Memory for unique personal events:
The roommate study

CHARLES P. THOMPSON
Kansas State University, Manhattan, Kansas 66506

Memory for naturally occurring episodic events was measured along with memory for the date of occurrence of those events. The effect of rehearsal was also measured. Participants in the experiment recorded unique personal events for themselves and their roommates for a semester (data were collected for approximately 14 weeks). They also rated the memorability of the events at the time the events were recorded. The roommates were unaware that they were to be tested on the events until approximately 1 week before the memory test. Accuracy in dating events decreased at the rate of roughly 1 day every week of the retention interval. Events rated as memorable were encoded better than events rated as unmemorable but were forgotten at the same rate. Increasing rehearsal decreased the rate of forgetting the events. Although the recorders selected the events to be recorded and knew they were to be tested on those events, they did not differ from their roommates on any of the memory measures.

Naturally occurring episodic memory consists, in part, of events shared with many people (e.g., the wedding of the Prince of Wales and Lady Diana Spencer in Great Britain, 1981). For the most part, however, an individual's episodic memory consists of personal events that are either entirely private or shared with relatively few people. For that reason, it has been difficult to measure episodic memory outside the laboratory. Most studies of naturally occurring memory have looked at semantic memory or a mixture of semantic and episodic memory. For example, one set of investigations has used a subset of memories that are commonly held (e.g. the Russian invasion of Afghanistan, the death of Peter Sellers) to study natural memory (e.g., Underwood, 1977; Warrington & Sanders, 1971). These events undoubtedly represent, for any single individual, a mixture of episodic (i.e., dated as an autobiographical experience) and semantic (i.e., general knowledge) memories. However, even if one could sort out the subset of episodic events, commonly held memories are not the most significant personal memories and, thus, are probably not particularly well remembered. Further, it is surprisingly difficult to identify a large set of events that are known to most of the participants in an experiment. Finally, a commonly held memory that is significant to one individual (e.g., an international relations expert) may be of little interest to another (e.g., an avid movie fan), and vice versa. All of these factors suggest that, although commonly held memories provide a useful tool for investigating naturally occurring memory, the common-memory approach is somewhat less than ideal for examining episodic memory.

A handful of studies have investigated personal events in naturally occurring episodic memory. They include studies on the cued recall of events, which are then dated approximately (e.g., Crovitz & Schiffman, 1974),

studies investigating recall of a few specific episodic events (e.g., Baddeley, Lewis, & Nimmo-Smith, 1978; Wagenaar, 1978), and studies using diaries (Linton, 1975; Thomson, 1930). In the diary studies, subjects recorded events from their daily lives and were later tested for memory of these events. The diary studies have the greatest potential for analytic examination of episodic memory because the procedure allows exact dating of a large number and a variety of episodic events. The diary procedure also has the advantage of capturing the personal and private nature of episodic memory. Finally, the procedure measures memory using very strong cues for memory retrieval. That is, the subject is provided a brief description of the event as a cue for retrieval.

With that background in mind, the experiment presented here used the diary procedure with naive subjects in order to answer a number of questions about naturally occurring episodic memory. Because so few data are currently available about memory for personal events, we concentrated on some basic theoretical and methodological questions. In addition, the experiment was intended to provide data that would clearly describe the forgetting of a wide range of naturally occurring episodic events.

We wanted to determine, first, whether recording events improves memory for those events. If it does, we wanted to know why and how much. To investigate that question, half the participants in the experiment recorded events from their lives, and, in addition, they recorded events for someone with whom they lived (usually a roommate). The roommates were not informed about the procedure until just prior to the testing period. Under these conditions, at least four factors could operate to produce superior performance on the memory test for the recorders. That is, the recorders

knew that they were to be tested on a set of personal events (thus having the advantage of intentional as opposed to incidental recall), they recorded those events (thus introducing additional rehearsal), they selected the events to be recorded (thus introducing a potential event selection bias), and they could remember the description instead of the event (thus contrasting recognition with recall). The most reasonable prediction is that one or more of these factors will operate to produce superior memory performance for the recorders. Should that prediction hold, any estimation of memory for unique personal events would have to take into account the inflation produced by those additional factors.

Second, we wanted to determine whether the rate of forgetting differs for memorable and unmemorable events. As part of the procedure, the recorders rated the memorability of events at the time the events were recorded. With a set of events spanning more than 3 months, one can quite accurately evaluate whether the forgetting functions for memorable and unmemorable events are parallel or diverge over time. Most laboratory studies suggest that memories encoded at different levels or strengths are forgotten at the same rate (e.g., Nelson & Vining, 1978; Underwood, 1964). However, it must be noted that laboratory studies use materials that, even in their most varied form, are quite homogeneous when compared with the variety of events recorded from everyday experience. Nevertheless, the best assumption is that differences in memory for naturally occurring events reflect differences in encoding rather than rate of forgetting.

Third, a question arises regarding the accuracy with which recorders can estimate the recallability of events. The recorders rated both their own and their roommates' events. Also, previous investigators (Bower & Gilligan, 1979; Rogers, Kuiper, & Kirker, 1977) have shown that judging traits for oneself produces better recall than does judging traits for others. Presumably, this result occurs because subjects can store the self-judged traits in a very well differentiated schema. We would speculate that rated memorability (at the time of recording) is based on either the ease with which an event fits into the subject's personal memory schema or the personal importance of the memory (or both). In either case, we would predict that the recorder will judge the memorability of personal events more accurately than the memorability of events recorded for a roommate.

The fourth question was whether the clarity (or strength) of a memory is related to the accuracy of dating a remembered event. It is clearly possible that remembering the content of an event might be independent of and/or different from remembering when the event occurred. For example, dating an event might depend totally upon reconstruction, whereas recalling the content of an event does not. Thus, we do not assume that remembering when an event occurred will necessarily be related to how well the event is remembered. (We exclude the special case in which an item is totally forgotten and, of course cannot be accurately dated.) As a related point, we were interested in determining how accurate memory is for the date of occurrence of events. Since personal events provide most of the markers for human life, it is of some interest to determine how accurately the time of occurrence of those events is remembered.

Fifth, what is the effect of rehearsal on memory? If anything is well established, it is that rehearsal will increase performance on a memory test (except for a recall following primary rehearsal; see Craik & Watkins, 1973). However, rehearsal in a laboratory study usually occurs immediately and affects initial encoding. In contrast, frequent rehearsal of a natural event should be spread over time, with increasing intervals between rehearsals. This pattern of rehearsals should have the effect of preventing forgetting in a manner similar (or identical) to that demonstrated by Landauer and Bjork (1978). They showed that an increasing interval between rehearsals produced better performance than did equally spaced rehearsals when the information was not repeated for study. Similarly, we would expect items that are rehearsed frequently to show less rapid forgetting than items that are rehearsed infrequently or not at all.

Finally, we were interested in determining whether we could predict memory for personal events with a laboratory test (e.g., word-list memory) or a self-rating measure. One possibility is that individuals can accurately assess their memory ability. Thus, as a first approach to predicting performance, the memory self-rating scale devised by Herrmann and Neisser (1978) was administered to all participants in the experiment. Their self-rating was then correlated with their performance on the measures of memory used in the experiment.

METHOD

Subjects

Thirty-two people served as subjects in the experiment. Sixteen were members of an information processing class and received class credit for participation. The other 16 participated without any renumeration other than feedback on their performance. Six additional subjects (three class members and three roommates) were eliminated from the study. Two pairs were dropped because the majority of the materials prepared by the class members did not meet the requirement for uniqueness specified below. The other pair was dropped because the roommate did not wish to participate.

Materials

The materials were recorded by the 16 class members and consisted of unique events that occurred during the course of the semester. Each week, each class member attempted to record two events each day (Monday through Thursday) for two individuals, the class member and an individual who shared living quarters with the class member (usually a roommate). Events were recorded for Monday through Thursday because roommates were often separated over the weekend. There were three restrictions on the events. First, they were to be unique (i.e., they were expected to occur no more than once during the semester). Second, they could not be embarrassing. Third, they were to be described in three written lines or less.

Preratings

At the time the events were recorded, the individual recording the event estimated the recallability of the event on a 3-point scale, with 3 being extremely memorable and 1 being not very memorable. An extremely memorable event was defined as an event that would probably be remembered 1 year after it had occurred. A not very memorable event was defined as one that would probably be forgotten within 2 weeks. These initial memorability ratings were given for all recorded events.

The subjects recorded a wide range of events that, even to the most casual observer, obviously differed in significance and memorability. The nature of the events clearly varied, and the rating of events seemed to vary from recorder to recorder. Thus, it is difficult to attempt to describe typical events. For example, events rated as not very memorable included "Found an earring on the apartment shuttle bus," "Saw prisms in physics class," and "Played football with Ed and Kirk." By contrast, the following events were rated as extremely memorable: "Someone who was in my life in a very special way told me they no longer wanted to be," "Got elected to house disciplinary board," and "Lori stood me up, then came over and picked me up."

In addition, all participants (i.e., class members and roommates) were given the short inventory of memory experiences (SIME) developed by Herrmann and Neisser (1978) about 1 week prior to the recall test. SIME requires responding on a 7-point scale to a series of 32 questions such as: (1) How often do you remember something that somebody said to you, but forget just who said it? (2) When you put something away and then look for it a week or so later, how often do you forget where you had put it? The responses to the questions on SIME can be partitioned in a number of ways, with the major partitions being an overall forgetting score (based on 24 responses), a score for memory for recent events (4 responses), and a score for memory for remote events (4 respones). SIME was given so that we might attempt to relate the self-ratings of memory ability to performance on the recall test.

Procedure

Each of the 16 participating class members generated events as described beginning the 2nd week of the semester. The items were collected on Friday of each week. Testing began 2 weeks prior to the end of the semester. (The semester is 16 weeks long, not including spring break.) All participants were tested individually and events were recorded until the day preceding the memory test. The recorders knew that it was important that they not inform their roommates about the procedure until just before the memory test. The comments of the recorders when they were asked to inform their roommates suggest that the recorders had held to that requirement. However, there was no way to objectively assess whether they had or not. The recorders were asked to inform their roommates about the experiment approximately 1 week prior to the memory test, at which time the roommates were given the opportunity to withdraw. All participants then followed the same procedure for the memory test. The events were numbered sequentially, beginning with the first event recorded. They were then read to the subject in one of two random sequences generated by a TRS-80 computer.

As each event was read, the subject first determined whether or not the event was unique. This procedure was used because subjects obviously could not accurately date events that occurred more than once during the semester. Subjects had no problem making the uniqueness judgment. Indeed, nonunique events were usually obvious to the experimenter (e.g., "went out for a beer"). If the event was not unique, it was deleted and the subject went on to the next event. If the event met the requirement of being unique, the subject entered into the computer a rating reflecting how well it was remembered (on a 7-point scale). The SIME memory scale, used for this purpose, is as follows: not at all (1),

barely at all (2), not so well (3), fairly well (4), very well (5), almost perfectly (6), and perfectly (7). If the event was not remembered at all, the "not-at-all" rating was entered and the subject went on to the next event. Again, it would not make sense to ask for the date of occurrence or rehearsal ratings for items that were not remembered. If the event was remembered, subjects rated how many times the subject had thought or talked about the event on a 6-point rehearsal scale. The rehearsal scale had both descriptive and numeric labels, as follows: never (0) = 1, rarely (1-2) = 2, a few times (3-5) = 3, several times (6-10) = 4, quite often (11-20) = 5, and very often (over 20) = 6. Finally, a calendar with no information other than date and day of week was used by the subject to estimate the date of occurrence of the event.

Following testing, the computer printed out, in chronological order for those items that met the criterion of being unique, the subjects' memory and rehearsal ratings for each event, the magnitude of the dating error (if any), and the initial memorability rating for each event.

RESULTS

In all analyses that follow, the retention intervals are equivalent for all subjects. That is, they are measured from the date of testing for each subject. Further, the analysis of the data reflects two characteristics of the data that should be noted at the outset. First, every subject had at least some days on which no events were recorded, and there were a few cases in which as much as a 2-week block of data was missing. The data are therefore pooled over blocks of 1 or more weeks and represent the average of the available data for that block. Further, to capture as much of the day-by-day performance as possible and simultaneously avoid missing cells, I adopted the strategy of performing an overall analysis on the major variables (e.g., memory rating) using a smaller block size, followed by analyses including other variables (e.g., rehearsal rating) that used larger block sizes. (Note that the analyses are therefore somewhat redundant, because the major variables are included in the analyses on both smaller and larger block sizes.) Because subjects did not select events to systematically vary initial memorability rating and rehearsal rating, block size was selected by perusing the data and estimating the block size that would produce very few missing data points. In a few cases, block size was changed (increased) after subjects' scores made it clear that there would be many missing data cells. Even with that procedure, there are occasional missing cells. I used the convention of assigning the mean of the appropriate condition to the missing cell for purposes of analysis. The number of missing cells will be stated in those analyses in which they occurred.

Second, both the memory rating and the estimated date data tended to be very stable for the shorter retention intervals (3-4 weeks) and then to become much more erratic. This is reflected in the analyses, in that a larger block size is used for the longer retention intervals than is used for the shorter retention intervals. In addition, it seemed possible that the larger variance at the

longer retention intervals might mask an effect that would be detectable at the shorter retention intervals. For that reason, separate analyses were performed for the short and long retention intervals on all data subsequently described. The results of those separate analyses were identical to the overall analysis in every case except the dating error by initial memorability rating analysis (described below). Thus, the results of the separate analyses are reported only in the latter case. Unless otherwise indicated, all analyses are standard between/within-subjects ANOVAs, with type of subject (recorder vs. roommate) as the only between-subjects factor.

Missing and Nonunique Events

Because of the difficulty most recorders would have had in recording events for their roommates over spring break, that week was omitted for all subjects and is treated as missing data (i.e., the week is included in the appropriate retention block) in all analyses except the one to be described here. In order to compare the frequency of occurrence of missing and nonunique events for recorders and roommates, the number of missing and nonunique events in the first 100 opportunities for such events (excluding spring break) was recorded for each subject. The period covered is 13.5 weeks (i.e., 12.5 weeks of recorded events, at 8 events per week, plus the week of spring break).

More events were missing for the roommates (mean = 11.1 events) than for the recorders (mean = 3.3 events) [$F(1,31) = 7.13$, MSe = 68.51]. Remember that the roommates were not informed about the procedure until just prior to the memory test. The recorders had to be careful about querying their roommates or the roommates would discover the purpose of their questions. Thus, in addition to the times when roommates were away on trips, the recorders occasionally had a difficult time generating appropriate instances for their roommates. That difficulty appeared in the analysis on the nonunique events as well. More nonunique events were generated for roommates (13.9) than for recorders (8.0) [$F(1,31) = 5.18$, MSe = 53.35].

Estimating Date of Occurrence

Dating error. In estimating the date of occurrence of events, all subjects occasionally made unusually large errors (from 45 days to as much as 90 days). For that reason, the median error (in days) was used to capture the performance for a block. The blocks consisted of 1 week each for the first two blocks (extending back from the day of testing) and 3 weeks each for the last four blocks. The median (absolute) error in estimating the date of occurrence was then computed for each retention block for each subject.

The effect of retention interval on dating error is shown in Figure 1. These data show a strong linear component in the effect, with each week increasing the magnitude of the median dating error by approximately 1 day.
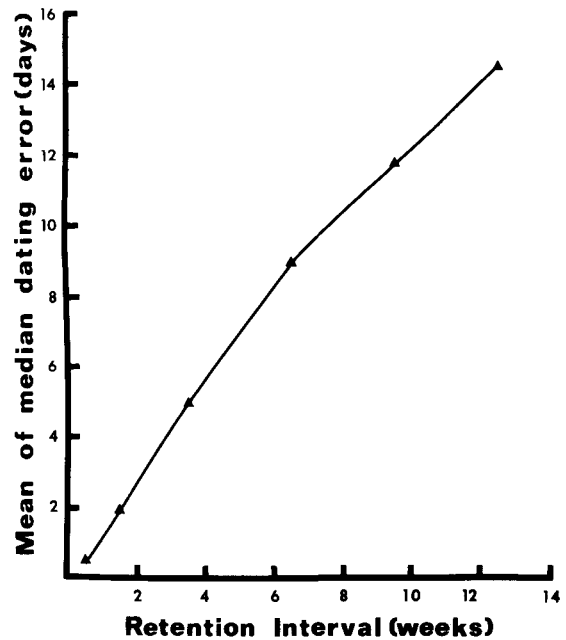

Figure 1. Median error (days) in estimating the date of occurrence of events as a function of retention interval.

The effect of retention interval was significant [$F(5,150) = 49.26$, MSe = 20.13]. There was no reliable effect of type of subject (recorder vs. roommate), nor was there a reliable Subject Type by Retention Interval interaction (Fs < 1.3).

Dating error and initial memorability rating. The subjects recording the events rated the perceived memorability of those events on a 3-point scale at the time the events were recorded. In order to assess the effect of perceived memorability on the accuracy of dating the event, the data were divided into two 7-week blocks. The median error in estimated date was then computed for each initial memorability rating (1, 2, or 3) in each block for each subject.

During the first 7-week block, the median error in estimating the date of an event increased as the perceived memorability of the event decreased [$F(2,60) = 5.81$, MSe = 14.10]. The median dating error was 6.5, 4.5, and 3.4 days for initial memorability ratings of 1, 2, and 3, respectively. Thus subjects were almost twice as accurate in estimating the date of events perceived to be extremely memorable as they were in estimating the date of events perceived to be unmemorable.

Note that an overall analysis of these data showed an effect only of retention interval [$F(1,30) = 58.90$, MSe = 50.54] (all other Fs < 2.1). Similarly, an analysis on the second 7-week block showed no effect of initial memorability rating (F < 1). The MSe was 37.83 for that effect in the second block analysis. The increase in error variance in the second block apparently was large enough to preclude detecting either a main effect or interaction involving initial memorability rating. Finally,

note that type of subject (roommate vs. recorder) did not produce a reliable effect in any of the analyses (Fs < 1.5).

Percent dated exactly. Because most subjects dated many events exactly and all subjects dated some events exactly, it is also possible to view the estimated date of occurrence data in terms of the percentage of events dated exactly. For that analysis, the data were partitioned into four 1-week blocks followed by five 2-week blocks. The percentage of items dated exactly was then computed for each subject for each block, and an analysis was performed on those data. The data from one subject were missing from one cell and were replaced as noted above.

The mean percent dated exactly for each retention interval is presented in Figure 2. These data show that the percent dated exactly dropped rapidly for the 1st month and fell more slowly thereafter. Between 5% and 10% of the items were such that they could be dated exactly after any interval (e.g., a birthday).

The drop in percent dated exactly was reliable $[F(8,240) = 35.75, MSe = 314.73]$. There was no reliable effect of type of subject and no reliable Subject Type by Retention Interval interaction (Fs < 1.6).

Correct on day but not date. In dating events, there may be some (or many) occasions for which the subject knows the day of the week on which the event occurred but does not know the date. In order to determine whether subjects tended to remember the day, but not the date, correctly, the percentage of errors in dating that were on the correct day was computed for each subject. This measure will be called the "correct day" measure. Note that the base level for the correct day measure would be expected to be 25%, because events were recorded for 4 days each week. These data showed no reliable difference between roommates and recorders $[F(1,30) = 1.99, MSe = 87.77]$. Further, the overall mean for the correct day measure was 28.1, which, given the variance noted above, is clearly not reliably different from the base level of 25.
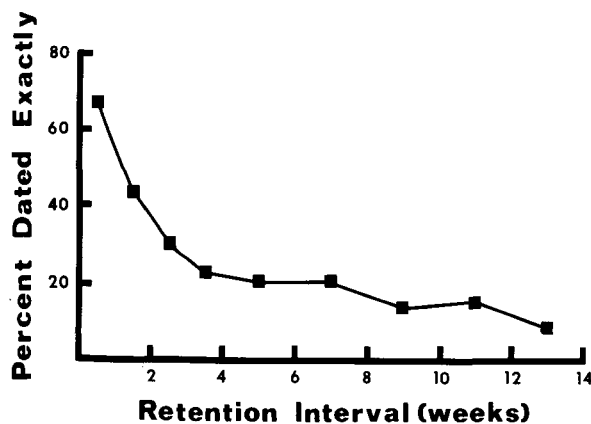

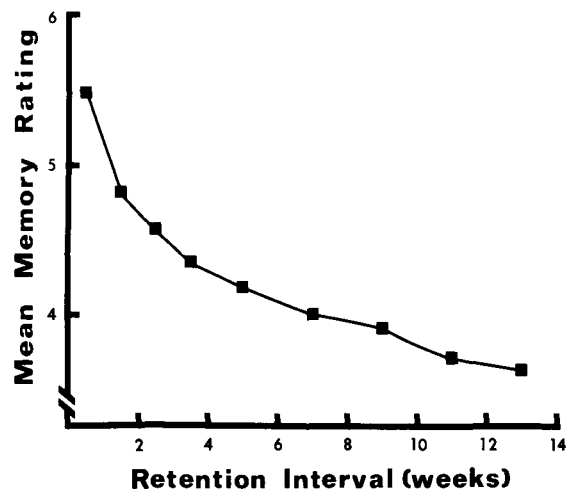Figure 2. Mean percent dated exactly for each retention interval.

Figure 3. Mean memory rating for each retention interval.

There were great individual differences in the correct day measure. Although the correct day measure and median dating error are logically independent, it seemed likely that subjects with high overall accuracy would also tend to choose the correct day when they did make an error. To check this hypothesis, we looked at the correlation between median dating error and the correct day measure. As anticipated, there was a moderate correlation between the two measures that was statistically reliable $[r(30) = -.462, p < .01]$.

Memory Ratings

Overall retention. The memory ratings at the time of test were on a 7-point scale, as follows: not at all (1), barely at all (2), not so well (3), fairly well (4), very well (5), almost perfectly (6), and perfectly (7). To evaluate changes in the subject's memory ratings over time, the retention interval was divided into nine blocks, with 1 week in each of the first four blocks and 2 weeks in each of the remaining five blocks. The mean memory rating for each block was then computed for each subject. The entry for one cell for one subject was missing and was computed as indicated above.

The mean memory rating for each retention interval is shown in Figure 3. As can be seen, the memory rating data showed a classical Ebbinghaus function, with an initial rapid decline and a slower reduction thereafter.

The change in ratings over retention interval was reliable $[F(8,240) = 42.21, MSe = .277]$. No other main effect or interaction was significant (Fs < 1.5).

Retention as a function of initial memorability rating. To evaluate the effect of initial memorability rating on memory rating at the time of test, the memory rating data were divided into four retention blocks. The first block contained 2 weeks and the last three blocks each contained 4 weeks. The memory rating for each initial rating in each block was calculated for each subject. Two subjects (one recorder and one roommate) were discarded

because one subject used no "1" preratings and the other used no "3" preratings. In addition, eight data cells were missing and were replaced with the procedure described earlier. The mean memory ratings for each prerating at each retention interval are shown in Table 1.

These data show that the memory ratings for both recorders and roommates increased as the initial memorability ratings increased. However, the mean memory ratings for the records covered a wider range than the mean ratings for the roommates. As predicted, the recorder could judge the memorability of personal events more accurately than the memorability of events recorded for a roommate.

The Subject Type by Initial Memorability Rating interaction described above was reliable [$F(2,56) = 8.20$, $MSe = .339$]. Moreover, the mean memory rating increased as initial memorability rating increased [$F(2,56) = 92.40$, $MSe = .339$]. Consistent with the previous analysis, mean memory rating decreased as the retention interval increased [$F(3,84) = 74.17$, $MSe = .501$]. The mean memory ratings were 5.13, 4.35, 3.89, and 3.68 for the first through the fourth retention intervals, respectively. The Retention Interval by Type of Subject interaction approached significance [$F(3,84) = 2.51$, $p < .05$], but the variation in the means was not systematic. No other main effect or interaction was statistically reliable in this analysis ($Fs < 1.4$).

Effect of Rehearsals

As noted in the Procedure section, subjects rated rehearsals on a 6-point scale having both numeric and descriptive labels. Specificially, a rating of 1 was given for "never (0 rehearsals)," 2 for "rarely (1-2 rehearsals)," 3 for "a few times (3-5)," 4 for "several times (6-10)," 5 for "quite often (11-20)," and 6 for "very often (over 20 rehearsals)." A tabulation of the data showed that the rehearsal ratings occurred with different frequencies for different initial memorability ratings. Subjects tended to give rehearsal ratings of 1 or 2 for events with an initial memorability rating of 1 and tended to use rehearsal ratings greater than 2 for events prerated as 3. Hence, in order to be able to look at retention interval and initial memorability rating together with rehearsal rating without an inordinate number of missing data cells, rehearsal ratings were blocked into 1-2 ratings vs. 3-6 ratings. The mean number of remembered events occurring in each rehearsal rating block for each initial

Table 1
Mean Memory Rating for Each Memorability Prerating
Presented Separately for Each Type of Subject

Subject Type	Memorability Prerating			
	1	2	3	Mean
Recorders	3.66	4.32	4.97	4.32
Roommates	3.88	4.12	4.60	4.20
Mean	3.77	4.22	4.79	4.26

Table 2
Mean Number of Remembered Events by Memorability
Prerating and Rehearsal Block

Rehearsal Block	Memorability Prerating			
	1	2	3	Total
1-2	12.7	15.5	7.3	35.5
3-6	8.8	17.2	22.7	38.7
Total	21.5	32.7	30.0	84.2

Table 3
Mean Memory Rating for Each of Two Rehearsal Rating Blocks
Under Each of Two Retention Interval Blocks

Rehearsal Rating	Retention Interval (in Weeks)		
	1-7	8-14	Mean
1-2	4.37	3.50	3.94
3-6	4.85	4.37	4.61
Mean	4.61	3.94	4.27

memorability rating category is shown in Table 2. Items rated as totally forgotten were not given a rehearsal rating.

To investigate the effect of number of rehearsals, the memory rating data were partitioned into 12 blocks representing all combinations of retention interval block (two 7-week blocks), initial memorability rating (1, 2, or 3), and rehearsal block (1-2 rating vs. 3-6 rating). The mean memory rating was then computed for each subject for each of the 12 blocks. Three subjects (one recorder and two roommates) had to be eliminated because a substantial number of data points were missing in their data. Even after removing the three subjects, 17 data points were missing and were computed in the manner described earlier.

The major finding of the analysis performed on these data was a reliable Retention Interval by Rehearsal Block interaction [$F(1,27) = 16.50$, $MSe = .196$]. The means for that interaction are shown in Table 3.

As can be seen, the retention loss was smaller for those events that were rehearsed more often. As predicted, an increase in rehearsal decreased the rate of forgetting.

Table 3 also shows that the mean memory rating was higher for those items that were rehearsed more often [$F(1,27) = 55.77$, $MSe = .694$]. Consistent with previous analyses, the mean memory rating for the block representing the shorter retention interval was higher (4.61) than that for the longer retention interval (3.94) [$F(1,27) = 88.85$, $MSe = .441$]. Once again, there was a reliable effect of initial memorability ratings [$F(2,54) = 56.45$, $MSe = .366$], with the mean memory rating increasing as the initial memorability rating increased. The mean memory rating was 3.90, 4.19, and 4.73 for initial memorability ratings of 1, 2, and 3, respectively.

There was no effect of type of subject ($F < 1$), nor was any other interaction reliable ($Fs < 2.7$).

Predicting Memory Performance by SIME

As a first attempt to predict performance on memory for unique personal events, all participants filled out the SIME developed by Herrmann and Neisser (1978). The responses to SIME may be partitioned into three major classes that produce scores for overall forgetting (OF), memory for recent events (R3), and memory for remote events (R1). (SIME and its subsections are described in detail in the Herrmann and Neisser paper). A comparison of roommates and recorders showed no reliable differences for any of these measures (Fs < 1). The overall mean scores were 4.63, 4.64, and 4.96 for OF, R3, and R1, respectively. The values for OF and R3 are quite comparable to the norms of 4.74 and 4.80, respectively, reported for those measures by Herrmann and Neisser. The mean for R1 (4.96) is somewhat higher than the norm (4.32) reported by Herrmann and Neisser, but it is well within the confidence limits suggested by their data (i.e., they report standard deviations greater than 1 for each of the four R1 items).

In order to determine how well these SIME measures predicted performance, the participants' scores on these three measures were correlated with their mean memory ratings, mean (over blocks) of the median dating error, and mean percent dated exactly. The recorder and roommate scores were pooled for purposes of computing these correlations. For the OF measure, the correlations ranged between −.001 and .005 (df = 30) and, of course, were not statistically significant. For the R3 measure, the correlations ranged between −.163 and .238 (df = 30) and, again, were not significant. In the case of the R1 measure (memory for remote events), the correlation with mean memory rating [r(30) = .326] was modest but was statistically significant by a one-tailed test (p < .05). The other correlations with R1 ranged from −.251 to .024 (df = 30) and were not statistically significant.

In short, the partitions of SIME used in this study do not clearly predict memory for unique personal events. There is a relationship between the SIME measure of memory for remote events and mean memory rating for the events recorded in this experiment, but it is weak at best.

Informal Observations

Following the recording of data described above, all participants received a copy of their responses and dating accuracy. Virtually all participants were eager to discuss their memory performance. These informal conversations produced two observations that are worth noting.

First, most participants indicated that they organized events in time with respect to certain key events. In the most dramatic instance, one participant who had broken with her boyfriend indicated that she thought of events as "before Jeff and after Jeff." The notion that episodic memory is organized around landmark events is not new.

However, the comments of these participants make it apparent that most subjects believe they order events in time using key events. Whether that organizational strategy can be demonstrated in some objective way remains to be seen.

Second, those participants who were particularly accurate in dating events identified themselves as individuals who either used the calendar a great deal for keeping track of their schedules or, for other reasons, were able to specifically date a large number of events. In the second category, for example, a pitcher for the Kansas State baseball team knew the exact dates of each game and, thus, was able to date any game event. In addition, he indicated that other events tended to be organized around those dated events (e.g., "That took place during the week before the Missouri game.").

DISCUSSION

Our first objective was to determine whether recording events improves memory for those events. Clearly, several laboratory phenomena all lead to the prediction that recorders would perform better than roommates. Surprisingly, in spite of the fact that the recorders selected the events, recorded the events, and knew they were going to be tested for memory of the event, their dating accuracy and rated memory for events were not reliably better than those of their roommates. While these data cannot be used to argue that there is no effect of recording events, the data strongly suggest that any such effect is, at best, very small.

The failure to find a difference between roommates and recorders has a very nice methodological implication. Specifically, the data indicate that a simpler single-subject diary procedure will work as well as the present procedure in estimating the dating accuracy of, or rated memory for, personal events.

Our next question was whether the rate of forgetting would be the same for events rated as memorable and events rated as unmemorable. Although the events recorded for this study were far more heterogeneous than are those used in laboratory studies, the data were completely consistent with previous laboratory findings. That is, the data suggest that the degree of encoding differs for memorable and unmemorable events but that the rate of forgetting thereafter is the same for both types of events. However, the initial retention block takes in a relatively long period (2 weeks). Such a block might not detect a condition in which memory for memorable and unmemorable events starts at an equivalent level and then diverges rather rapidly for a relatively short time (e.g., over 24 h), with the rate of forgetting equivalent thereafter. While these data allow for that possibility, it is not very probable, given that the vast majority of laboratory data support the hypothesis that memories encoded at different levels are forgotten at the same rate (e.g., Underwood, 1964).

An important methodological concern with the memory rating procedure used in this study is that subjects might be basing their initial memorability ratings, test memory ratings, and rehearsal ratings on the description of the event rather than on memory of the event (or rehearsal of the event). That is, the subject's rating could be, essentially, an evaluation of whether that event should have been remembered (or rehearsed). That would, of course, explain the failure to find any difference between recorders and roommates. The systematic change in memory ratings over retention intervals greatly reduces, if not completely eliminates, that methodological concern. If test memory ratings were based entirely on the description of the event rather than on memory for the event, there would be no effect of retention interval. Even putting aside that obvious argument, it would be difficult to understand how memory ratings based on the description of an event, rather than on memory for the event, would produce the classical Ebbinghaus forgetting function. Thus, the evidence strongly indicates that the memory ratings in this experiment were based on memory for the event.

Our third question regarded the accuracy with which recorders could estimate the recallability of their own events and their roommates' events. As it turned out, the only reliable difference involving roommates and recorders was an interaction reflecting the fact that the recorder could more accurately assess the memorability of personal events (at the time they were recorded) than the memorability of events recorded for a roommate. It is not particularly surprising to find that individuals can assess their memory for a personal event better than they can assess someone else's memory for a personal event. At the same time, that result does raise the interesting possibility that the accuracy of such assessment might be a useful measure of social sensitivity.

The fourth question was whether the clarity (or strength) of a memory is related to the accuracy of dating an event. These data show that memory for when the event occurred and memory for what occurred are related for the first few weeks but then become relatively independent (i.e., they appear to be related only in that one cannot date an event that is totally forgotten). The change over time may occur because time and event memory are lost at different rates. Although the distinction between time and event memory is very speculative, one of the most interesting findings of the study was the strong linear component in the rate of forgetting for the dating error data. Over the range tested, subjects increased their dating error by about 1 day for each week of the retention interval. This outcome stands in sharp contrast to the general assumption that the classical Ebbinghaus function is the standard for modeling forgetting functions.

This study also showed that those subjects most accurate in dating events were most likely to be partly right when they made an error. That is, one possibility in dating an event is that one might remember the day of the week on which the event occurred but not remember the date. We checked for this possibility by looking at the proportion of correct day judgments when an error was made on the date. There was no overall tendency to get the day, but not the date, right. However, there was a moderate (and reliable) tendency for overall accuracy in dating events to be positively related to getting the day of the week correct on dating errors.

Next, we asked about the effect of rehearsals on memory. The present data suggest that increased rehearsals slow the rate of forgetting in addition to increasing memory for the event. The increase in memory produced by rehearsal is well-known, but the reduction in rate of forgetting when rehearsal is spread over time is not. The most relevant study is by Landauer and Bjork (1978). In that study, they showed that an increasing interval between successive rehearsals produced better memory performance than did equally-spaced rehearsals. It is likely that naturally-occurring rehearsal shows a pattern of increasing intervals between rehearsals.

Note that the interaction described above also suggests that subjects are not basing their rehearsal estimates simply on memory for the event. If that were the case, then rehearsal ratings should parallel memory ratings, producing parallel functions, as in the memory rating data.

Finally, we were interested in predicting memory for personal events with a laboratory test (such as word-list memory) or a self-rating measure. As a first attempt, we administered the self-rating scale (SIME) developed by Herrmann and Neisser (1978) to all the subjects in the experiment. This attempt to predict memory performance was not very successful. There is a modest relationship between one of the subscales of SIME and rated memory, but that relationship is simply not strong enough to be of much use in predicting memory performance. As is often the case, participants' judgments about their abilities are not closely related to their performance.

Let me conclude by suggesting several reasons for investigating naturally occurring episodic memory. First, we need to establish the ecological validity of our laboratory findings (e.g., Brunswik, 1956; Neisser, 1978). It is not a foregone conclusion that results from laboratory studies generalize to memory in a natural setting. For example, we expected to find a difference in performance between recorders and roommates and did not. Further, the almost linear forgetting function found for the dating error data stands in sharp contrast to the Ebbinghaus type of function generally assumed to be the model for forgetting.

Second, we can look at phenomena that may be difficult or impossible to bring into the laboratory. Consider, for example, the potential effect on memory of

the emotional content of many of the events reported in this study. It is not feasible to artificially create the equivalent of the emotional content of an event such as "Someone who was in my life in a very special way told me they no longer wanted to be."

Third, the data base for naturally occurring episodic memory is extremely small and needs to be expanded so that theories of memory can be validated against data drawn from everyday experience.

Finally, the present study used a different set of events for each individual, with the events ranging dramatically in their importance and presumed memorability. Clearly, the data base for naturally occurring episodic memory describes the kind of performance that all of us engage in every day. The data, therefore, have personal as well as psychological significance.

REFERENCES

BADDELEY, A. D., LEWIS, V., & NIMMO-SMITH, I. When did you last . . . ? In M. M. Gruneberg, P. E. Morris, & R. M. Sykes (Eds.), *Practical aspects of memory.* New York: Academic Press, 1978.

BOWER, G. H., & GILLIGAN, S. G. Remembering information related to one's self. *Journal of Research in Personality,* 1979, 13, 420-432.

BRUNSWIK, E. *Perception and the representative design of psychological experiments.* Berkeley: University of California Press, 1956.

CRAIK, F. I. M., & WATKINS, M. J. The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior,* 1973, 12, 599-607.

CROVITZ, H. F., & SCHIFFMAN, H. Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society,* 1974, 4, 517-518.

HERRMANN, D. J., & NEISSER, U. An inventory of everyday experiences. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory.* New York: Academic Press, 1978.

LANDAUER, T. K., & BJORK, R. A. Optimum rehearsal pattern and learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory.* New York: Academic Press, 1978.

LINTON, M. Memory for real-world events. In D. A. Norman & D. E. Rumelhart (Eds.), *Exploration in cognition.* San Francisco: Freeman, 1975.

NEISSER, U. Memory: What are the important questions? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory.* New York: Academic Press, 1978.

NELSON, T. O., & VINING, S. K. Effect of semantic versus structural processing on long-term retention. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, 4, 198-209.

ROGERS, T. B., KUIPER, N. A., & KIRKER, W. S. Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology,* 1977, 35, 677-688.

THOMSON, R. H. An experimental study of memory as influenced by feeling tone. *Journal of Experimental Psychology,* 1930, 13, 462-467.

UNDERWOOD, B. J. Degree of learning and the measurement of forgetting. *Journal of Verbal Learning and Verbal Behavior,* 1964, 3, 112-129.

UNDERWOOD, B. J. *Temporal codes for memories: Issues and problems.* Hillsdale, N.J: Erlbaum, 1977.

WAGENAAR, W. A. Recalling messages broadcast to the general public. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory.* New York: Academic Press, 1978.

WARRINGTON, E. K., & SANDERS, H. I. The fate of old memories. *Quarterly Journal of Experimental Psychology,* 1971, 23, 432-442.