Pictorial and Conceptual Representation of Glimpsed Pictures

Mary C. Potter, Adrian Staub, and Daniel H. O'Connor
Massachusetts Institute of Technology

Pictures seen in a rapid sequence are remembered briefly, but most are forgotten within a few seconds (M. C. Potter, A. Staub, J. Rado, & D. H. O'Connor, 2002). The authors investigated the pictorial and conceptual components of this fleeting memory by presenting 5 pictured scenes and immediately testing recognition of verbal titles (e.g., *people at a table*) or recognition of the pictures themselves. Recognition declined during testing, but initial performance was higher and the decline steeper when pictures were tested. A final experiment included test decoy pictures that were conceptually similar to but visually distinct from the original pictures. *Yes*es to decoys were higher than *yes*es to other distractors. Fleeting memory for glimpsed pictures has a strong conceptual component (conceptual short-term memory), but there is additional highly volatile pictorial memory (pictorial short-term memory) that is not tapped by a gist title or decoy picture.

Long-term memory for pictured scenes viewed for a few seconds is remarkably good (Nickerson, 1965; Shepard, 1967; Standing, 1973). However, when such scenes are presented in a rapid sequence for durations of 125–333 ms (in the range of normal eye fixations), most of them are forgotten by the time recognition is tested, shortly after presentation (Intraub, 1979, 1980; Potter, 1975, 1976; Potter & Levy, 1969). For example, Potter and Levy found that when each picture was shown for 167 ms (six pictures per second), only about 25% of the pictures were correctly recognized when tested a minute or two later. Was that because few of the pictures could even be recognized with such a brief presentation? Further experiments, in which viewers searched for verbally described targets (e.g., *a picnic*), showed that it was relatively easy to spot such targets presented among other pictures at the same rates that were associated with memory failure. For example, at 167 ms per picture, a verbally described target—a picture that had never been seen before—was detected on more than 70% of the trials (Potter, 1976). It was even easier to spot the target when the picture itself was shown in advance. Figure 1 shows a typical set of results, contrasting detection with later recognition memory. In a more stringent test, in which the target was characterized negatively (e.g., as *not an animal*), detection was still substantially higher than later recognition (Intraub, 1981). Because none of the pictures had been viewed previously in these experiments (except in the picture-target condition), participants' success in picking out the targets suggests that these rapidly presented pictures must have

been understood at a conceptual level at least momentarily—even though most of the nontarget pictures were soon forgotten.

If scenes can be momentarily understood in a glimpse as short as ⅙ s, but most are subsequently forgotten, how quickly does the forgetting occur? The phenomenon of change blindness suggests that forgetting of pictorial details occurs almost instantly: After an interruption of only 80 ms between views, participants had difficulty detecting a feature change in a picture (Rensink, O'Regan, & Clark, 1997, 2000). Even after an 8-s preview of a picture (ample time to encode the picture), a change was difficult to detect. However, information about the prechange state may have been encoded but not immediately compared with the postchange state (Hollingworth, 2003a; Simons, Chabris, Schnur, & Levin, 2002), so the extent of forgetting in a typical change blindness experiment is unclear. In change blindness studies, most of the picture remains unchanged, and the gist of the scene is assumed to be retained. In contrast, when a sequence of different pictures is presented rapidly, the whole picture is likely to be forgotten, not just details.

To find out just how quickly pictures are forgotten, Potter, Staub, Rado, and O'Connor (2002) presented sequences of five pictures (plus a final masking picture that was regarded as a filler) at a rate of 173 ms (about six pictures per second). An immediate recognition test followed, consisting of the five old pictures randomly mixed with five new pictures—distractors. Figure 2 shows the results as a function of relative serial position in the test sequence. There was a striking drop in performance over the course of the test,[1] with a suggestion of an asymptote after three or four old pictures and three or four distractors had been tested.[2]

It is surprising that, although there was a marked effect of serial position in testing, there was no corresponding evidence for a

Mary C. Potter, Adrian Staub, and Daniel H. O'Connor, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology.

Adrian Staub is now at the Department of Psychology, University of Massachusetts. Daniel H. O'Connor is now at the Department of Psychology, Princeton University.

Correspondence concerning this article should be addressed to Mary C. Potter, NE20-453, Massachusetts Institute of Technology, Cambridge, MA 02139. E-mail: mpotter@mit.edu

[1] The drop in performance over the test was not accounted for by the higher false-alarm rate (false *yes*es to distractors) early in the test, because the effect persisted after correction for guessing (Y_{corr}; see *Data analyses* section of Experiment 1), as shown in Figure 2.

[2] Although there is only a suggestion of an asymptote in Figure 2, most of our experiments confirm a leveling off after three or four old test items and a corresponding number of distractors.
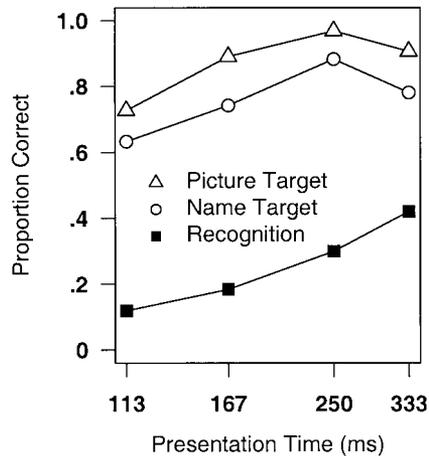
Figure 1. Proportions of correct detections of specified pictures compared with recognition memory for pictures. In the detection task, targets were specified by showing the picture in advance or by giving a short verbal title. Recognition memory was corrected for guessing. Presentation time is on a log scale. Adapted from "Short-Term Conceptual Memory for Pictures," by M. C. Potter, 1976, *Journal of Experimental Psychology: Human Learning and Memory, 2,* Figure 1, p. 511. Copyright 1976 by the American Psychological Association.

recency effect in presentation. The only serial position effect was higher recall for the first picture presented—a primacy effect that we replicated in most of the experiments in the present study. This primacy effect was probably found because attention to the first presented picture is immediate, whereas attention to later pictures requires a switch from processing the previous picture. The flat serial position effect for the remaining pictures (excluding the final filler picture) suggests that all the masking interference from a following picture occurs immediately, and additional presentation pictures have no further interfering effect, at least with presentation sequences of up to 20 items (Potter et al., 2002).

Potter et al.'s (2002) next question was whether the rapid drop in memory over 8 s of testing was due to decay or interference. The answer was both. An unfilled delay of 5 s before the beginning of the test resulted in a drop in recognition, but not as great a drop as after 5 s of testing (Potter et al., 2002). In sum, rapidly presented pictures remain in memory for a short time, even when they are slated to be forgotten within a few seconds.

The present study addressed this question: What kind of information is represented in short-term memory for scenes? We considered two possible forms of short-term memory: visual, pictorial short-term memory (PSTM) and conceptual short-term memory (CSTM; Potter, 1993, 1999). We hypothesized that the high level of performance seen early in testing reflects both PSTM and CSTM. We proposed that PSTM is lost rapidly, leaving CSTM as the main basis of short-term recognition later in testing. By the end of about 8 s of testing, both CSTM and PSTM have been lost, and only the subset of pictures that have been consolidated in long-term memory (LTM) can still be recognized. To test this hypothesis about the time course of PSTM and CSTM, in the present study we contrasted pictorial and conceptual recognition tests of picture memory. The pictorial test was the usual yes–no recognition test of old pictures mixed with new (distractor) pictures,

whereas the conceptual test required participants to use descriptive titles as cues to remember the pictures they had just seen while rejecting titles of nonpresented pictures (distractors). None of the presented pictures (or the descriptive titles used in testing) had been seen previously. Whereas a test picture provides both visual–pictorial and conceptual information, a descriptive title provides only conceptual information (none of the titles named visual features per se, except indirectly via participants' knowledge of what objects and scenes look like).

We expected that recognition of test pictures would be more accurate than recognition of test titles, inasmuch as pictures provide full information whereas descriptive titles provide only abstracted picture gist. Our primary question was whether this expected main effect would interact with serial position in the test. If superior performance on the first few tested items (see Figure 2) is attributable to visual features that are quickly forgotten, then we would expect tests using titles to have a different profile, starting lower and showing less of an initial decline.

In a final experiment, a different method was used to distinguish between visual and conceptual characteristics of remembered pictures. One old picture in each test sequence was replaced by a conceptually similar picture that was visually distinct from the target. If participants make more false *yes* responses to such decoys than to other distractors, that would suggest that they retained conceptual information about the original picture but not full perceptual information. Again, our question was whether, if there were a decoy effect, it would interact with serial position on the recognition test: Specifically, if participants initially retain visual information but quickly forget it as the test continues, then the decoy effect should be reduced in early test positions.

In Experiment 1, using two groups of subjects, we established a baseline for performance on the title test relative to the standard
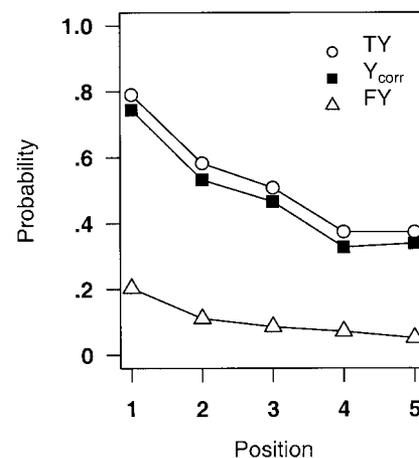
Figure 2. Probability of recognizing an old picture (TY), falsely recognizing a new picture (FY), and recognizing an old picture corrected for guessing (Y_{corr}) as a function of relative serial position in the recognition test in Potter et al. (2002, Experiment 1). Adapted from "Recognition Memory for Briefly Presented Pictures: The Time Course of Rapid Forgetting," by M. C. Potter, A. Staub, J. Rado, and D. H. O'Connor, 2002, *Journal of Experimental Psychology: Human Perception and Performance, 28,* Figure 2B, p. 1166. Copyright 2002 by the American Psychological Association.

picture recognition test under conditions in which most of the pictures are likely to be in LTM because they are shown for 1 s each. Experiment 2 replicated Experiment 1 using a higher rate of presentation (173 ms per picture), at which there is likely to be short-term memory for most of the pictures but LTM for only a fraction (Potter, 1976; Potter et al., 2002). Experiment 3 combined Experiments 1 and 2 by showing pictures for 173 ms plus a blank interstimulus interval (ISI) totaling 1 s per picture to test the hypothesis that consolidation time, not viewing time, accounts for the advantage of a 1-s presentation. In Experiment 4, recognition tests with pictures and titles were randomly mixed within the same experiment to rule out possible strategy effects. In Experiment 5, with only pictures as test items, one of the old pictures in each test sequence was replaced by a conceptually similar but visually distinct decoy picture. The general method used in all of the experiments is described in detail in Experiment 1.

Experiment 1

The purpose of Experiment 1 was to assess whether memory for a pictured scene can be successfully tapped by recognition testing with a verbal title rather than the picture itself. In previous research, memory for line drawings of single objects has been tested by presenting the object's name in a recognition test, with recognition results that are similar to although often somewhat lower than those for recognition tests using the pictures themselves (e.g., K. O'Connor & Potter, 2002; Snodgrass & McClure, 1975). The to-be-remembered pictures in those studies consisted of single, readily recognized objects with conventional basic-level names, whereas the scenes in the present study were complex, were being viewed for the first time, and did not have conventional names. It was unclear to what extent a recognition test using titles of such pictures would be a valid index of memory. Therefore, in the first experiment, we compared pictures and verbal titles as recognition-memory cues after presenting the pictures for 1 s each, a rate at which most pictures are likely to be in LTM (Potter & Levy, 1969).

Method

Participants. Two groups of 10 volunteers from the Massachusetts Institute of Technology community were paid for their participation. All reported normal or corrected vision. No participants took part in more than one of the experiments in the present study, and none had been in earlier studies with pictures in this laboratory.

Materials and apparatus. The pictures were 660 color photographs with widely varied content, chosen from commercially available compact discs. They included pictures of animals, people engaged in various activities, nature scenes, and city scenes. They were presented on an Apple PowerPC 7500/100 computer with a 17-in. (43.18-cm) monitor set at a resolution of 832 × 624 pixels and a refresh rate of 75 Hz, using MacProbe software (Hunt, 1994). The pictures were 300 × 200 pixels as displayed (10.6 × 7.1 cm), subtending approximately 13° of visual angle horizontally and 9° vertically when viewed from the normal distance of 45 cm. Pictures were shown against a medium gray background, which was present throughout. The room was dimly illuminated.

For the test using descriptive titles, two research assistants, working together, constructed a short (one- to seven-word) descriptive title for each picture in the experiment, both distractors and presented pictures. The titles were generated without knowledge of which pictures were presented pictures and which were distractors. Each title was intended to capture the

central theme or gist of the picture as concisely as possible. Explicit references to visual features, such as color and shape, were prohibited. Examples include *pond with trees*, *child near shelves*, *avocado*, *people standing in a row*, *hands working on electrical board*, *football game*, *short wall with plants*, and *fox*. In the recognition test, the titles were displayed in Palatino bold 20-point font in the center of the screen.

Design and procedure. There were 60 experimental trials in each group. For the group tested with pictures, each trial consisted of 5 presentation pictures plus a 6th picture that served as a visual and conceptual mask and 10 test pictures consisting of the 5 presentation pictures and 5 distractors. Pictures were assigned randomly (without replacement) from the full set to each trial and to the role of presentation picture or distractor in that trial. For the title condition, it was necessary to switch 20 of the 660 pictures between trials to avoid having conceptually similar pictures (and, hence, very similar or identical titles) on the same trial, whether as presentation pictures or distractors. The rearrangement was made after the group tested with pictures had already completed the experiment; in all subsequent experiments, the rearranged sequences were used.

The serial positions of the 5 presentation pictures were counterbalanced across subjects so that each picture appeared in each serial position. Half of the 10-item recognition tests (30 trials) began with an old picture, half with a distractor. Of the 30 trials beginning with an old picture, that picture was drawn equally often from each of the 5 serial positions in the presentation sequence. The remaining old and new test pictures were randomly ordered on each trial, with the constraint that no more than 2 old or 2 new pictures appeared consecutively. The order of the old and new pictures in each test trial was reversed for half of the subjects, counterbalanced with serial position in presentation: The 1st old picture tested became the 5th tested, the 2nd became the 4th, and so on; the distractors were handled in the same manner.

Before each trial, the word *ready* appeared on the screen, indicating that the participant could press the space bar to begin the trial. The sequence began with a red fixation cross in a black rectangle framed by red edging that was the same size as the pictures. The fixation array was presented for 293 ms, followed by a blank of 200 ms and six pictures in sequence for 1 s each. The sixth picture was regarded as a filler and was not tested for recognition. Immediately following the last picture, a white rectangle slightly larger than the pictures was presented for 293 ms, followed by the first test picture. (The purpose of the white rectangle was to signal the end of the presentation sequence and the beginning of the test sequence.)

Each test picture was presented for 400 ms, followed by a blank screen until the participant responded by pressing a *yes* or a *no* key on the keyboard. The reason for the relatively short duration of the test picture was to encourage the participant to make a rapid decision and to minimize the elapsed time during the test. The participant's response was followed by a 107-ms delay before the next test picture appeared, and the cycle repeated for the 10 test pictures. Participants were instructed to view the presentation sequence and then respond to the test pictures as rapidly as they could, consistent with accuracy. They were told that they did not need to be absolutely sure that a picture was an old picture in order to press *yes*. They were not given any information about how many of the test pictures in each trial were old pictures. Between trials, there was a brief interval while the computer loaded the pictures for the next trial, after which the word *ready* appeared on the screen. There were four practice trials, each using a different set of pictures. No picture was repeated in the experiment, except as an old test picture.

The procedure for the group tested with titles of the pictures was identical to that for the picture group, except that titles replaced the pictures in the recognition test. The titles appeared for 800 ms, rather than 400 ms, to allow participants sufficient time to read them. Participants were instructed to press *yes* if a title corresponded to a picture that had appeared in the presentation sequence and otherwise to press *no*.

Data analyses. In this and the following experiments, we analyzed the data using three measures, calculated separately for each subject in each

cell of the design. The first measure was the proportion of *yes* responses, separately for true *yes*es (TYs) and false *yes*es (FYs). This analysis allowed us to look at serial position in both presentation and test. Second, we used a high-threshold guessing correction that is often used in studies of picture memory: $P(Y_{corr}) = [P(TY) - P(FY)]/[1 - P(FY)]$, where P is proportion.

Third, we used A', which is an alternative to d' that can be calculated even when the false-alarm rate is zero for a given participant in a given condition (Donaldson, 1993; Grier, 1971; Macmillan & Creelman, 1991; Pollack & Norman, 1964; Snodgrass & Corwin, 1988). Because the results from the Y_{corr} analyses and the A' analyses were in most cases similar, we report only $P(Y_{corr})$, except when only one of the two analyses was significant at the .05 level or better; in those cases, we specify which analysis was significant and which was not.[3] Because we were particularly interested in serial position effects in testing, we calculated $P(Y_{corr})$ for each subject at each of the five relative serial positions of the old versus the new pictures in the test. For analyses of serial position effects in the presentation sequence, uncorrected $P(TY)$ was used (unlike test serial position, there is no measure of false *yes*es that corresponds to each serial position in presentation).[4]

Results and Discussion

As expected, the group given the picture recognition test performed very well: Overall, they correctly recognized 89% of the old pictures and falsely recognized only 4% of the distractors. The group tested with titles did less well, although their performance was surprisingly good: Overall, 75% of the titles of old pictures were correctly recognized and 14% of distractor titles were falsely recognized. Figure 3 shows $P(Y_{corr})$ for the two groups as a function of serial position in the test. An analysis of variance (ANOVA) revealed a significant difference between the groups, $F(1, 18) = 13.82, p < .01$, and a significant effect of test position, $F(4, 72) = 4.93, p < .01$, with no interaction ($F < 1$). Trend analyses were carried out separately for pictures and titles. The linear trend was significant for both groups: pictures, $F(1, 9) = 11.89, p < .01$; titles, $F(1, 9) = 5.93, p < .05$.

An analysis was carried out on correct *yes*es in order to evaluate the effects of serial position in the presentation sequence. Both serial position in presentation and in testing were variables, as was test type. As before, there was a significant main effect of pictures versus titles and of test position, but no main effect of presentation position. The only significant interaction was Test Type × Presentation Position, $F(4, 72) = 3.24, p < .05$. Inspection indicated that there was a first-item primacy advantage for pictures only, whereas for titles there was a slight recency effect. Neither effect was significant, however.
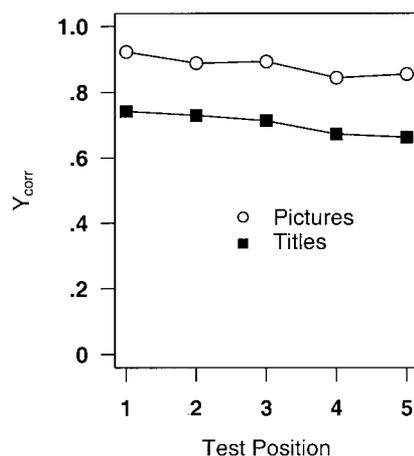
In sum, when a sequence of 6 pictures is displayed at a rate of 1 s per picture, most test pictures are presumably consolidated in LTM so that they are readily recognized, and there is little or no forgetting during the test period. Corrected for false *yes*es, 88% of the pictures were remembered. These results are consistent with previous findings demonstrating a high level of recognition when the duration of each picture is 1 s or more; Potter and Levy (1969) found that about 82% of the pictures in their study were subsequently recognized when presented at a rate of 1 per second, with little memory loss over a 32-item recognition test (see also Hollingworth, 2003b; Intraub, 1980). In the title condition in the present experiment, the overall level of performance was lower than for the picture condition. After correction for false *yes*es, 70% of the titles successfully cued recall of the corresponding picture.

These results suggest that when a series of pictures is presented at the rate of 1 per second, participants are not only able to remember the visual features of these pictures, but they are also able to encode and remember the meanings of the pictures in some detail. Both perceptual and conceptual information is retained for at least 14 s (the approximate duration of the presentation and test periods), and probably for considerably longer. Titles do not cue memory for the pictures as successfully as the pictures themselves because titles provide only gist, not the full pictorial and conceptual detail that is encoded and consolidated in 1 s. However, titles clearly capture some important aspects of the information in memory.

Experiment 2

Experiment 2 tested the hypothesis that the fleeting memory for rapidly presented pictures observed by Potter et al. (2002; see the present Figure 2) depends on visual–pictorial rather than solely on conceptual information. Having shown in Experiment 1 that descriptive titles can tap LTM for pictures, in Experiment 2 we used the same contrast between picture and title tests to examine fleeting memory—PSTM or CSTM—when pictures are presented at a high rate, almost six pictures per second (173 ms per picture).

Figure 3. Probability of recognizing an old picture corrected for guessing (Y_{corr}) as a function of relative serial position in the recognition test, separately for the group given a picture recognition test and the group given a title test, in Experiment 1. Presentation rate was 1 picture per second.

---

[3] The main reason for reporting the Y_{corr} results rather than A' is that the former measure is transparently related to the raw $P(TY)$ and $P(FY)$ results. The second reason is that previous studies of picture memory (Potter, 1974) have indicated that the variances of the distributions of the familiarity of old and new pictures differ, such that the variance is much greater for old pictures, suggesting that measures like A' and the two-high-threshold correction that assume symmetry between these distributions are inappropriate.

[4] The design, procedure, materials, and analyses used in the picture test condition of Experiment 1 were like those in Experiment 1 in Potter et al. (2002), although the presentation duration was 1 s rather than the 173-ms duration in the previous study.

Method

The method was like that of Experiment 1, except that the sequence was presented for 173 ms per picture. Because of a programming error, the test pictures in the picture group were presented to 4 subjects for 800 ms each, rather than 400 ms; these subjects performed similarly to other subjects. There were 20 participants in each group.

Results and Discussion

As expected, performance was much worse in Experiment 2 than in Experiment 1. As in Experiment 1, the group tested with pictures did better overall than the group tested with titles. Both groups, however, showed a decline in performance over the test, as shown in Figure 4. It is important to note that the benefit of testing early in the sequence was greater for the group tested with pictures, showing that features of the presented picture not represented in the title—presumably visual features—were available early in testing but were quickly lost.

In the group tested with pictures, 51% of the old pictures were recognized and 14% of the distractors were falsely recognized; in the group tested with picture titles, 48% and 18% were recognized, respectively. The guessing-corrected scores, $P(Y_{corr})$, are shown in Figure 4 as a function of test serial position. In the analysis of $P(Y_{corr})$, the picture group was more accurate than the title group, $F(1, 38) = 8.20$, $p < .01$; the effect of test position was significant, $F(4, 152) = 75.41$, $p < .001$; and the Group × Test Position interaction was also significant, $F(4, 152) = 6.68$, $p < .001$. Separate analyses of the two groups showed a significant effect of serial position in each group: For the picture group, $F(4, 76) = 86.66$, $p < .001$, and a trend analysis showed significant linear and quadratic components (both $ps < .001$); for the title group, $F(4, 76) = 16.82$, $p < .001$, and there was a significant linear trend ($p < .001$), but the quadratic trend was not significant ($p = .08$).

In an analysis of the *yes* responses to old pictures in which both presentation serial position and test serial position were variables, the main effect of test serial position and the Test Serial Position ×
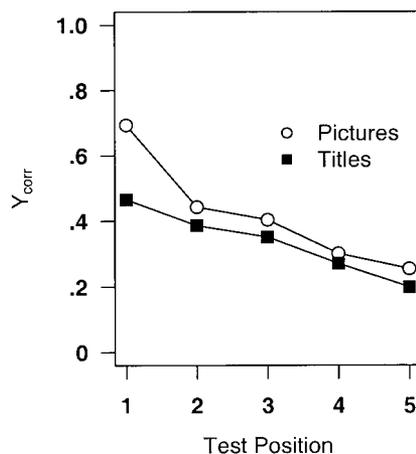
Test Group interactions were highly significant, as in the analysis of $P(Y_{corr})$. The serial position of presentation had a significant effect, $F(4, 152) = 19.33$, $p < .001$, consisting of a primacy benefit for the first presented picture. There was no recency effect (but note that the last picture in presentation was treated as a filler). There were no Presentation Serial Position × Test Group (pictures vs. titles), Presentation Serial Position × Test Serial Position, or Presentation Serial Position × Test Serial Position × Test Group interactions (all $Fs < 1$).[5]

The performance with pictures versus titles supports the hypothesis that the fleeting early benefit in the recognition test is primarily due to short-lived visual–pictorial information—PSTM—that is not tapped by a title. The relative similarity of performance with pictures and titles in the later part of the test suggests that the more durable short-term memory for a picture is primarily conceptual when pictures are presented at a rate of about six per second. Test pictures and titles tap the same type of information, except right at the beginning of the test, when the picture test evidently taps considerably more information that is quickly lost if testing comes a little later. However, the finding that performance showed a significant decline throughout testing, not only in the picture test but also in the title test, suggests that conceptual information—CSTM—is also being forgotten over the approximately 8 s of testing, although at a slower rate than PSTM.

The results of the group tested with titles show that information about the pictures' semantic contents was often remembered well enough for participants to match verbal titles to these memories, even when the pictures were presented for only 173 ms each. Indeed, a comparison of the two groups in Experiment 2 (Figure 4) indicates that, apart from the beginning of the test, titles were almost as recognizable as pictures. This result is surprising, and it is strong support for the claim that conceptual gist is extracted very early in the viewing of a picture. Potter (1975, 1976) and Intraub (1981) showed that participants can readily pick out a picture target when given a descriptive title in advance of viewing a rapid sequence of pictures—although they do still better when shown the target picture itself in advance (see Figure 1). However, this is the first time that the reverse has been shown: that a title is an effective retrieval cue for a previously glimpsed picture.

Not surprisingly, performance was worse overall in Experiment 2 (with pictures presented for 173 ms) than in Experiment 1 (with pictures presented for 1 s). It is interesting to note, however, that at the higher rate of presentation, results with titles were more nearly equivalent to those of the picture tests.[6] This outcome is what one might expect if gist information (corresponding to the title) is encoded early in the processing of a picture (e.g., in the first 173 ms), whereas more detailed visual or conceptual infor-

Figure 4. Probability of recognizing an old picture corrected for guessing (Y_{corr}) as a function of relative serial position in the recognition test, separately for the group given a picture recognition test and the group given a title test, in Experiment 2. Presentation rate was 5.8 pictures per second (173 ms per picture).

---

[5] The overall pattern of results for the picture group—rapid forgetting of some of the pictures during the test, in particular in the early part of the test—replicated the results of a similar experiment in Potter et al. (2002; see Figure 2 in the present article).

[6] An ANOVA of $P(Y_{corr})$ combining Experiments 1 and 2 showed a highly significant difference between the experiments and significant main effects of type of test and test serial position, as well as significant interactions. It is important to note that the Experiment × Test Type × Test Serial Position interaction was also significant ($p < .05$), consistent with the conclusion that title and picture tests were more nearly equivalent at the high rate, except at the beginning of the test.

mation is acquired later. In Experiment 2, relatively little extra detail could be encoded and consolidated in the 173 ms of presentation, reducing the difference between picture and title tests.[7]

In one important respect, however, the results support the assumption that picture details (beyond gist) play a significant role in the recognition test, even with a 173-ms presentation (Experiment 2). Performance in the picture test was markedly better than that in the title test in the early part of the recognition test, specifically for the first old picture tested. This outcome is most readily explained by assuming that nongist pictorial information (PSTM) was available if the picture was tested early. Later in the test, however, there was a reliance on gist information (CSTM), putting titles and pictures more nearly on par. Because both kinds of information— PSTM and CSTM—had ample time to be consolidated into LTM at the slow rate of presentation in Experiment 1, a test picture conveyed more useful information than a title, and this difference was stable throughout the test.

Experiment 3

One question about the comparison between Experiments 1 and 2 is whether the benefit from 1 s rather than 173 ms of viewing comes from the extra time to inspect the pictures or from extra time to consolidate the information. Intraub (1980) presented pictures sequentially for 110 ms, with blank ISIs ranging from 0 to 5.9 s; memory performance improved markedly as the ISI increased from 20% with a 0-s ISI to 84% with an ISI of 1.5 s (see also Potter, 1976, Experiment 3). This finding suggests that the main benefit of a long presentation duration is to allow consolidation of information extracted early in viewing. To test this hypothesis, in Experiment 3, we presented each picture for 173 ms, followed by a blank ISI of 827 ms, so the stimulus onset asynchrony (SOA) was 1 s, as in Experiment 1, but the presentation duration of each picture was 173 ms, as in Experiment 2.[8]

Method

The method was like that of Experiments 1 and 2, except that each picture in the sequence was presented for 173 ms, followed by a blank ISI of 827 ms, for a total SOA of 1 s. The blank ISI was a medium gray, the same as the background. There were 10 participants in each group.

Results and Discussion

The results of Experiment 3, corrected for guessing, are shown in Figure 5 as a function of test serial position. Despite the brief exposure duration of each picture, performance was not much different from that in Experiment 1, in which pictures were viewed for a full 1 s (Figure 3), and was much better than in Experiment 2, in which pictures followed each other with no ISI (Figure 4). As in Experiment 1, the group tested with pictures did better, overall, than the group tested with titles. Only the picture group showed a significant (although small) decline in performance over the test, as shown in Figure 4.

In the group tested with pictures, 90% of the old pictures were recognized and 7% of the distractors were falsely recognized; in the group tested with picture titles, 72% and 16% were recognized, respectively. In the analysis of $P(Y_{corr})$, the picture group was more accurate than the title group, $F(1, 18) = 50.23$, $p < .001$; the effect of test position was significant, $F(4, 72) = 2.89$, $p < .05$;[9]
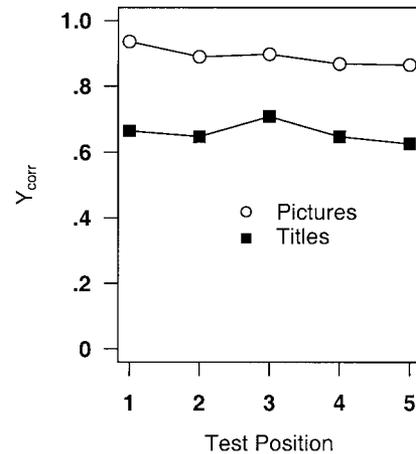
Figure 5. Probability of recognizing an old picture corrected for guessing (Y_{corr}) as a function of relative serial position in the recognition test, separately for the group given a picture recognition test and the group given a title test, in Experiment 3. Presentation duration was 173 ms followed by a blank interstimulus interval of 827 ms, for an overall rate of 1 picture per second.

but the Group × Test Position interaction was not significant ($F < 1$). Separate planned analyses of the two groups showed a significant effect of test position in the picture group, $F(4, 36) = 3.24$, $p < .05$;[10] a linear trend analysis was not significant ($p < .10$). For the title group, neither test position nor linear trend was significant.

In an analysis of the *yes* responses to old pictures in which both presentation serial position and test serial position were variables, there was a significant main effect of pictures versus titles ($p < .001$) and of test position ($p < .05$), but no main effect of presentation position and no interactions. These results are consistent with those of Experiment 1.

Analyses comparing Experiments 1 and 3 were carried out separately for the picture and title groups on $P(Y_{corr})$. For the picture groups, there was no significant difference between the two experiments, the effect of test position was significant ($p < .001$), and there was no interaction. For the title groups, there was again no significant difference between the experiments ($p = .24$), test position was marginally significant ($p = .05$), and there was no interaction.

Overall, the comparison between Experiments 1 and 3 shows that when there is an SOA of 1 s between pictures, having the picture in view for only 173 ms of that interval is little different from being able to view the picture for the full 1 s and that in both

---

[7] A different possibility, contradicted by the results of Experiments 1 and 2, is that because a title provides a relatively impoverished representation of a picture, only a robust memory representation could successfully match it. On this hypothesis, a scanty picture memory provided by a short, 173-ms glimpse would provide a poor match with a title test. One would then expect performance with titles relative to pictures to be much worse in the 173-ms condition than in the 1-s condition—contrary to what we observed.

[8] We thank a reviewer for suggesting this experiment.

[9] Test position was not significant in the A' analysis.

[10] Test position was not significant in the A' analysis.

cases performance is much better than when pictures are presented for 173 ms with no ISI (Experiment 2). The opportunity in Experiment 1 to make two or three eye movements to different parts of a picture did not give a significant memory advantage. This suggests that the major benefit of a slower rate of presentation is not in the extra time to inspect the picture but the extra time to consolidate information picked up in the initial 173 ms of viewing. This conclusion is consistent with results of Intraub (1980), described earlier.

Experiment 4

In Experiment 4, we tested an alternate interpretation of Experiments 1–3. It seems possible that participants adopt two different encoding strategies when they are tested with pictures versus titles. Perhaps performance in the title condition was relatively good because participants were able strategically to encode the pictures in a verbal or propositional form, even when pictures were presented at a rate of six per second in Experiment 2. To test this hypothesis, in Experiment 4, the two types of trials were randomly intermixed within participant, so that participants did not know until the test began whether they would be seeing pictures or titles. If the nearly equivalent level of performance for Experiment 2's two groups was the result of task-specific encoding strategies, performance would be expected to decline in one, the other, or both conditions in Experiment 4. As in Experiment 2, pictures were presented for 173 ms with no ISI.

Method

The method was the same as that in Experiment 2, except that type of recognition test was a within-subject variable. Of the 60 trials, half were tested with titles and half with pictures, counterbalanced between subjects. Serial position in presentation and serial position in test were counterbalanced with type of test. Trials were randomized so that participants did not know until the recognition test began whether it would consist of pictures or titles. There were 20 participants.

Results and Discussion

Overall, the results of Experiment 4, with titles and pictures as a within-subject variable, were similar to those for Experiment 2, which had a between-subjects design. If anything, however, the advantage of a picture test early in testing was more marked in Experiment 4.

In the picture test condition, 56% of the old pictures were recognized and 18% of the distractors were falsely recognized; in the title condition, 47% and 24% were recognized, respectively. The guessing-corrected scores, $P(Y_{corr})$, are shown in Figure 6 as a function of test serial position. In the analysis of $P(Y_{corr})$, performance in the picture condition was more accurate than performance in the title condition, $F(1, 19) = 8.59$, $p < .01$; the effect of test position was significant, $F(4, 76) = 15.80$, $p < .001$; and the Condition × Test Position interaction was also significant, $F(4, 76) = 10.08$, $p < .001$. Separate analyses of the two test conditions showed a significant effect of serial position in each condition: for the picture condition, $F(4, 76) = 22.25$, $p < .001$, and a trend analysis showed significant linear ($p < .001$) and quadratic components ($p < .01$); for the title condition, $F(4, 76) =$
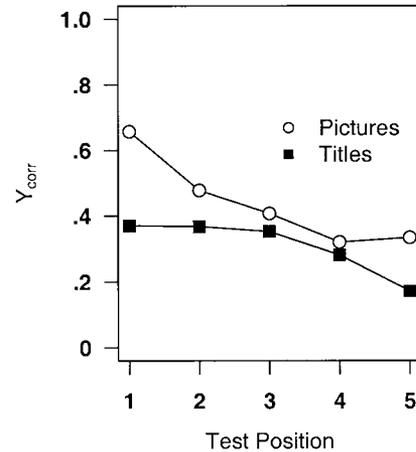
Figure 6. Probability of recognizing an old picture corrected for guessing (Y_{corr}) as a function of relative serial position in the recognition test, separately for picture recognition tests and title tests, in Experiment 4. Trials with picture tests and with title tests were intermixed within subjects. Presentation rate was 5.8 pictures per second (173 ms per picture).

7.33, $p < .001$, and there were significant linear ($p < .001$) and quadratic components ($p < .05$).

In an analysis of correct *yes*es in which test condition was crossed with serial position in presentation and in test, our main interest, as before, was in serial position in presentation and its possible interaction with test condition and with test serial position. There was a significant main effect of presentation position, $F(4, 76) = 10.85$, $p < .001$, with the usual primacy effect and no recency effect; there was no significant interaction with test position, with test condition (pictures versus titles), and no triple interaction. This pattern of results was highly similar to that in Experiment 2.

In sum, intermixing trials with picture tests and title tests changed performance only minimally in comparison with Experiment 2, in which the two test conditions were presented to different groups. It seems unlikely, on the basis of this result, that success in recognizing titles depends on strategic encoding of the pictures in verbal or propositional form. To the contrary, the evidence we have presented suggests that participants standardly encode the meanings of pictures so that a title is an effective match to that meaning in many cases. The evidence suggests, however, that in addition to a picture's overall meaning or gist, for a brief time participants maintain enough detailed visual information to make recognition of the first old test item much better when it is a picture than when it is a title.

Experiment 5

In Experiment 5, we evaluated the conceptual basis of memory in another way, by including in the recognition test occasional pictures that matched the title—the gist—of one of the old pictures. We called such a picture a *decoy*; two examples are shown in Figure 7. The decoy was visually different from the old picture it replaced: When they were side by side, it was easy to tell that the pictures were not the same. If participants rely on a visual match early in the test, then they should not be susceptible to the decoys;
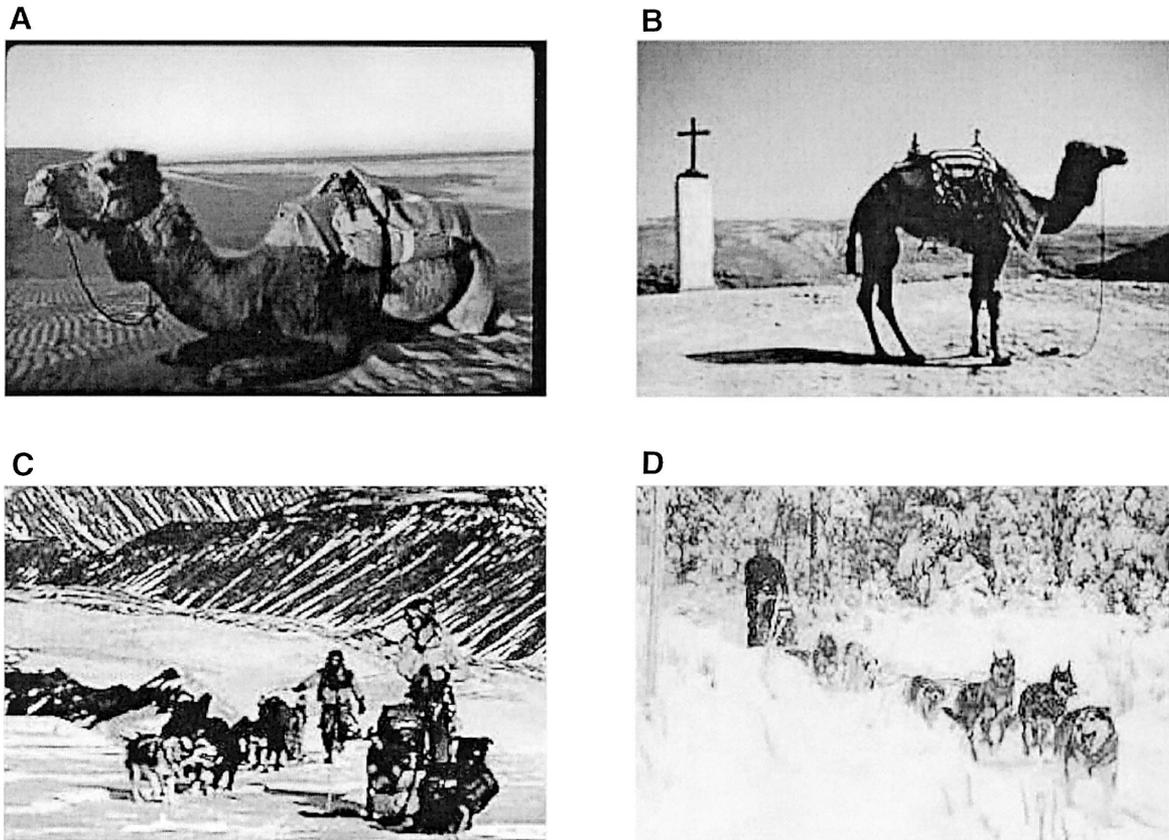
Figure 7. Examples of original pictures and decoys used in Experiment 5. A and B: *Camel*; C and D: *Dogsled*. The original picture in each case is on the left. In the experiment, the pictures were in color.

to the extent that they rely on a conceptual or gist representation of the presented pictures, however, they should make false *yes* responses more often to decoys than to new distractors that are not conceptually similar to any of the pictures they have just seen. Moreover, if participants rely in part on CSTM (as Experiments 2 and 4 suggest), then one would expect a falloff in response to decoys in the course of the test like that for responses to old pictures.

Method

There were 20 participants. The method was the same as that for Experiment 2's picture group, except for the following: A decoy picture replaced one of the five old pictures in each recognition test. The decoy was selected so as to satisfy the same short (one- to seven-word) conceptual title that was assigned to the corresponding presentation picture in Experiments 1–4; if a presentation picture was assigned the title *skiers on a mountain*, this would also be the main theme of the corresponding decoy. However, the decoy was always selected so as to be visually dissimilar to the corresponding presentation picture. A minimal criterion was that the two pictures should be sufficiently dissimilar that they would be easily discriminable if presented beside each other. The pairs varied in their degree of dissimilarity: Some did share some visual features, such as dominant colors, but most were alike at only an abstract level. Figure 7 shows example pairs of decoy and old pictures. Like the original set of pictures, the decoy pictures were gathered from various sources, including commercially available CDs and the Internet.

To test the discriminability of old pictures and decoys, we had a separate group of 7 participants make same–different judgments of sequentially presented pairs of pictures preceded and followed by visual masks consisting of other pictures cut up into small squares and reassembled randomly. On 120 of the 240 trials, the two pictures were identical. On 60 trials, the two pictures were unrelated. On the remaining 60 trials, the two pictures were the old picture and the decoy, in that order. The trials were intermixed randomly. Each trial began with a fixation cross for 307 ms, followed by a blank for 200 ms, then a mask for 173 ms, the first picture, a second mask, the second picture, and a third mask. The participant responded by pressing one of two keys to indicate *same* or *different*. For the identical pairs, 96.1% of the responses were *same*; for the unrelated pairs, 98.3% of the responses were *different*; and for the decoy pairs, 97.4% of the responses were *different*. There was no significant difference in accuracy between the decoy and related-picture pairs. Thus, participants were almost always able to distinguish decoys from old pictures when these were presented for 173 ms and separated by a 173-ms visual mask. Clearly, the two pictures were not highly similar visually. However, they were certainly more similar than two random pictures would have been, so we cannot entirely rule out the possibility that there was sufficient visual similarity to produce some level of visual (rather than solely conceptual) matching once the memory for the first picture had been degraded by test interference.

In Experiment 5, the serial positions of the decoy pictures were counterbalanced so that for each subject, the decoys appeared equally often in the place of the first-tested old picture, the second-tested old picture, and so on. The position of the presentation picture that was replaced by the

decoy was similarly counterbalanced: For each subject, the first-presented picture was replaced in the test sequence exactly as often as the second-presented, and so on. As in the previous experiments, the position of a given picture in the presentation sequence was counterbalanced between subjects. Test position was counterbalanced between subjects in the same way as in Experiment 1 and subsequent experiments.

In sum, each test sequence consisted of five new pictures (the standard distractors), four old pictures, and one decoy picture that occupied the test position vacated by one of the old pictures. Participants were not informed that there would be decoys: They were instructed as before to press *yes* only if they were reasonably confident that they had seen a given picture in the preceding sequence. Thus, the correct response to the decoys was *no*.

Results and Discussion

Participants made significantly more false *yes* responses to decoys than to other distractors, although there were substantially more true *yes*es to old pictures than false *yes*es to decoys. Thus, participants generally remembered more than just the gist of a picture, and so they were able to reject a gist-sharing decoy much of the time. It is important to note that the benefit of testing early was much greater for old pictures than for decoys (Figure 8).

Responses to old pictures, decoy pictures, and new pictures were scored separately. Overall, participants recognized old pictures 52% of the time and falsely recognized nondecoy distractors 15% of the time. They responded *yes* falsely to the decoy pictures 30% of the time.

In calculating $P(Y_{corr})$, we used the false-*yes* rate for nondecoy distractors both for *yes* responses to old pictures and for *yes* responses to decoys. For old pictures, the $P(Y_{corr})$ analysis of test position was significant, $F(4, 76) = 61.26$, $p < .001$, with a sharp drop-off in performance (as shown in Figure 8) that was similar to that in the picture test conditions in Experiments 2 and 4. A trend analysis showed significant linear and quadratic components (both ps $< .001$). For decoys, test position was also significant, $F(4, 76) = 12.21$, $p < .001$, with a drop-off that was less marked than that for old pictures. A trend analysis showed a significant linear

component ($p < .001$); the quadratic component was not significant. In an analysis of $P(Y_{corr})$ comparing old pictures and decoys, there was a significantly greater probability of saying *yes* to an old picture than to a decoy, $F(1, 19) = 268.42$, $p < .001$; there was a main effect of test position, $F(4, 76) = 50.87$, $p < .001$; and there was an interaction, $F(4, 76) = 4.41$, $p < .01$. The interaction is shown in Figure 8: The advantage of seeing an old picture (relative to a decoy) was greater when tested early than when tested later.

In an analysis of correct *yes* responses to old pictures as a function of serial positions in presentation and in test, there was, as usual, a significant primacy effect for the first picture presented, with no other serial position effects, $F(4, 76) = 5.31$, $p < .001$, and a significant effect of test position, $F(4, 76) = 78.88$, $p < .001$, with performance high early in the test and dropping rapidly thereafter. There was no interaction between these effects ($F < 1$). In the corresponding analysis of *yes* responses to decoys, the only significant effect was for test position, $F(4, 76) = 20.58$, $p < .001$. Although the presentation position of the picture corresponding to the decoy had no significant effect, the probability of responding *yes* to a decoy was higher when the corresponding picture was the first one presented, suggesting a small primacy effect.

In sum, the decoys were clearly more effective than random, conceptually unrelated distractors in producing (false) *yes*es, but they were less effective than the genuine old pictures. This result suggests either that decoy pictures are only moderately successful at accessing the stored representations underlying correct recognition or that they successfully lead to retrieval, but the retrieved representation includes enough additional information to allow the viewer to reject the decoy.

Experiment 5 shows clearly that gist information is not all that is retained from a series of briefly presented, conceptually masked pictures: Visual or conceptual detail is remembered as well. When participants are presented with a test picture that fails to capture this visual detail (or includes different details), they are fairly likely to reject it, even if it is conceptually similar to an old picture. Because conceptual as well as visual memory for the presented picture falls off during the test, false recognition of the decoy also falls. As with titles in Experiments 2 and 4, the initial falloff is more rapid for presented pictures than for decoys, showing that there is extra information about the presented picture early in testing—PSTM—that helps in recognition of an old picture but is not similar enough to the decoy to increase false *yes*es to it.

General Discussion

The present experiments investigated the conceptual versus visual–pictorial nature of the fleeting memory for briefly glimpsed scenes reported by Potter et al. (2002). The question was whether (a) this short-term memory consists of just one type of information that is rapidly forgotten or (b) there are two distinct types of short-term memory with different time courses, one visual–pictorial—PSTM—and the other conceptual—CSTM. To investigate this question, we used two methods—tests using titles versus pictures and tests with decoy pictures versus old pictures—to distinguish between purely conceptual memory representations (titles or conceptually similar decoys) and representations that include visual as well as conceptual information (the pictures themselves). Our main interest in all of the experiments was the
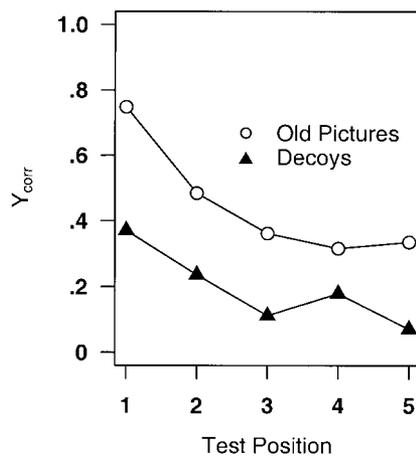
Figure 8. Probability of recognizing an old picture and falsely recognizing a decoy picture, corrected for guessing (Y_{corr}), as a function of relative serial position in the recognition test in Experiment 5; nondecoy distractors were used to assess the guessing rate in both cases. Presentation rate was 5.8 pictures per second (173 ms per picture).

relation between test position and performance with conceptual versus perceptual memory probes.

In Experiments 1–4, participants viewed short sequences of pictures and were then tested, either by viewing the same pictures mixed with new distractors or by viewing titles of the old pictures mixed with titles of distractors. The titles did not provide any explicit visual information, only the conceptual gist of each picture. Thus, to the extent that participants remember mainly gist, performance with titles should be similar to that with pictures in recognition testing. In Experiment 1, the pictures were presented at a rate of 1 per second, slow enough so that performance was near ceiling for the picture test, showing that the pictures were in LTM. Performance was lower for the title test, but still very good, indicating that titles can successfully tap memory for pictures. There was a significant decline in accuracy over the course of the recognition test, although it was a small effect.

In the subsequent experiments, pictures were presented for 173 ms each. In Experiment 2, overall performance was much worse than in Experiment 1, as expected from previous work, and performance in the title test was again significantly worse than that in the picture test. It is surprising to note, however, that the difference between titles and pictures was, if anything, smaller than in Experiment 1. There was a significant decline in performance over the course of the yes–no test, much more marked than that in Experiment 1, showing that there was short-term memory for the pictures that was lost over the first few seconds of testing. A Test Serial Position × Test Type interaction was significant in Experiment 2, with a much larger picture-test advantage early in the test.

In Experiment 3, the pictures were presented for 173 ms, as in Experiment 2, but they were followed by an 827-ms blank ISI, for an SOA between pictures of 1 s, as in Experiment 1. Performance was very similar to that in Experiment 1, showing that total processing time per picture determines whether or not the picture is consolidated in LTM, not actual presentation time (Intraub, 1980; Potter, 1976). Experiment 4 replicated the results of Experiment 2, using a within-subject design. These results support Potter et al.'s (2002) claim that when presentation is rapid, most pictures are remembered at least briefly, but many are forgotten over the course of the recognition test.

With respect to the questions addressed in the present study, the results of Experiments 2 and 4 support the hypothesis that the fugitive memory for a glimpsed picture consists in part of pictorial properties of the picture, information that is not tapped by a conceptual title. That is, for those pictures that are slated to be rapidly forgotten, a picture test given immediately can show that they were momentarily remembered, whereas an immediate title test does not produce the same high initial level of recognition. Performance continues to drop after the first test, but at about the same rate for pictures and titles, suggesting that conceptual information is also being forgotten. The relatively small difference between titles and pictures suggests that most of the information on which recognition is based (after the first test items) is conceptual—at least for very briefly presented pictures. (That a conceptual title is sufficient for recognition does not show that there is no memory for visual information about that picture, but it suggests that the initial visual information has fallen to a level at which there is little or no redundancy gain in testing with a picture rather than a title.) Toward the end of the 10-item test, performance approached an asymptote, suggesting that pictures still remem-

bered after 6–8 s of testing are in LTM, where they can be tapped either by a test picture or by a gist title.

In Experiment 5, we used a second method to address the same questions. A decoy picture with the same gist as one of the five old pictures was substituted for the old picture in the recognition test so that the test consisted of four old pictures, one decoy, and five distractors. The question was whether a picture with the same gist as one that had been presented would be falsely recognized—even though it was visually different from the presented picture. False *yes* responses to the decoy picture would be expected only if a participant retained conceptual information about the original picture but not specific visual information. If a participant remembered no more than conceptual gist equivalent to a title, such as *dogsled*, then a different picture of a dogsled should be falsely recognized as frequently as the original picture is correctly recognized. However, if specific pictorial information or conceptual detail is remembered, then little or no false recognition should be observed, because there is little or no specific pictorial overlap between decoys and the original pictures. In a control study, decoys and original pictures were easily discriminated when one was presented 173 ms after the other.

The results indicated that false recognition of decoys did occur, although it occurred less frequently than true recognition. As had been the case with title tests in the earlier experiments, the superiority of true old pictures was greatest at the beginning of the recognition test. However, just as memory for old pictures continued to decline during the test, false recognition of decoys also declined. This suggests that, as with recognition tests using titles, gist is being forgotten during the test, as well as pictorial properties.

Do Titles Tap Conceptual Memories or Visual Memories?

An alternative account of how titles are used to cue picture memory is the following.[11] Suppose that what is remembered is a visual–pictorial representation of the picture. When a title is presented, it is understood conceptually and is used to retrieve or construct an appropriate image (e.g., of a camel). That image is then compared with pictorial representations in recent memory, and if a match is found, the response is *yes*. The main problem with this alternative hypothesis is that a title has many (indeed, an infinite number of) potential matching pictures, only one of which would be imaged by the viewer of the title—so the chances of a good visual match with the actual picture shown are small.[12] In contrast, the probability is high that the title captures the conceptual gist of the picture. So, matching the meaning of the title to the remembered conceptual gist of the picture is likely to be easy.

PSTM, CSTM, and Visual Short-Term Memory (VSTM)

We have assumed that the initial peak in recognition of pictures is due to brief persistence of visual–pictorial information (PSTM), not conceptual information (CSTM). We cannot, however, rule out

---

[11] We thank a reviewer for this suggestion.

[12] There are also other problems with the imaging hypothesis, such as that images take seconds to generate and lack the vividness of a color photograph.

the possibility that conceptual details are the basis for the peak—details that are rapidly forgotten. Whereas the titles and decoys provide conceptual gist that matches presented pictures, they do not match the presented picture in conceptual detail. For example, the fact that the camel is lying down may be part of the conceptual information conveyed by Figure 7A, and the presence of two dogsleds may be part of the conceptual information in Figure 7C. If we assume that this additional conceptual information is quickly lost during the test, that could account for the initial peak for test pictures relative to titles or decoys.

Although there is no gold standard for distinguishing between conceptual and perceptual knowledge, there is some agreement that novel and meaningless combinations of visual features, such as color and shape, or novel patterns, such as partly filled checkerboard-like arrays, are remembered visually rather than conceptually. The visual properties of the array seem to exhaust the conceptual content.[13] That is why studies of VSTM have typically used colored geometric objects and the like. However, memory for such materials is usually poor except for the most recent stimulus (e.g., Phillips & Christie, 1977), so studies of VSTM capacity have normally looked only at memory for the most recent array (e.g., Luck & Vogel, 1997). Thus, the evidence that viewers can remember fleetingly four or five presented pictures in the present experiments (and up to 20 pictures in Potter et al., 2002) indicates that PSTM has different properties than VSTM as studied by Phillips and Luck, among others (Luck & Vogel, 1997; Phillips & Christie, 1977).

In the present study, visual information is bound to the conceptual identity of meaningful objects or scenes. This bound information is more distinctive and therefore easier to retain than an arbitrary combination of a limited number of shapes and colors. That may be why Hollingworth and Henderson (2002) found that visual properties, such as the orientation in depth of an object in a realistic scene, can be retained accurately over multiple intervening fixations on other objects in the scene. However, the evidence in their study that the information about orientation typically persists in LTM (see also Hollingworth, 2003a) leaves open the question of whether there is also disruptible short-term memory for object orientation that is like PSTM and, if so, whether it survives one or more subsequent fixations on other objects. In Hollingworth and Henderson's experiments, there was evidence that the probability of detecting an object change is lower when the object is viewed for a shorter time, but there was little evidence that increased time between initial viewing and the test affected memory. Using a new technique in which the viewer follows a dot that moves to a new object in a scene about once per second, Hollingworth (2003b; see also Hollingworth, 2004) found that most objects are remembered stably in LTM, although there is significantly better memory for the two most recent objects viewed, suggesting a VSTM capacity of two objects. This result seems consistent with the time course of PSTM in the present study.

Although it is not clear from previous work whether visual information is distinct from conceptual information, either in kind or in time course of forgetting, we adopt a widely accepted assumption that rich visual detail is lost more rapidly than more holistic and categorical semantic information. This assumption is consistent with the present evidence that picture recognition tests, if administered immediately, tap information about presented pictures that is not tapped by a title test or a decoy picture.

Do Titles Cue Retrieval of Visual Information?

Just how are titles recognized? We can be certain that viewers are not encoding pictures verbally when they are presented at a rate of 6 per second (Intraub, 1979), so there is little likelihood that a verbal title was generated during viewing and matched to the test title. Instead, the test title must be understood and then matched to the encoded conceptual gist of a picture in memory. However, it is clear from the decoy experiment that participants remembered more than bare gist, equivalent to the titles we used. We propose that a title's gist is the basis for initial retrieval of the matching picture but that more information about the picture is then recalled, including visual information: The picture is redintegrated (Horowitz & Prytulak, 1969). This process accounts for the strong intuition that we can picture a recalled scene. When the test picture is a decoy, its gist also results in retrieval of the relevant picture, but subsequent redintegration of the retrieved picture is sufficient in most cases to show that it does not match the decoy—the participant recalls to reject (e.g., Clark & Gronlund, 1996).

Conclusion

Pictures glimpsed briefly in a rapid sequence mimic the input from successive eye fixations in an idealized situation in which there is no overlap in content from one simulated fixation to the next. Earlier work (Potter, 1976) showed that such pictures can be momentarily understood, but they are then likely to be forgotten within minutes. In more recent work, Potter et al. (2002) showed that forgetting was not instantaneous: As many as 20 pictures in a sequence could be remembered with a high probability as long as they were tested immediately, but as the test continued, there was rapid forgetting. In the present study, we showed that this short-term memory for pictures has two components, one visual–pictorial (PSTM) and the other conceptual (CSTM). These two components coexist initially, but PSTM is lost very early in testing, within 2–3 s, whereas CSTM lasts somewhat longer, for 5–6 s. Further forgetting is very gradual, showing that the surviving pictures were those whose visual and conceptual properties were consolidated into LTM during presentation. As viewers make each new fixation under normal viewing conditions, information about the visual as well as conceptual contents of the last few fixations may be available, although fated to be forgotten shortly. Such transient information—initially both perceptual and conceptual, then primarily conceptual—may be important in maintaining and updating a coherent representation of the scene.

---

[13] We distinguish between conceptual information and verbal information; whereas most concepts (as well as most visual properties) can be expressed in words, the words themselves are not concepts: They are pointers to concepts.

References

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3,* 37–60.

Donaldson, W. (1993). Accuracy of *d′* and *A′* as estimates of sensitivity. *Bulletin of the Psychonomic Society, 31,* 271–274.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75,* 424–429.

Hollingworth, A. (2003a). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 388–403.

Hollingworth, A. (2003b, May). *Short- and long-term memory contributions to the online visual representation of natural scenes.* Poster presented at the Third Annual Meeting of the Vision Sciences Society, Sarasota, FL.

Hollingworth, A. (2004). *The relationship between online visual representation of a scene and long-term scene memory.* Manuscript submitted for publication.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 113–136.

Horowitz, L. M., & Prytulak, L. S. (1969). Redintegrative memory. *Psychological Review, 76,* 519–531.

Hunt, S. M. J. (1994). MacProbe: A Macintosh-based experimenter's workstation for the cognitive sciences. *Behavior Research Methods, Instruments, & Computers, 26,* 345–351.

Intraub, H. (1979). The role of implicit naming in pictorial encoding. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 78–87.

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.

Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 604–610.

Luck, S. J., & Vogel, E. K. (1997, November 20). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge, England: Cambridge University Press.

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

O'Connor, K. J., & Potter, M. C. (2002). Constrained formation of object representations. *Psychological Science, 13,* 106–111.

Phillips, W. A., & Christie, D. F. M. (1977). Components of visual memory. *Quarterly Journal of Experimental Psychology, 29,* 117–133.

Pollack, I., & Norman, D. A. (1964). A nonparametric analysis of recognition experiments. *Psychonomic Science, 1,* 125–126.

Potter, M. C. (1974). [Normal–normal slope of ROC curves for picture recognition]. Unpublished raw data.

Potter, M. C. (1975, March 14). Meaning in visual search. *Science, 187,* 965–966.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 509–522.

Potter, M. C. (1993). Very short-term conceptual memory. *Memory & Cognition, 21,* 156–161.

Potter, M. C. (1999). Understanding sentences and scenes: The role of conceptual short term memory. In V. Coltheart (Ed.), *Fleeting memories* (pp. 13–46). Cambridge, MA: MIT Press.

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

Potter, M. C., Staub, A., Rado, J., & O'Connor, D. H. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1163–1175.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368–373.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition, 7,* 127–145.

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

Simons, D. J., Chabris, C. F., Schnur, T. T., & Levin, D. T. (2002). Evidence for preserved representations in change blindness. *Consciousness & Cognition, 11,* 78–97.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Psychology: General, 117,* 34–50.

Snodgrass, J. G., & McClure, P. (1975). Storage and retrieval properties of dual codes for pictures and words in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 521–529.

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25,* 207–222.