Effect of retention interval on implicit and explicit memory for pictures

JOAN GAY SNODGRASS
New York University, New York, New York

and

AIMEE SURPRENANT
Yale University, New Haven, Connecticut

A picture-fragment completion task and a yes–no recognition task were used to determine whether rates of forgetting in perceptual (implicit) and episodic (explicit) memory systems were the same. In Session 1, identification thresholds for 60 fragmented pictures were measured. In Session 2, which occurred after one of four delay periods (immediate, 2 days, 7 days, or 14 days), the subjects were presented with half of the pictures shown in the previous session mixed with an equal number of new pictures, and their identification thresholds were measured. A yes–no recognition test for the picture names (90 names of pictures shown in the picture-fragment completion task, 30 new-picture names) was given after the identification phase, and both speed and accuracy of recognition were measured. Perceptual learning was measured by relative savings in identification of repeated, compared to new, pictures. Episodic learning was measured in two ways: by the corrected recognition score for names of pictures presented in Session 1 alone (hits minus false alarms) and by the difference in hit reaction times between Session 1 pictures and the new pictures presented for identification in Session 2. These three learning measures showed approximately equal rates of forgetting; furthermore, the episodic learning measure based on accuracy showed more similarity to the perceptual learning measure than it did to the episodic learning measure based on reaction time. Accordingly, the results of this experiment do not support functional dissociation of implicit and explicit memory systems by different forgetting rates.

Separate memory systems have been proposed to account for observed dissociations between performance on an implicit test such as perceptual identification and an explicit test such as recall or recognition (Tulving, 1985; Tulving, Schacter, & Stark, 1982). The terms implicit and explicit memory were introduced by Graf and Schacter (1985) to characterize the differences between the two types of memory tests in a theoretically neutral way: *implicit memory* is exhibited by an increase in a subject's skill in processing a previously presented stimulus; *explicit memory* is exhibited by a subject's awareness that he or she encountered that stimulus in a prior episode. The two types of memory are thus functionally distinguished by the type of retrieval test.

The major experimental evidence for asserting that implicit and explicit memory constitute separate (but related) memory systems is based on findings of functional and stochastic independence between performance on implicit and explicit memory tests. The most compelling evidence for this functional dissociation comes from findings that amnesic patients show preserved learning on implicit memory tasks, relative to their disastrous performance on explicit memory tasks (e.g., Cohen & Squire, 1980; Graf, Squire, & Mandler, 1984; Schacter & Graf, 1986) and a slower rate of decay of implicit relative to explicit memory

in normal subjects (Gollin, 1966; Jacoby, 1983; Jacoby & Dallas, 1981; Tulving et al., 1982). Recently, however, Sloman, Hayman, Ohta, Law, and Tulving (1988) were unable to replicate the Tulving et al. (1982) finding that subjects revealed no forgetting in implicit memory over a 1-week period, and they concluded that the conditions that produced such resistance to forgetting are at present unknown.

The present experiment was designed to study the rate of forgetting in both implicit and explicit memory tasks across a range of delays up to 14 days. We used picture-fragment completion as the implicit memory task and picture-name recognition as the explicit task. The picture-fragment completion task was similar to the word-completion task used by Tulving et al (1982), except that the pictures were shown at more than one level of completeness so that if the subject could not identify the picture at the most fragmented level, successively more complete pictures were presented until correct identification was achieved. This methodology provides a more finely graded measure of performance for each picture than simple success or failure. We also attempted to measure both picture-fragment completion and recognition-memory performance (accuracy and speed) on comparable scales so that forgetting rates could be compared across implicit and explicit measures.

METHOD

Subjects and Design
The 32 subjects were volunteers from an introductory psychology course at New York University who participated in the experiment

to fulfill a requirement for the course. The subjects were randomly assigned to conditions and tested individually.

The experiment was designed to test both perceptual and episodic memory on separate sets of pictures in order to avoid intertest transfer effects (Shimamura, 1985). Accordingly, the order of the tests was not varied; instead, the recognition test always followed the picture-fragment completion test. The training session (Session 1) consisted of picture-fragment completion on two sets of 30 fragmented pictures. One set served as the repeated set for the fragmented-picture test and is denoted S_{11}; the second set served as the repeated set for the recognition-memory test and is denoted S_{10}. Set S_{11} was presented in both the training and the fragmented-picture test; set S_{10} was presented only in the training session.

In the fragmented-picture test administered during Session 2, two sets of 30 pictures were also presented: S_{11r} (the repeated set from Session 1) and S_{01}, a completely new set. Finally, four sets of picture names were presented for the recognition test: S_{11}, S_{01}, S_{10}, and S_{00} (names of never-presented pictures).

Apparatus and Materials

The 120 pictures used in the fragmented-picture task were taken from Snodgrass and Vanderwart's (1980) picture set. Eight levels of fragmentation are possible, with Level 8 being the complete picture and Level 1 the most fragmented. All stimuli were presented on the Apple Macintosh microcomputer at the most fragmented level (1) and were subsequently completed one level at a time until the subject correctly identified the picture by typing its name (Snodgrass, Smith, Feenan, & Corwin, 1987).

The pictures were divided into four sets of 30 pictures each. The sets were matched on the basis of perceptual identification thresholds from a previous study (Snodgrass & Corwin, 1988). Within each delay group, the assignment of sets to training condition was counterbalanced across subjects. All sequences were randomized by the computer.

Procedure

The 32 subjects were randomly assigned to one of the four delay conditions until each delay condition contained 8 subjects. The subjects were informed that they would be asked to try to identify pictures of common objects from fragments. The subjects were then asked to sign a consent form. The remaining instructions were presented on the computer.

During Session 1, the subjects performed picture-fragment completion of 60 fragmented pictures. Each fragmented-picture series was presented by a systematic method of limits in which the most fragmented picture was presented first, followed by increasingly more complete representations of the picture. If the subjects thought they knew the picture's name, they typed it into the computer. If the answer was correct, the trial was completed and the identification threshold was recorded by the computer. In the case of pictures for which more than one name was possible (e.g., gorilla or ape), the computer had the acceptable alternatives stored in memory.

During Session 2, the subjects received a picture-fragment completion test and a recognition-memory test of the pictures' names. For the immediate (no delay) group, Session 1 and Session 2 were separated by a 10-min number-cancellation distractor task. The remaining delay groups returned after the appropriate number of days (2, 7, or 14). In the picture-fragment completion test, 30 old pictures (set S_{11} shown in the previous session) and 30 new pictures were presented, using the same instructions and procedure as for Session 1. The subjects were not informed that some of the pictures would be repeated from Session 1. Picture sequences in both sessions were presented in a different random order for each subject.

Following the picture-fragment completion test, the subject was given a surprise recognition test. A total of 120 words were presented on the Macintosh screen, one at a time. The subjects were instructed to choose whether or not they had seen the picture corresponding to the word in either session by pressing the slash (/) key for old (seen before) and the "z" key for new (not seen before). They were told that their reaction times were being measured so they should

be as fast as possible. Reaction times (RTs) were recorded to the nearest 16 msec.

The stimuli in the recognition test consisted of 90 names of pictures seen in one or both of the previous sessions, and 30 names of pictures that had not been seen in either session. Old names thus comprised 75% of the pictures. The subjects were presented with a fixation point (+), which appeared in the middle of the screen for 2 sec, a blank screen for 1 sec, and then the name. The name stayed on the screen until the subject indicated whether it was old or new, at which time the next trial began with the presentation of the fixation point. All stimulus sequences were presented in a different random order for each subject. At the end of Session 2, the subjects were debriefed and thanked.

RESULTS AND DISCUSSION

To make the data from picture-fragment completion more comparable to name recognition, thresholds were converted to P values by the linear transformation: $P = (9-T)/8$, where T is the threshold. Because T varied from 1 to 8, P varied from 1.0 to .125.

The mean proportions identified for the four sets of fragmented pictures are shown in Table 1. Because nothing distinguished the to-be-repeated set (S_{11}) from the not-to-be-repeated set (S_{10}) in Session 1, we expected performance to be equivalent on these two sets, and it was. We expected a slight improvement in performance from Session 1 to Session 2 on the new set (S_{01}) because of task- or skill-learning effects, but this was not evident in this experiment except for a small effect for the 7-day group. However, as expected, the subjects showed considerable learning for the repeated set (S_{11r}) from Session 1 to Session 2.

Because this was a between-subjects design, the groups differed somewhat in their performance levels. In particular, the 2-day group showed worse performance in picture identification than did any of the other groups. Because priming effects are normally measured with respect to a new-item baseline, we analyzed perceptual learning using the two measures shown in Table 1. The absolute measure is the difference between old and new pictures during Session 2 and represents the usual measure of priming. The relative measure is the absolute difference divided by 1 minus the new-picture performance. Elsewhere, we have argued that this particular relative-difference measure is a better measure of perceptual learning than the absolute measure when baseline differences occur (Snodgrass, 1989; Snodgrass, Corwin, & Feenan, 1988). Both the absolute- and the relative-difference measures showed a

Table 1
Proportion of Picture Fragments Identified as a Function of Delay

	Session 1		Session 2		Perceptual Learning	
	S_{11}	S_{10}	S_{11r}	S_{01}	Absolute	Relative
Immediate	0.59	0.60	0.79	0.59	0.21	0.51
2-day	0.54	0.54	0.73	0.55	0.18	0.40
7-day	0.58	0.58	0.74	0.62	0.12	0.32
14-day	0.58	0.57	0.70	0.58	0.12	0.29

Note—S_{11} = set presented in the first session and to be repeated in the second session; S_{10} = set presented only in the first session; S_{01} = set presented only in the second session; S_{11r} = set repeated in second session (same pictures as S_{11}); Absolute = $S_{11r}-S_{01}$; Relative = $(S_{11r}-S_{01})/(1-S_{01})$.

Table 2
Proportion of Picture Names Recognized as Old as a Function of Delay

| Delay | Frequency/Recency | | | | Episodic Learning |
	S_{11}	S_{01}	S_{10}	S_{00}	P_r
Immediate	0.99	0.98	0.93	0.13	0.80
2-day	1.00	0.97	0.89	0.14	0.75
7-day	0.98	0.98	0.76	0.30	0.46
14-day	0.99	0.95	0.80	0.28	0.51

Note—S_{11} = set presented in both sessions; S_{01} = set presented only in the second session; S_{10} = set presented only in the first session; S_{00} = set presented as new in the recognition-memory task; $P_r = S_{10} - S_{00}$.

Table 3
Mean Reaction Time in the Recognition-Memory Test as a Function of Delay

| Delay | Frequency/Recency | | | | | Episodic Learning: Difference |
	S_{11} Hits	S_{01} Hits	S_{10} Hits	S_{00} FAs	S_{00} CRs	
Immediate	792	821	896	1055	1091	75
2-day	896	1001	1147	1315	1375	145
7-day	736	757	908	1018	1171	150
14-day	1070	1133	1555	1830	2221	422

Note—S_{11} = set presented in both sessions; S_{01} = set presented only in the second session; S_{10} = set presented only in the first session; S_{00} = set presented as new in the recognition-memory task; FAs = false alarms; CRs = correct rejections; difference = $S_{10} - S_{01}$.

monotonic relationship with retention interval, although only the relative-difference measure showed a drop between the 7-day and the 14-day retention intervals.

One-way analyses of variance (ANOVAs) were performed on both measures. In these and subsequent analyses, three planned comparisons were conducted when the ANOVA was significant. The first planned comparisons compared the immediate group with all delayed groups combined, the second compared the 2-day group with the 7- and 14-day groups combined, and the third compared the 7-day group with the 14-day group. All reported differences are significant at $p < .05$.

Both the absolute- and the relative-difference measures showed exactly the same effects. There was a significant effect of delay [$Fs(3,28) = 7.18$ and 6.13, for the absolute and relative measures, respectively]. The subjects in the immediate condition performed significantly better than did those in the delayed conditions [$Fs(1,28) = 12.77$ and 13.83, respectively], the 2-day group performed significantly better than the 7- and 14-day groups combined [$Fs(1,28) = 8.87$ and 4.13, respectively], and the 7- and 14-day groups did not differ. In summary, we conclude that perceptual memory in this experiment decayed with retention interval, but seemed to reach an asymptote after about a week.

The proportions of picture names recognized as old in the yes–no recognition task are shown in Table 2. The conditions are listed in descending order by frequency and/or recency, so the proportions of recognitions are expected to decrease from left to right, as they do. The values in the S_{00} column represent false-alarm rates. Hit rates for twice-presented (S_{11}) and recently presented (S_{01}) picture names were very close to ceiling; however, hit rates for the names of pictures presented only in the first session

were lower, particularly for the delayed groups. The false-alarm rates were considerably higher than the miss rates, which indicates that the subjects exhibited a yea-saying bias. This is consistent with the design of the yes–no recognition test, in which 75% of the picture names were old.

Because we were interested in episodic memory for only the delayed stimuli, corrected recognition scores (P_r) were calculated by subtracting false alarms from hits for the S_{10} set. A one-way ANOVA on the P_r scores showed a significant effect of delay [$F(3,28) = 8.38$]. Planned comparisons showed that the immediate group performed significantly better than the delay groups [$F(1,28) = 10.95$], the 2-day group performed significantly better than the 7- and 14-day groups combined [$F(1,28) = 13.84$], but the 7- and 14-day groups did not differ ($F < 1$). In summary, the pattern of results for this measure of episodic learning is identical to that observed for perceptual learning—episodic memory in this experiment decayed with retention interval, but reached an asymptote after about a week.

Mean RTs in the recognition test are shown in Table 3. Outliers, defined as RTs greater than four standard devia-
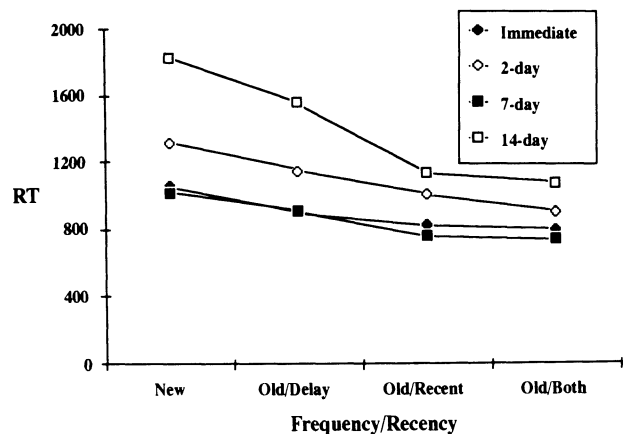

Figure 1. Reaction times to names of old pictures as a function of recency and/or frequency for each delay group.
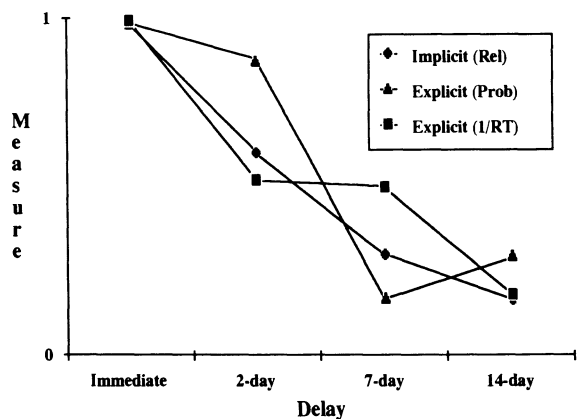

Figure 2. Three measures of learning for the four delay intervals. The measure of implicit learning is the relative measure computed by $(S_{11r} - S_{01})/(1 - S_{01})$; the probability measure of explicit learning is P_r, computed as $S_{10} - S_{00}$; the 1/RT measure of explicit learning is computed as $1/(S_{10} - S_{01})$. All measures have been linearly transformed to equate their ranges.

tions above a subject's mean for that condition, and all RTs greater than 10 sec, were replaced by the maximum acceptable RT before computing the means (less than 1% of the data). For the old-picture sets, only hit RTs are shown, as the miss rates were too low to produce stable miss RTs. However, for the new-picture set, RTs for both false alarms and correct rejections are shown. The conditions are ordered by frequency and/or recency, so the RTs are expected to increase from left to right, as they do. It is clear that the RT measure is sensitive to the frequency/ recency manipulation within each delay group. It is notable that the false-alarm RTs were actually faster than the correct-rejection RTs, even though the probability of a correct rejection was higher than the probability of a false alarm.

Figure 1 shows RTs for the four types of yes responses for each delay group, plotted so that performance increases (RT decreases) from left to right. There were large group differences in overall RT, with the 7-day group responding as quickly as the immediate group. To correct RT for baseline differences, we looked for a transformation that would show the expected ordering of learning effects across delay groups. The simplest transformation, shown in the last column of Table 3, was to subtract RTs to the recently presented set (S_{01}) from those to the delayed set (S_{10}). We used the recently presented set as the baseline because the study/test delay was the same for all four groups. This difference measure increased monotonically with delay interval, which is what we would expect if strength of explicit memory decays monotonically with delay and if RT is a direct measure of memory strength.

The results of a one-way ANOVA on the difference RTs showed that the overall effect of delay was significant [$F(3,28) = 4.71$]. Planned comparisons between the immediate and the delay groups were significant [$F(1,28) = 4.07$], the 2-day group's performance was not significantly different from that of the 7-day and 14-day groups combined, and the 7-day group showed significantly less forgetting than did the 14-day group [$F(1,28) = 7.41$]. Thus, this RT difference is the only measure of learning that showed forgetting beyond 7 days.

Figure 2 shows the three measures of learning—relative perceptual learning from the fragment-completion task, episodic learning using the P_r measure, and episodic learning using the reciprocal of the RT difference measure—plotted as a function of delay. Values of the three measures have been linearly transformed to equate their ranges. Although all three measures show substantial forgetting with delay, the pattern varies somewhat depending upon the measure. Both the implicit and explicit P_r measures show large differences between the 2-day and the 7-day groups, with smaller or even negative differences (for the explicit measure) between the 7-day and the 14-day groups. In contrast, the RT measure shows no difference between the 2- and the 7-day groups, but a large difference between

the 7- and the 14-day groups. Thus, the two measures of explicit memory show less similarity with one another than the P_r measure does with the measure of implicit memory. The results of this experiment, then, do not support the notion of a functional dissociation between explicit and implicit memory systems in normal subjects when retention interval is the independent variable.

REFERENCES

COHEN, N., & SQUIRE, L. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210, 207-209.

GOLLIN, E. S. (1966). Serial learning and perceptual recognition in children: Training, delay and order effects. *Perceptual & Motor Skills*, 23, 751-758.

GRAF, P., & SCHACTER, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic patients. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 11, 501-518.

GRAF, P., SQUIRE, L., & MANDLER, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 10, 164-178.

JACOBY, L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 9, 21-38.

JACOBY, L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 110, 306-340.

SCHACTER, D. L., & GRAF, P. (1986). Preserved learning in amnesic patients: Perspectives from research on direct priming. *Journal of Clinical & Experimental Neuopsychology*, 8, 727-743.

SHIMAMURA, A. P. (1985). Problems with the finding of stochastic independence as evidence for multiple memory systems. *Bulletin of the Psychonomic Society*, 23, 506-508.

SLOMAN, S. A., HAYMAN, C. A. G., OHTA, N., LAW, J., & TULVING, E. (1988). Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 14, 223-239.

SNODGRASS, J. G. (1989). Sources of learning in the picture fragment completion task. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues*. Hillsdale, NJ: Erlbaum.

SNODGRASS, J. G., & CORWIN, J. (1988). Perceptual identification thresholds for 150 fragmented pictures from the Snodgrass and Vanderwart picture set. *Perceptual & Motor Skills*, 67, 3-36. (Monograph Supplement 1-V67)

SNODGRASS, J. G., CORWIN, J., & FEENAN, K. (1988, November). *Some issues in measuring priming effects*. Paper presented at the meetings of the Psychonomic Society, Chicago, Il.

SNODGRASS, J. G., SMITH, B., FEENAN, K., & CORWIN, J. (1987). Fragmenting pictures on the Apple Macintosh computer for experimental and clinical applications. *Behavior Research Methods, Instruments, & Computers*, 19, 270-274.

SNODGRASS, J. G., & VANDERWART, M. (1980). A standardized set of 260 pictures: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, 6, 174-215.

TULVING, E. (1985). How many memory systems are there? *American Psychologist*, 40, 385-398.

TULVING, E., SCHACTER, D. L., & STARK, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 8, 336-342.