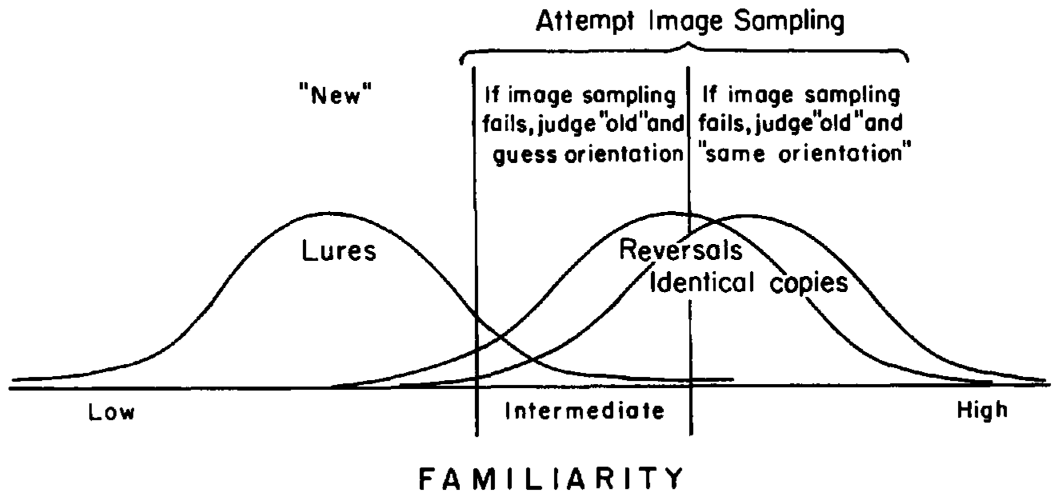

Figure 1. A hypothetical dual-process hypothesis for picture recognition and left–right confusions. (Shown are familiarity distributions for identical test items [right-most distribution], reversed test items, and entirely new lures [left-most distribution], and vertical lines representing high-[right-most line] and low-[left-most line] familiarity criteria.)

probability of "same" classifications of identical-copy items averaged .74, whereas the probability of "different" classifications of left–right-reversed items averaged only .55, a highly reliable difference, $F(1, 49) = 52.3$.[1]

A Dual-Process Hypothesis for Orientation Retrieval

The small effect of orientation on old–new judgments, and the larger, asymmetric effect of orientation on same–different classifications can be approached theoretically via several recent models of how recognition judgments are made (e.g., Gillund & Shiffrin, 1984; Mandler, 1980; Ratcliff, 1978; Tulving, 1983). Despite their basis in findings obtained with verbal materials, these models share a notion that is applicable to pictures: Recognition judgments made to previously encountered stimuli are affected by the overall match or overlap between features of a recognition cue and one or more traces in memory. The degree of such match presumably determines perceived familiarity, or some similar variable (resonance, ecphory) that plays a large role in determining old–new decisions. However, under at least some conditions, more active retrieval processes of image sampling (or conscious recollection) might be used by subjects.

Applying the concepts of familiarity and image-sampling processes to the previously described phenomena of orientation retrieval, we offer a dual-process hypothesis that embodies three claims: First, the familiarity process used in recognizing pictures is only weakly sensitive to orientation information in memory. Second, because of this property of the familiarity process, image sampling is important as a basis for orientation classifications. Third, in cases in which image sampling fails, high familiarity leads to a "same-orientation" classification. These three claims of the dual-process view are illustrated in Figure 1 and explicated below.

The first claim of the dual-process hypothesis is that the perceived familiarity of previously viewed pictures is only slightly affected by orientation-specific features (e.g., left–right locations of objects within pictures). Instead, it is more strongly affected by orientation-independent information (e.g., specific details, Loftus & Bell, 1975, and even verbal information, Bartlett et al., 1980). In consequence, perceived familiarity is only slightly greater for identical than for reversed items. It is, of course, substantially greater for both identical and reversed items than for entirely new lures.

The second claim of the dual-process hypothesis follows naturally from the first. Because familiarity is only slightly greater for identical than reversed test items, classifying test items by same–different orientation benefits from an additional process. We describe this process as one of sampling an image of a previously viewed picture (Gillund & Shiffrin, 1984), though we could just as easily describe it as conscious recollection based on the products of a retrieval cue interacting with a memory trace (Tulving, 1983). We assume that image sampling is attempted when (a) a subject judges that a test item is at least moderately familiar (its familiarity exceeds the lower criterion in Figure 1) and (b) the task is to classify this item as "same" or "different" orientation. Such image sampling can lead to recovery of attributes of a previously viewed picture, followed by analytical comparison of these attributes with the test item. Thus, if image sampling succeeds, the subject is likely to detect an orientation match (leading to a "same-orientation" classification) if the picture is an identical item, or an orientation mismatch (leading to a "different" classification) if the item is a reversal. The subject might also detect mismatches beyond those of orientation (leading to a judgment of "new"), though we do not address this possibility here.

[1] A report including this and another pilot study is available on request.

This second claim of the dual-process hypothesis might seem to contradict evidence that many phenomena of recognition memory can be explained without appealing to image-sampling processes (Gillund & Shiffrin, 1984), or that image sampling (or search) is attempted only if familiarity is moderate in value (Atkinson & Juola, 1973). However, the present dual-process hypothesis applies to recognition tasks requiring subjects not simply to make old–new judgments, but also to classify old items by same–different orientation. In such tasks it is plausible—unless data suggest otherwise—that in order to maximize classification accuracy, subjects attempt image sampling for all items that seem old (i.e., have familiarity exceeding a low criterion). The extent of image sampling in standard recognition tasks requiring only old–new judgments is left open by this hypothesis.

The third claim of the dual-process hypothesis is that high familiarity—that is, familiarity exceeding some high criterion—sometimes leads to a "same" classification. More specifically, as shown in Figure 1, when familiarity is high but image sampling fails (i.e., when it cannot be performed or produces ambiguous information), both an "old" judgment and a "same" classification are made. When familiarity is intermediate and image sampling fails, an "old" judgment is made and a "same" or a "different" classification is guessed. When familiarity is low, image sampling is not attempted; the subject simply makes a "new" response.

This third claim of the dual-process hypothesis is perhaps the most critical, as it accounts for the finding that "same" classifications of identical items are more frequent than "different" classifications of reversals, even if there is no response bias in classifications of lures. Note first that, according to the hypothesis, if an identical item has high familiarity, it will be correctly classified as "same" regardless of whether image sampling succeeds. In contrast, if a reversal has high familiarity, it will be correctly classified as "different" only if image sampling succeeds; otherwise it will be wrongly classified as "same." Thus, reversals are classified with relatively low accuracy because image sampling—a failure-prone process—is a more important basis for "different" than for "same" classifications.

Note second that although a proportion of lures might be moderately familiar, only a relatively small number will be highly familiar. Hence, same–different classification of most falsely recognized lures either is based on image sampling, or else is simply guessed. The consequence is that little or no bias favoring "same" classifications need be manifest in subjects' responses to lures.

Predictions of the Dual-Process Hypothesis

Although the dual-process hypothesis is certainly speculative, it has an implication that can fruitfully guide research: There should be independent variables that (a) affect image sampling but not familiarity and (b) affect familiarity but not image sampling. These two types of variables should affect recognition as shown in Table 1.

Consider a hypothetical experimental manipulation that improves image sampling without affecting familiarity (Table 1, Row 1). According to the hypothesis, such a manipulation should substantially increase "different" classifications of reversals while having comparatively little effect on "same" classifications of identical items. This is because image sampling is needed for "same" classifications only if familiarity is moderate (see Figure 1). Moreover, an image-sampling manipulation should have no effect on judgments that identical items and reversals are "old." The reason is that those identical and reversed items that exceed the lower, familiarity criterion will be called "old" even if image sampling fails.

Now consider a hypothetical experimental manipulation that does not affect image sampling, but does increase familiarity of both identical and reversed items (but not lures). "Old" judgments to both identical and reversed items should clearly be affected. This is because familiarity of both item types should more frequently exceed the lower criterion in Figure 1. In addition, "same" classifications of identical items should be increased, because a boost in familiarity should place more of these items in the region above the high criterion, where a "same" classification is always made. However, we should not expect an increase in classifying reversals "different," because a boost in familiarity should place more reversals in the region above the high criterion, which should work to increase erroneous "same" classifications (when the attempt at image sampling fails). This negative effect might be ameliorated by reversals of relatively low familiarity. An overall boost in familiarity might move many such items from the low-familiarity zone, where a

Table 1
Predictions of the Dual-Process Hypothesis for an Experimental Condition Improving Image Sampling, Increasing Familiarity of Identical and Reversed Items, or Both

	Measure and item type			
	"Old" judgments		Correct classifications	
Effect of condition	Identical items	Reversals	Identical items	Reversals
Image sampling improved	no effect[a]	no effect[a]	little or no effect	increase
Familiarity increased	increase	increase	increase[b]	little or no effect
Both	increase	increase	increase[b]	increase

[a] Image sampling might contribute to old–new discrimination by reducing false alarms to new items. [b] This prediction may not hold for the conditional probability of "same" judgments to identical items, given recognition of these items as "old."

"new" judgment is made, to the intermediate-familiarity zone, where a "different" classification is at least possible (through either guessing or successful image sampling). Nonetheless, an overall increase in correct "different" classifications is likely to be small if present at all.

A Test of the Dual-Process Hypothesis

Upon initiating our studies of left–right confusions, we were faced with a problem: Even assuming the validity of the dual-process hypothesis, it was unclear what variables might (a) affect image sampling but not familiarity or (b) affect familiarity but not image sampling. We therefore conducted some pilot investigations, one of which showed that a 1-week delay of the recognition test produced effects similar to the image-sampling pattern (Table 1, Row 1): Whereas retention interval did not affect "old" judgments to identical or reversed items, correct classifications were significantly less probable with a 1-week delay than a 5-min delay. Although these findings appeared promising, the predictions of the dual-process hypothesis (Table 1) are based on an assumption that appeared to be questionable in our pilot investigation. This assumption is that criterion placement—what a subject perceives as intermediate familiarity and high familiarity (Figure 1)—does not change between conditions. Yet in our pilot study, retention interval was a between-subjects variable, and it was possible that the criteria used by 5-min-retention subjects differed from those used by 1-week-retention subjects. Thus, a more definitive test of retention-interval effects was necessary.

A major goal for the present experiment was to test the main implication of our pilot investigation: A 1-week delay of the recognition test affects image sampling but not familiarity of old items. To examine effects of retention interval while controlling for those of criterion placement, this experiment used the following procedure: An input list of pictures was presented to subjects, followed after 1 week by a second input list. The recognition test, which included old items from both of these lists, occurred approximately 3 hr after the second input list. Because retention interval was a within-subjects variable, and because short- and long-retention-interval items were intermixed in the test, it seemed safe to assume that retention interval would not be confounded with criterion placement. On the basis of the dual-process hypothesis together with our pilot results, we predicted that the 1-week condition—as compared with the 3-hr condition—would show a reliable reduction in correct "different" classifications, but relatively little or no reduction in correct "same" classifications, and no reduction in "old" judgments to identical or reversed items (Table 1, Row 1).

A second goal for this experiment was to test the effects of familiarity along with those of image sampling. In pursuit of this goal, half of the pictures in each input list were presented two times (with a lag of four intervening items). Virtually all theories of recognition memory claim that repeating input items should increase familiarity (e.g., Gillund & Shiffrin, 1984; Mandler, 1980). Hence, we expected that whereas the effects of retention interval would fit the predictions for an image-sampling variable (Table 1, Row 1), the effects of repetition would (a) fit the predictions of a familiarity variable (Table 1, Row 2) or (b) match the pattern expected if both familiarity and image sampling were improved by repetition (Table 1, Row 3). That is, we predicted that two input presentations would increase "old" judgments to both identical and reversed items, and would improve classification of the former as "same" (Table 1, Row 2). We did not predict whether two presentations would improve classification of reversed items as "different."

A third goal was to clarify the currently mixed evidence that reversing test pictures affects familiarity. Hence, we planned to examine whether identical items would receive more "old" judgments than left–right-reversed items. We were interested also in whether identical items would be judged as more frequent and/or more recent than reversals. Hintzman and Stern (1977) found that reversing test pictures reduced their judged frequency. If this somewhat surprising finding proves to be replicable, it would suggest that orientation affects familiarity and that familiarity in turn plays some role in frequency judgments. In order to examine this issue, which is important for extending the dual-process model beyond the limited domain of orientation and "old" judgments, the recognition task we give to our subjects required classifying "old" items not only by orientation, but also by frequency as well as recency.

Method

Subjects

The 32 subjects were 21 female and 11 male students taking psychology courses at The University of Texas at Dallas. The majority (30) were undergraduates who participated as one alternative means of fulfilling a course requirement (the remaining 2 subjects were graduate students who participated on a volunteer basis). The average age in the sample was 29 years (SD = 7.7), and all but 3 subjects were right-handed.

Materials

The stimulus materials were 96 scenic color slides (35 mm) of landscapes and cityscapes; each was a cropped version of a magazine photograph and entirely devoid of writing. The slides were selected by the criterion of being reasonably asymmetric, with a salient object or feature appearing on the left side but not the right side, or vice versa.

Design

The experiment involved three separate sessions: (a) a first input session immediately preceding or following a psychology class, (b) a second input session immediately preceding the class 1 week later, and (c) a testing session following the class that same day. Owing to variation in length of classes, the testing session was given 1.5 to 3.5 hr (M = 2.7) after the second input session.

Within each of the two input sessions (1 week apart), a unique (nonoverlapping) set of 24 scenic color photographs was shown. Within each set, half (12) of the pictures were presented once, whereas the remainder were presented twice in the same orientation with a lag of four intervening items. Thus, each input set contained 36 slides depicting 24 unique scenes. The sequencing of once- and twice-presented items was random with the constraint that the first presentation of a twice-presented picture could occur no later than serial position 31. Thus, the second presentation of a twice-presented item could occur no later than serial position 36, the last position in the list.

The recognition test included all 48 old items, in addition to 48 entirely new lures. The 48 old items comprised six pictures assigned to eight different conditions. These conditions were defined by the factorial combination of repetition (one vs. two input presentations), retention interval (first vs. second input list), and orientation (identical-copy vs. reversal). Note that when pictures were presented twice at input, they were shown in the same orientation both times. However, half of these twice-presented pictures (as well as half of the once-presented pictures) were reversed when shown at test. Thus, repetition, retention interval, and orientation-at-test of old items were orthogonal factors.

The test pictures were randomly sequenced. Further, the "old" test pictures were counterbalanced over the eight conditions such that each of them served in each combination of repetition, retention-interval, and orientation-at-test condition with a subgroup of 4 subjects. The 48 new pictures were the same for all subjects.

Procedure

The instruction for the input lists was to examine carefully each picture presented. Although the subjects were aware this was a memory experiment, no details regarding the forthcoming test were given. The instruction for the recognition test was to judge each picture as "old" or "new," and to classify each "old" item as (a) same or reversed in orientation, (b) once or twice presented, and (c) recent ("today") or nonrecent ("last week"). Responses were made on answer sheets containing seven columns, the first and second columns labeled "same" and "different," the third and fourth columns labeled "once" and "twice," the fifth and sixth columns labeled "today" and "last week," and the seventh column labeled "new." Subjects responded to each item either by marking the seventh column (for a "new" judgment), or by marking one or the other of the first and second columns, one or the other of the third and fourth columns, and one or the other of the fifth and sixth columns (for an "old" judgment and classifications of orientation, frequency, and recency). The slides were presented for 3 s per item at input, and 10 s per item at test, with an interstimulus interval equal to the projector's change time (about 1 s).

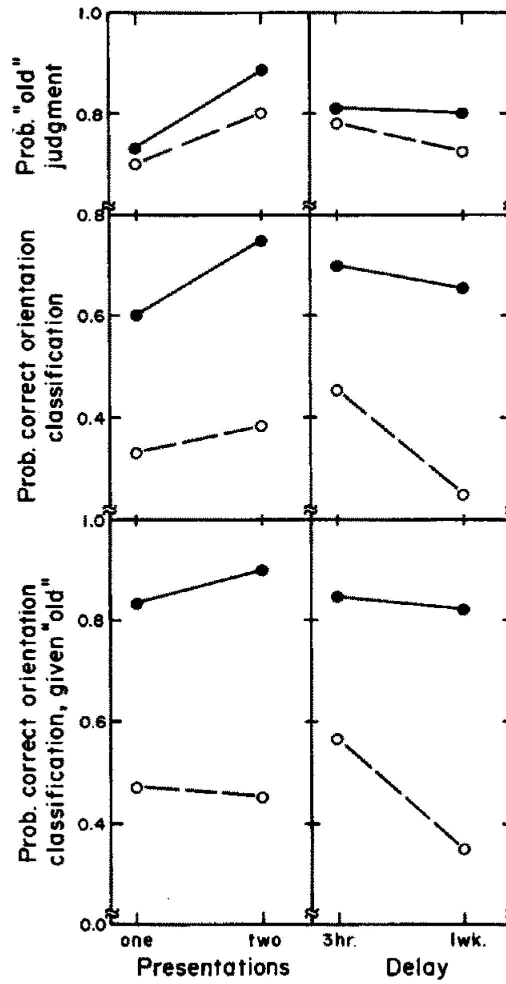

Figure 2. Mean probabilities of "old" judgments (top panels), correct orientation classifications (middle panels) and correct classifications, given an "old" judgment (bottom panels) for identical-copy items (filled circles) and left–right reversals (empty circles) by repetition condition (left panels) and retention interval (right panels).

Results

The principal empirical questions (see Table 1) concerned "old" judgments and orientation classifications made to identical and reversed items in the recognition test. Hence, we first report analyses of variance (ANOVAs) of effects of orientation at test (same vs. reversed), repetition at input, (one vs. two presentations), and retention interval (3 hr vs. 1 week), using each of these two measures. We then turn to some analyses based on (erroneous) "old" judgments and same–different classifications made to entirely new lures. Finally, we report a treatment of "twice" (versus "once") judgments and "today" (versus "last week") judgments to identical and reversed items, as well as to lures. The α level was .05.

Old Judgment Probabilities

The top two panels of Figure 2 show probabilities of "old" judgments to identical and reversed items as a function of repetition (left panel) and, with data regrouped, retention interval (right panel). An ANOVA supported what is clear from the Figure: "Old" judgments were more probable for identical than reversed items, $F(1, 31) = 5.22$, $MS_e = .038$, and for twice-presented items than for once-presented items, $F(1, 31) = 43.7$, $MS_e = .025$. However, there was no reliable difference between more and less recent items, $F(1, 31) = 2.04$, $MS_e = .051$, $p > .10$, replicating the findings of our pilot study. Moreover, the ANOVA supported no reliable interactions, $Fs(1, 31) < 2.4$, $ps > .10$.

Orientation Classifications

The two middle panels of Figure 2 show the probabilities of correct orientation classifications of identical items ("same") and reversals ("different") by repetition (left panel) and retention interval (right panel). As suggested by the figure, an ANOVA supported main effects of all three variables, $F(1, 31) = 124.2$, $MS_e = .054$ for orientation, $F(1, 31) = 17.2$, $MS_e = .036$ for repetition, and $F(1, 31) = 13.2$, $MS_e = .079$ for retention interval. More important, there were reliable interactions between orientation and repetition, $F(1, 31) = 4.98$, $MS_e = .035$, and between orientation and retention interval, $F(1, 31) = 14.0$, $MS_e = .034$. Whereas repetition increased "same" classifications of identical items, $t(31) = 4.51$, $p < .001$, without markedly affecting "different" classifications of reversals, $t(31) = 1.39$, $p > .10$, retention interval produced the opposite pattern: a robust effect on "different" classifications of reversals, $t(31) = 5.18$, $p < .001$, but not on "same" classifications of identical items, $t(31) < 1$. No other effects in the ANOVA were reliable. Referring to Table 1, it can be seen that repetition produced the pattern expected if familiarity were being manipulated, whereas retention interval produced the pattern expected if an image-sampling process were being manipulated.[2]

The two bottom panels of Figure 2 show the conditional probabilities of correct orientation classifications, given recognition of items as "old." The effects of repetition—but not those of retention interval—apparently were attenuated with this measure. Indeed, an ANOVA performed on these conditional probabilities supported no effects involving repetition, although it did support a main effect for orientation, $F(1, 31) = 92.4$, $MS_e = .097$, a main effect for retention interval, $F(1, 31) = 13.5$, $MS_e = .067$, and an Orientation × Retention Interval interaction, $F(1, 31) = 15.6$, $MS_e = .038$. Because such conditional probabilities should be insensitive to familiarity effects (such probabilities pertain to items selected to be at least moderately familiar), the pattern is sensible: It supports the conclusion that repetition affected perceived familiarity, whereas retention interval affected image sampling.

"Old" and "Same–Different" Responses to Lures

The probability of "same" classifications of entirely new lures averaged .23, whereas that of "different" classifications of entirely new lures averaged only .12. These two figures sum to .35, which is the overall probability of "old" judgments to lures. What is most important here is the indication of response bias favoring "same" classifications over "different" classifications. Indeed, the response bias effect was reliable, $t(31) = 5.34$, $p < .001$.

To clarify the consequences of the response bias favoring "same" classifications (which was not present in our pilot investigations), we used responses to lures to divide our subjects into a biased subgroup ("same" and "different" classification probabilities averaging .28 and .07, respectively, $n = 15$) and a nonbiased subgroup ("same" and "different" probabilities averaging .18 and .17, respectively, $n = 17$). Having made this division, the (raw) probabilities of correct orientation classification of identical and reversed test items were subjected to an additional ANOVA in which response bias was a between-subjects factor. All previously reported main effects and interactions were once again reliable, although there was also a bias main effect, $F(1, 30) = 11.9$, $MS_e = .075$, qualified by a Bias × Orientation interaction, $F(1, 30) = 6.28$, $MS_e = .046$: Correct "same" classifications exceeded correct "different" classifications by a somewhat greater margin in the response-biased subgroup (mean probabilities = .65 and .26, respectively) than in the nonbiased subgroup (mean probabilities = .70 and .44, respectively). Note, however, that (a) the orientation effect was substantial even in the nonbiased subgroup, and (b) as shown in Table 2, the two subgroups showed the same pattern of repetition and retention-interval effects. We conclude that response bias might strengthen the asymmetry in orientation classification, but it is not a necessary prerequisite for this asymmetry, or for the qualitative pattern of repetition and retention-interval effects.[3]

Table 2
Probabilities of Classifying Identical Items "Same" and Reversals "Different" for each Repetition and Retention-Interval Condition by Subjects Showing Response Bias Versus No Response Bias for "Same" Classifications

	Repetition and retention-interval condition			
	Presentations		Retention interval	
Item type	Once	Twice	3 hr	1 week
	Biased subgroup ($n = 15$)			
Identical items	.56	.74**	.67	.63
Reversals	.23	.28	.31	.20*
	Nonbiased subgroup ($n = 17$)			
Identical items	.64	.76**	.73	.68
Reversals	.42	.46	.59	.29**

* $p < .05$. ** $p < .01$.

"Twice" and "Today" Judgments

Our major concern here was with the effects of orientation on judged frequency and judged recency. A straightforward way to assess these effects is to examine probabilities of "twice" and "today" judgments to identical and reversed items by repetition and retention interval. "Twice" and "today" judgments to lures are analyzed separately.

Figure 3 (top panels) displays the mean probabilities of "twice" judgments to identical and reversed items. An ANOVA of these data showed a main effect of repetition, $F(1, 31) = 89.5$, $MS_e = .055$ (as would be expected), and also an effect of orientation, $F(1, 31) = 21.1$, $MS_e = .026$; identical items received

[2] A question relevant to interpreting subjects' orientation-classifications is whether their orientation memory was greater than chance. The probabilities of correct orientation classifications (Figure 2, middle panels) do not answer this question, as such probabilities have no well-defined values of chance expectations. However, it is informative to compare probabilities of "same" classification responses to identical items (hit rate) versus reversed items (false-alarm rate). Collapsing across repetition and retention-interval condition, the "same" response hit rate exceeded the "same" response false-alarm rate (.68 versus .40, respectively), supporting above-chance memory for orientation of pictures. The poorest orientation memory was found in the single-presentation/1-week-interval condition, but even there the hit rate and false-alarm rate were different (.57 vs. .42), $t(31) = 2.46$, $p < .05$.

[3] We note that in the ANOVA the Bias × Orientation × Retention Interval interaction approached significance, $F(1, 30) = 4.05$, $MS_e = .031$, $.05 < p < .10$. As shown in Table 2, the Orientation × Retention Interval interaction appeared more pronounced among the nonbiased subjects.

"twice" judgments more often than reversals. Yet, as with "old" judgments, there was no effect of retention interval and there were no reliable interactions. The main effects of repetition and orientation were also reliable in an ANOVA of the conditional probabilities of "twice" judgments, given recognition that items were "old" (Figure 3, bottom panels). Though not shown in Figure 3, this second ANOVA also showed a Retention Interval × Repetition interaction, $F(1, 31) = 4.57$, $MS_e = .057$: Correct "twice" judgments to twice-presented items were similar in the short- and long-retention conditions (.57 and .55, respectively), but erroneous "twice" judgments to once-presented items were lower in the short-retention condition (.24 vs. .34). Thus, the shorter retention interval produced fewer confusions of once- with twice-presented items.

Figure 4 (top panels) displays mean probabilities of "today" judgments to identical and reversed items. An ANOVA revealed an expected main effect for retention interval, $F(1, 31) = 124.0$, $MS_e = .103$, and a (much smaller) main effect for repetition, $F(1, 31) = 10.4$, $MS_e = .027$. Not shown in the figure was a Retention Interval × Repetition interaction, $F(1, 31) = 6.92$, $MS_e = .038$: The probabilities of "today" judgments were .61 for twice-presented short-retention-interval items, .48 for once-presented short-retention-interval items, and .10 for both twice- and once-presented long-retention-interval items. Thus, repetition apparently improved knowledge that recent items were in fact recent. Of greater importance, there was no main

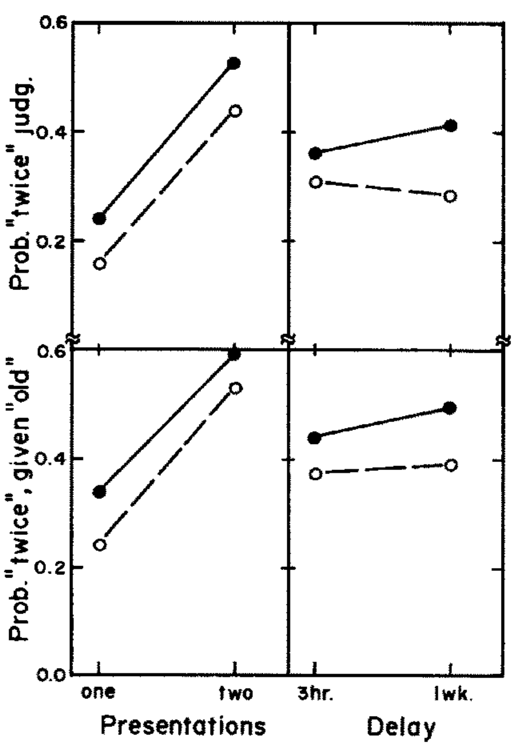

Figure 3. Mean probabilities of "twice" judgments to identical-copy items (filled circles) and left–right reversals (empty circles) by repetition condition (left panels) and retention interval (right panels). The bottom panels show the mean conditional probabilities of "twice" judgments, given recognition of items as "old."

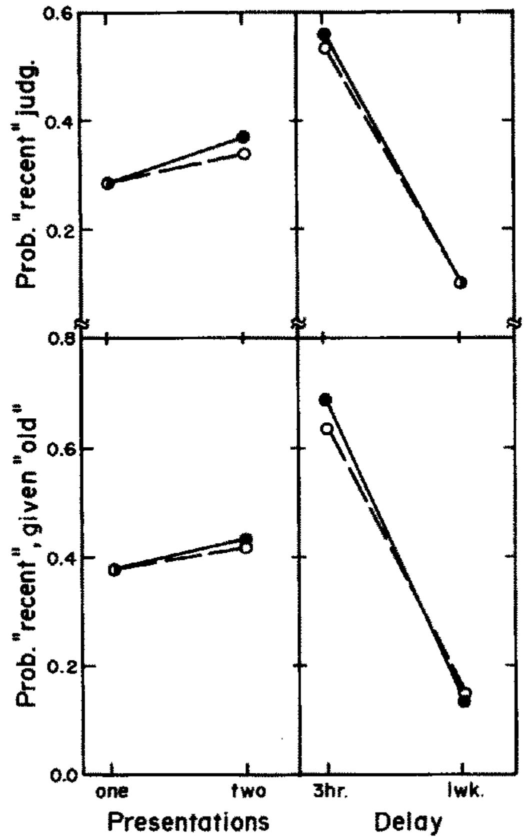

Figure 4. Mean probabilities of "recent" ("today") judgments to identical-copy items (filled circles) and left–right reversals (empty circles) by repetition condition (left panels) and retention interval (right panels). The lower panels show the mean conditional probabilities of "recent" ("today") judgments, given recognition of items as "old."

effect for orientation ($F < 1$), nor any interactions involving that variable. Thus, although orientation affected frequency judgments, it did not affect recency judgments. An ANOVA of the conditional probabilities (Figure 4, bottom panel) revealed a similar pattern of findings, except the repetition main effect (but not the Repetition × Retention Interval interaction) was no longer reliable.

We note in closing that the mean probability of "twice" judgments to lures was .08, and that of "today" judgments to lures was .09. The corresponding probabilities of "once" and "last week" judgments were .27 and .26, respectively. These data suggest biases against "twice" and "today" judgments favoring "once" and "last week" judgments, although the cause of these biases is unknown.

General Discussion

Two observations from previous picture-memory studies were replicated by our findings. First, there was a small but reliable difference in "old" judgments to identical versus reversed items. Although this small effect has not always been detectable, (e.g., Dallett et al., 1968; Standing et al., 1970), its presence here as well as in prior work (see especially McKelvie, 1983) suggests it is real albeit subject to unknown boundary conditions. Second, we found an asymmetry in orientation classification: Classifying identical items as "same orientation" was substantially more probable than classifying reversals as "different."

Given a small effect of orientation on "old" judgments and a marked asymmetry in orientation classification, we can consider what these phenomena imply about picture memory. Although the asymmetric classification phenomenon has been attributed to response bias (Klatzky & Forrest, 1984), we found this phenomenon among a subset of our subjects who showed no response bias in their classification of lures (Table 2). Furthermore, two pilot studies for the present experiment showed the asymmetric classification phenomenon without any evidence for response bias in subjects' responses to lures. Finally, it can be seen in Figure 2 (middle panels) that the magnitude of the identical–reversed difference was not constant, but was greater for twice-presented items than for once-presented items, and for items presented 1 week before than for more recently presented items. No simple version of a response bias hypothesis would have predicted such a pattern.

If response bias does not explain the asymmetry in orientation-classification, what might be a better account? We argued earlier that this asymmetry, along with the small effect of orientation on "old" judgments, implies a dual-process account that embodies three claims. First, global familiarity contributes to picture recognition and is only slightly higher for identical than reversed items. Second, because mean familiarity of identical and reversed items differs only slightly, classification of items by orientation is improved by image sampling: that is, conscious recovery of attributes of a previously studied picture, followed by analytical comparison of those recovered attributes to attributes of the test picture. Third, when image sampling fails, high familiarity is taken as evidence for "same orientation."

The dual-process hypothesis supported the general expectation that manipulating a given independent variable can produce an image-sampling pattern, a familiarity pattern, or an image-sampling-plus-familiarity pattern (Table 1). On the basis of results of a pilot investigation, we developed the hypothesis to make the specific prediction that a 1-week retention interval—compared to a 3-hr retention interval—will produce the image-sampling pattern. Additionally, prior theorizing predicted that repetition of input items—compared to single-presentation of input items—will produce either the familiarity pattern or the image-sampling-plus-familiarity pattern.

The data supported both predictions. Retention interval reliably affected the probability of "different" classifications of reversed items, but not the probability of "same" classifications of identical items or the probabilities of "old" judgments to identical and reversed items (the image-sampling pattern). In direct contrast, repetition did not reliably affect the probability of "different" classifications of reversed items, but it did affect the (raw) probability of "same" classifications of identical items, as well as the probabilities of "old" judgments to identical and reversed items (the familiarity pattern). In sum, the data were consistent with the dual-process hypothesis (Figure 1) with the added assumptions that (a) retention interval (between 3 hr and 1 week) affects primarily image sampling, whereas (b) repetition (one vs. two presentations) affects primarily familiarity of old items.

An attractive feature of the dual-process hypothesis is that it is easily elaborated to provide an account of the differing patterns of "twice" and "today" judgments shown in Figures 3 and 4. Within the dual-process framework, the finding that identical test items received more "twice" responses than reversed test items (see also Hintzman & Stern, 1977) suggests that judgments of frequency were based at least partly on perceived familiarity, resonance, or some other correlate of overall cue–trace match or similarity. One possible suggestion—allied in spirit with prior multitrace theories (Hintzman, 1976)—is that subjects can judge a picture's input frequency by estimating the number of traces that resonate to the picture's presentation beyond some preset criterion. Reversing a test picture might reduce the probability that resonance produced by one or more traces exceeds this preset criterion.

The finding that test-item orientation did not affect "today" responses suggests that recency judgments, unlike frequency judgments, were unaffected by perceived familiarity or resonance. Instead, the dual-process hypothesis suggests that recency judgments can be based on (a) recovery of temporal attributes from memory traces via the image-sampling process or (b) the ease of image sampling or the vividness of the sampled images, either of which might be correlated with recency and therefore serve as a cue for recency. Of course, evaluating these notions is a matter for future research.

The effects of retention interval require some additional comment. These effects might initially seem to contradict studies by Dallett et al. (1968), Shepard (1967), and Standing et al. (1970), all of which showed effects of delay on old–new recognition of pictures. However, both Shepard and Standing et al. used a forced-choice test of picture recognition, and though Dallett et al. used a free-choice test, they reported only overall frequencies of recognition errors, collapsing over old and new pictures. Therefore, what these prior studies showed is not that retention interval affects familiarity of old pictures, but rather that it affects discrimination of old from new pictures. In at least some cases, effects of retention interval on old–new discrimination might pertain to rejectability of lures as "new" (perhaps reflecting effects of image sampling) instead of recognizability of targets as "old."

We currently know of only four studies that have examined effects of retention interval on old pictures and new pictures separately. One of these studies (Laughery, Fessler, Lenorvitz, & Yoblick, 1974) found no main effect of retention interval on either hits or false alarms. However, this study was a test of face recognition that was procedurally different from the present investigation (e.g., an input "list" of just one face was followed by a lengthy test). The remaining three studies (Deffenbacher, Carr, & Leu, 1981; Nickerson, 1965, 1968) all showed effects of retention interval, but the test delays were either all rather short (ranging from about 3 to 17 min, Nickerson, 1965), or covered a greater range than that used here (2 min to 2 weeks in Deffenbacher et al., 1981, and 1 to 360 days in Nickerson, 1968). Fortunately, Nickerson (1968) included two retention intervals—1 and 7 days—that are comparable to the 3-hr and 1-week intervals used here. And a comparison of just these two retention intervals shows only a very small effect on the proportion of "old" judgments to old items (an estimate from his Figure 1 puts the difference at approximately .06; there also was only a small effect on false "old" judgments to lures). In contrast, and consistent with the present findings, there was a large effect of repetition on "old" judgments to old items at these two retention intervals (a difference of about .22).

It seems likely given Nickerson's (1965, 1968) findings that recognition of old items follows a negatively accelerated, Ebbinghaus-type forgetting function, and that the retention intervals used in the present experiments span a nearly horizontal portion of this function. If shorter retention intervals are sampled (3–17 min, Nickerson, 1965), or if the manipulation of retention interval is simply made more drastic (Nickerson, 1968), effects of this variable on recognizing old items might be found. It also is probable that repetition facilitates classification of reversals under some conditions. But if this is true, we are led to conclude that recognizing "oldness" and detecting reversals have different forgetting and learning *functions*, in accordance with a dual-process view.

It is perhaps easier to find evidence for two different memory processes than to characterize accurately these processes and their functioning. Thus, we might be mistaken that the processes used in orientation retrieval are familiarity and image sampling, and, even accepting this characterization, we may be misled about the necessary conditions for image sampling to occur. It is possible that in some situations familiarity must exceed some high criterion for image sampling to be possible. In terms of Tulving's (1983, chap. 14) theorizing, image sampling might be possible only when retrieval cues interact with stored memory traces to produce "ecphoric information" of particularly high quality. Alternatively, image sampling might be used in some cases only when familiarity is intermediate in value (cf. Atkinson & Juola, 1973). Although we have found no basis in the present data to suggest such restrictions on the image-sampling process, experiments directly addressed to this matter will be necessary in the future.

Another task for the future is to develop and extend the dual-process hypothesis to deal with several different issues, most especially that of the internal representations of pictures in long-term memory. An attractive albeit speculative hypothesis is that familiarity and image sampling might be based on different memory codes. Indeed, a dual-code view is compatible with our finding that retention interval affected classification of reversals but not old judgments to identical items and reversals. This pattern might imply the existence of some type of pictorial code that (a) is useful for detecting that reversals are different but (b) is rapidly forgotten in some sense.

References

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583–612). New York: Academic Press.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Bartlett, J. C., Till, R. E., Gernsbacher, M., & Gorman, W. (1983). Age-related differences in memory for lateral orientation of pictures. *Journal of Gerontology, 38,* 439–446.

Bartlett, J. C., Till, R. E., & Levy, J. C. (1980). Retrieval characteristics of complex pictures: Effects of verbal encoding. *Journal of Verbal Learning and Verbal Behavior, 19,* 430–449.

Corballis, M. C., & Beale, I. L. (1976). *The psychology of left and right.* New York: Halstead.

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

Deffenbacher, K. A., Carr, T. H., & Leu, J. R. (1981). Memory for words, pictures, and faces: Retroactive interference, forgetting and reminiscence. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 299–305.

Farrell, W. S. (1979). Coding left and right. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 42–51.

Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology, 17,* 324–363.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–68.

Hintzman, D. L. (1976). Repetition and memory. *The Psychology of Learning and Motivation, 10,* 47–88.

Hintzman, D. L., & Stern, L. D. (1977). Spacing, mirror-image repetition, and memory for pictures. *Bulletin of the Psychonomic Society, 10,* 321–324.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 1–12.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition, 13,* 289–303.

Klatzky, R. L., & Forrest, F. H. (1984). Recognizing familiar and unfamiliar faces. *Memory & Cognition, 12,* 60–70.

Laughery, K. R., Fessler, P. K., Lenorvitz, D. R., & Yoblick, D. A. (1974). Time delay and similarity effects in facial recognition. *Journal of Applied Psychology, 59,* 490–496.

Loftus, G. R., & Bell, S. M. (1975). Two types of information in picture memory. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 103–113.

Madigan, S. (1983). Picture memory. In J. C. Yuille (Ed.), *Imagery, memory and cognition: Essays in honor of Allan Paivio* (pp. 65–89). Hillsdale, NJ: Erlbaum.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

McKelvie, S. (1983). Effects of lateral reversal on recognition memory for photographs of faces. *British Journal of Psychology, 74,* 391–407.

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

Nickerson, R. S. (1968). A note on long-term recognition memory for pictorial material. *Psychonomic Science, 11,* 58.

Nickerson, R. S., & Adams, M. J. (1979). Long-term memory for a common object. *Cognitive Psychology, 11,* 287–307.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

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

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science, 19,* 73–74.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Received May 11, 1986
Revision received May 23, 1986 ■