Scale effects in memory for the time of events

WILLIAM J. FRIEDMAN
Oberlin College, Oberlin, Ohio

and

ARNOLD J. WILKINS
Medical Research Council Applied Psychology Unit, Cambridge, England

This study addressed the question of how people remember the time of past events. Stimuli were 10 news events that had occurred from 6 months to 20 years before the study. In contrast to previous studies of memory for time, subjects were asked to provide estimates of the stimulus events on multiple time scales, including year, month, day of the month, day of the week, and hour. If judgments are based on direct information about the age of the memory, accuracy should decrease monotonically as one moves to finer scales. Alternatively, if subjects reconstruct the time from fragmentary information associated with the event, one would expect that estimates on finer time scales would often exceed grosser scales in accuracy. Results for accuracy, confidence, and number of recall cues supported the latter position. In addition, subjects reported a variety of types of recall cues, the most common being memory for personal experiences or events that were contiguous with the news event.

In contrast to the progress made in describing many memory phenomena, our understanding of how people answer "when" questions is still quite poor. We know that elapsed time is a potent influence on the ability to remember, but we know little about how temporal information is extracted when people try to recall the time of an event. In some memory models (e.g., Lindsay & Norman, 1977), time is a concept attached to event nodes, but the nature of the information embedded in that concept is far from clear. Among the possible accounts of the storage of time information in memory are (1) that events are organized in memory in a time-ordered format, (2) that explicit time tags are laid down at the time of encoding, (3) that decay of trace strength provides time information, (4) that events are interrelated by order codes, and (5) that idiosyncratic contextual information associated with a trace is used to deduce the time of events (see Hintzman, Block, & Summers, 1973, Linton, 1975, Tzeng & Cotton, 1980, and Underwood, 1977, for discussions).

Most of the research relevant to the issue has involved asking subjects to judge the serial position of individual words or the relative recency of pairs of words from sequentially presented lists. These laboratory studies generally are designed to minimize external information about the time of presentation so that models such as (1), (2),

and (3) can be evaluated. The research indicates considerable sensitivity to relative time of presentation even when subjects do not know in advance that temporal judgments will be required. However, a number of findings are troublesome for Models 1, 2 and 3, including a strong "primacy effect" on the accuracy of judgments. This and other findings (see, e.g., Guenther & Linton, 1975, Hintzman et al. 1973, Tzeng & Cotton, 1980, and Underwood, 1977) suggest that incidental contextual information plays an important role in judging time of occurrence.

Because of apparent context effects even in the relatively uniform test situations and because of the very short time scales used, there is reason to doubt the relevance of the serial learning studies to everyday recall of the time of events. It is therefore important to consider studies in which subjects judge the time of events experienced outside the laboratory. In seven studies, subjects were asked to recall the date of public or personal events (Baddeley, Lewis, & Nimmo-Smith, 1978; Ferguson & Martin, 1983; Lieury, Aiello, Lepreux, & Mellet, 1980; Lieury, Caplain, Jacquet, & Jolivet, 1979; Linton, 1975; Thompson, 1982; Underwood, 1977). The main aim of most of these studies, and one that was clearly fulfilled, was to demonstrate the usual forgetting curve in a context outside the laboratory. Absolute error in judgments of the date on which events occurred was found to be a positive function of the length of the retention interval. Unfortunately, the forgetting curve itself provides little information about the processes involved in dating events. For example, it seems compatible with all of the accounts mentioned above. Another common finding is the underestimation of the age of older events in the ranges tested, a finding that seems similarly neutral with respect to these models. However, introspective reports in the Baddeley

Part of this work was conducted while the first author was Visiting Scientist at the Medical Research Council Applied Psychology Unit and received additional support from a research and development grant from Oberlin College. Thanks are due Jo Whitehead for her patient coding of the responses and to Alan Baddeley for his helpful comments on a previous draft of the paper. Requests for reprints should be sent to William J. Friedman, Department of Psychology, Oberlin College, Oberlin, OH 44074.

et al., Ferguson and Martin, Lieury et al. (1980), Linton, and Thompson studies appear to narrow the range of possible models by suggesting that often information about time is not directly available for the target event. Instead, subjects seem to remember fragments of the event and to use this information, frequently coupled with personal "landmark" events, to reconstruct when the event must have occurred. If subjects had direct access to time information, as several accounts imply, it seems unlikely that they would engage in so circuitous a method.

The introspective data, of course, are insufficient for us to decide between direct access models, such as Models 1, 2 and 3, and reconstructive ones. One approach that may distinguish between them comes from altering the tasks used in the previous studies and requiring separate estimates on different time scales. In each of the previous studies, subjects were asked to give a date for each event. The particular information requested varied from study to study. For example, Underwood (1977) asked subjects to write the month and year, Thompson (1982) asked for the date and day of the week and provided a blank calendar, and Baddeley et al. (1978) had subjects circle the year and month and indicate the day of the month by marking a point on a 10-cm line. But with the exception of one analysis in Thompson's study (showing that day of the week was no better recalled than calendar date), time estimates were treated as single points on a temporal continuum. This is convenient for describing forgetting curves but obscures the multiple scales on which humans apparently experience, and perhaps remember, time. Conventional time systems, human activities, and, to some extent, the natural environment are structured on time scales that include parts of days, days of the week, days of the month, months of the year, and entire years. If memory for the time of events is a reconstruction coupling fragmentary information from the trace with other knowledge about time, we might expect that accuracy would vary from scale to scale. For example, an estimate might be inaccurate by years but be quite accurate for month, or it might be inaccurate for day of the week or day of the month but be accurate for time of day. Such a finding would be difficult for models based on the age of the memory, such as Models 1 and 3. If age of the

memory contributes the critical information for judgments of the date, one would expect finer scale information to be less accurate. By analogy, if an estimate of duration is given by the method of reproduction, the timer should show greater accuracy in seconds than in tenths of seconds.

In this study, we examined scale effects in estimates of the time of events. Subjects attempted to recall the year, month, day of the month, day of the week, and hour of 10 public events of the preceding 20 years. In addition to producing estimates on each time scale, subjects were asked to rate their confidence in the estimates and to report how they arrived at them. If judgments are based on information about the age of the memory, one would expect dramatic decreases in accuracy and confidence as one moves to finer time scales. Violations of a pattern of monotonic decreases would be consistent with a reconstructive process. Although there is no obvious basis for predicting the likelihood of exploiting particular scales, we suspected that time of day and time of year would be especially well remembered. Therefore, we expected greater accuracy, confidence, and numbers of recall cues for month and hour than would appear in a scale decrement pattern.

EXPERIMENT 1

Method

Subjects. Nineteen women served in the study. They ranged in age from 32 to 66 years, with a mean age of 50. All were members of the Applied Psychology Unit subject panel and resided in or near Cambridge, England.

Materials and Procedure. Ten news events were chosen for subjects to recall. The particular events were selected because (1) they were punctate and (2) they were of sufficient significance for most people to have been aware of the events soon after they had occurred. The events, their dates, and their approximate times of occurrence are given in Table 1. We cannot be certain, of course, when or if subjects actually learned of the events, and it is not safe to assume that subjects lacked subsequent exposure to information about the time of the events.

The description of each event (Table 1) was listed on the top of a sheet of paper, and sheets were ordered randomly. Below were five rows labeled for each scale: year, month, day of the month, day of the week, and hour. Two events were assigned randomly

Table 1
Stimulus Events for Experiment 1 and Times of Occurrence

Event Description	Day of Week	Date	Approximate G.M.T.
1 John F. Kennedy is assassinated.	Friday	November 22, 1963	19:00
2 Donald Campbell is killed in a speedboat accident on Lake Windermere.	Wednesday	January 4, 1967	9:00
3 The first spaceman sets foot on the moon.	Monday	July 21, 1969	3:00
4 Lord Mountbatten is killed by an explosion aboard his boat.	Monday	August 27, 1979	12:00
5 Former Beatle John Lennon is shot.	Tuesday	December 8, 1980	4:00
6 Pope John Paul II is shot.	Wednesday	May 13, 1981	16:00
7 Prince Charles and Lady Diana exchange rings.	Tuesday	July 29, 1981	12:00
8 Egyptian president Anwar Sadat is assassinated.	Thursday	October 6, 1981	13:00
9 Prince William is born.	Tuesday	June 22, 1982	21:00
10 The ship <i>Mary Rose</i> rises above the surface.	Monday	October 11, 1982	9:00

to each of five scale orders for each subject. The five orders on the page were random, with the restriction that each scale appear once in each of the five positions.

In each row, the subjects were asked to estimate on the appropriate scale the time when the event happened. They were also asked to indicate their confidence in this estimate by marking one of five points along a line whose ends were labeled "pure guess" and "very confident." To the right of the estimate and confidence rating was a space provided for subjects to "List the things you thought of in arriving at your estimate." At the bottom of each page, the subjects were asked whether or not they remembered what they were doing when they learned of the event.

Group testing took place during May 1983, in a single session lasting approximately 50 min. After they had been acquainted with the task, the subjects were additionally instructed to work in the order given, to leave no blanks—to guess if they had to, to give estimates of when the event happened (in British time), and to give exact estimates and not phrases.

Scoring. Time estimates on each scale were converted to error scores in the following way. Year error score was the absolute value of the estimate minus the actual year. For the other four scales, error scores were the proportion of the deviations expected by chance. These were calculated by finding the temporal distance from the estimate to the nearest occurrence of the actual time and dividing this by the distance that would be expected by chance alone. For example, if a subject estimated the month of the Kennedy assassination to be February, the distance would be 3 months, the chance expectation is always three months, and so the proportion error would be 1.0. For an estimate of December or October, the proportion error would be .33. Chance deviations for day of the month, day of the week, and hour are 7.5 (assuming 30-day months), 1.71, and 6 respectively. The proportion deviation calculations employ the simplifying assumption that the estimate on a given scale is always a deviation from the closest occurrence of the actual time on that scale—estimates on different scales are calculated independently. The year error score cannot be expressed as a proportion because there is no reasonable way of specifying the maximum range of possible years.

Confidence ratings were converted to integers from 1 ("pure guess") to 5 ("very confident"). Marks between reference points were rounded down.

Recall cues were recorded as present or absent for each scale for each event and, when present, were assigned to one of five categories. In Category 1, the subjects reported that they remembered a personal experience that was contiguous with the target event or that the target occurred at approximately the same time as another personally meaningful event whose time was known. Examples from

pilot studies include "I was attending a once annual party," "It was wash day," and "It was within a few weeks of the birth of my first child." Category 2 included instances in which some aspect of the event itself was cited, for example, "They were wearing light clothes" or "The Prince stayed up all night." In Category 3, some logical constraint was placed on when the event could have occurred. An example is "They would have picked good weather for the launch." In Category 4, the subjects reported that the target was some specified or implicit interval before or after another event whose time was known, for example, "(Prince William was born) about a year after the July wedding." Category 5 consisted of recall cues that were not assigned to the other categories.

Results and Discussion

The first set of analyses concerned the effect of time scale on the accuracy of estimates. Mean absolute error for year and mean proportion error for month, day of the month, day of the week, and hour are presented in Table 2. The four proportion error scores were subjected to a repeated-measures ANOVA with event and scale (month, day of the month, day of the week, and hour) as crossed, within-subject factors. Proportion error cell means were substituted for missing data points for the approximately 5% of time estimates that were incomplete or left blank on response sheets, and the data of two subjects who were missing more than half of their estimates were deleted from the analysis. Event [$F(9,144) = 7.44$, $p < .0001$], scale [$F(3,48) = 24.18$, $p < .0001$], and their interaction [$F(27,432) = 6.36$, $p < .0001$] were all highly significant factors. Nonmonotonicity of scale was shown by a planned comparison for quadratic trend [$t(48) = 2.63$, $p < .02$].

Table 2 shows substantial inconsistency with a scale decrement pattern. First, one would expect month estimates to be accurate only when the corresponding year estimates are accurate. For the 2 events for which year error is less than .25 (Events 7 and 9), month error is significantly below chance ($ps < .05$). However, for 5 of the remaining 8 events (Events 3, 4, 5, 6, and 10), month estimates are also significantly below chance ($ps < .05$). Second, as one moves from month to day of the

Table 2
Mean Error of Time Estimates for Each Scale for Each Event in Experiment 1

Event	Actual Year	Absolute Deviation Year	Proportion of Chance Deviation				Mean of the Four
			Month	Day of Month	Day of Week	Hour	
1 (Kennedy)	1963	1.75	.98	.79	.62**	.64**	.76
2 (Campbell)	1967	3.78	1.51	1.08	.86	.57**	1.00
3 (Moon)	1969	2.67	.55***	.93	1.06	1.14	.92
4 (Mountbatten)	1979	.71	.10***	.93	1.10	.41***	.64
5 (Lennon)	1980	1.35	.76*	1.18	.86	1.34	1.04
6 (John Paul II)	1981	.76	.65**	.82	1.10	.45***	.76
7 (Wedding)	1981	.06	.04***	.77	.99	.14***	.48
8 (Sadat)	1981	.94	.92	1.32	.93	.31***	.87
9 (Prince)	1982	.06	.23***	.76	.65**	.85	.62
10 (Mary Rose)	1982	.29	.73*	.73*	.75	.55***	.69
Mean		1.24	.65	.93	.89	.64	

Note—Significance levels are for one-tailed t tests with the null hypothesis that entries are equal to or greater than chance deviations. Italicized entries are also significantly lower than the corresponding deviation scores in Experiment 2.

* $p < .05$. ** $p < .01$. *** $p < .001$.

month and day of the week to hour, one would expect only increases in errors. However, both month and hour are substantially below day of the month and day of the week in proportion error. The scale pattern is not an artifact resulting from the inclusion of events from different time zones. The scale means for the British events (Events 2, 4, 7, 9, and 10) showed the same scale-to-scale pattern as the set as a whole.

Variation in the accuracy of recall of an event appears to reflect the importance of the events to the subjects as well as the age of the events. The three events involving the royal family (Events 4, 7, and 9), for example, are dated more accurately than most intervening events. In addition, older events are dated less accurately for year [$r(8) = -.81, p < .01$] and month [$r(8) = -.56$]. Hour [$r(8) = -.21$], day of the month [$r(8) = .01$], and day of the week [$r(8) = .22$] fail to show forgetting curves. The greater the proportion of subjects who reported remembering what they were doing when they learned of a particular event, the more accurate the month estimates were [$r(8) = .64$]. But for the other scales, the correlations were less than .40. Age of the event was not the mediator of the correlation with month estimate, since it was only very weakly correlated with remembering one's activity [$r(8) = .22$].

The second dependent variable, confidence, showed a pattern of scale-to-scale variation similar to that of accuracy. Table 3 presents the mean confidence ratings, with low scores representing low levels of confidence in the corresponding estimate. Confidence ratings were subjected to a repeated-measures ANOVA with event and scale (year, month, day of the month, day of the week, and hour) as crossed, within-subject factors. Cell means were substituted for the approximately 8% missing data points. Event [$F(9,162) = 19.31, p < .0001$], scale [$F(4,72) = 43.02, p < .0001$], and their interaction [$F(36,648) = 3.20, p < .0001$] were all highly significant. Examination of the means column in Table 3 reveals that confidence decreases from year to month to day of the month and day of the week scales, consistent with a scale decrement pattern. But confidence for hour estimates is greater than that for day of the month and day

of the week estimates and only slightly lower than that for month estimates. A planned comparison showed a significant quadratic trend for scale [$t(72) = 3.21, p < .005$]. The last columns of Tables 2 and 3 are highly correlated [$r(8) = -.92, p < .01$], suggesting that event variation in accuracy and confidence are attributable to similar factors.

The results for the third dependent variable, recall cues, are shown in Table 4. The subjects reported recall cues on a total of 389 of 950 opportunities. This total is unevenly distributed by scale as well as by cue type. The column totals in Table 4 show a scale pattern similar to that of confidence. Year and month had frequent recall cues, day of the month and day of the week had few, and hour fell in between. The scale variation was tested statistically by coding each opportunity to report a cue as 1 (cue reported) or 0 (item left blank). The repeated-measures ANOVA had the same design as the ANOVA for confidence. Both scale [$F(4,72) = 29.78, p < .0001$] and event [$F(9,162) = 4.13, p < .001$] were significant, but their interaction was not. A significant quadratic trend [$t(72) = 2.45, p < .01$] was found, again supporting nonmonotonicity in the scale effect.

The row totals of Table 4 indicate the important role that remembered experiences and personal referents (Category 1) play in the dating of events. Similar methods have been reported in previous studies (Baddeley et al., 1978; Lieury et al., 1980; Linton, 1975; Thompson, 1982), and Baddeley et al. showed that their use is positively related to accuracy (see also Loftus & Marburger, 1983). The second most frequent response was Category 3, indicating that the task was not treated strictly as a memory task. Instead, the subjects appear to have used a variety of types of information to constrain the time that events could have occurred. Category 4 also illustrates the use of external information, in this case exploiting temporal patterns. For example, more than half of the subjects referred to the time of the royal wedding in explaining their estimates of the month in which Prince William was born.

The results of this study were consistent with the predictions of the reconstructive model. However, an assump-

Table 3
Mean Confidence Ratings for Each Scale for
Each Event in Experiment 1

Event	Scale					Mean
	Year	Month	Day of Month	Day of Week	Hour	
1. (Kennedy)	2.41	1.82	1.22	1.83	2.06	1.87
2. (Campbell)	1.47	1.24	1.06	1.00	1.37	1.23
3. (Moon)	1.83	1.53	1.00	1.23	1.38	1.39
4. (Mountbatten)	2.74	2.72	1.47	1.94	2.42	2.26
5. (Lennon)	2.06	1.69	1.35	1.12	1.65	1.57
6. (John Paul II)	2.22	1.41	1.12	1.35	1.68	1.56
7. (Wedding)	3.44	3.33	2.44	2.71	2.67	2.92
8. (Sadat)	2.58	1.47	1.18	1.17	1.71	1.62
9. (Prince)	3.77	3.32	1.79	1.59	2.17	2.53
10. (Mary Rose)	3.33	2.35	1.47	1.82	2.24	2.24
Mean	2.58	2.09	1.41	1.58	1.94	

Table 4
Number of Recall Cues Falling in Each Category for Each Scale in Experiment 1

Cue	Scale					Total
	Year	Month	Day of Month	Day of Week	Hour	
1 Personal experience or event	48	44	18	26	34	170
2 Some aspect of event itself	4	12	0	4	22	42
3 Logical constraint	18	24	4	7	19	72
4 Interval before or after another event	27	17	6	1	7	58
5 Other	17	7	4	12	7	47
Total	114	104	32	50	89	

tion of this model raises a possible problem for our interpretation of the scale effects in accuracy. Under the reconstructive model, it is assumed that subjects have access to general information about when particular types of events typically occur. The findings for recall cues indicate that subjects do use this information in estimating the time of events. If this is true, it seems likely that such information would be differentially available for different time scales. For example, weddings typically take place during the daytime and during the spring or summer but can occur on any day of the month. In this example, simply exploiting what one knows about weddings in general could lead to a more accurate estimate on the month and hour scales than on the day-of-the-month scale. Although such a phenomenon would be interesting in itself, it would make it difficult to draw conclusions about the accuracy of memory.

The following experiment was conducted to test the possibility that the scale effects in accuracy in Experiment 1 are attributable solely to general time knowledge. A new group of subjects were asked to imagine that they were writing a novel about each of a series of fictitious events. Each of these events was supposed to provide a generic description of a particular news event that we had used in the original study. The subjects were required to choose the time of day, day of the week, day of the month, month, and year in which they would have this event take place. These times allowed us to estimate the extent to which accuracy in the memory task could be attributed to general time knowledge.

EXPERIMENT 2

Method

Subjects. Twenty-seven members of the Applied Psychology Unit subject panel (23 women and 4 men aged between 22 and 67 years, mean age = 53) served in this experiment. One failed to understand the instructions, and her data were discarded.

Materials and Procedure. The subjects were tested on successive days as two groups of 12 and 15 people. Each subject received the following written instructions. "Please imagine you are writing a novel. Listed below are descriptions of fictitious stories. They do not refer to actual events. Please write below each one when you think the event is most likely to have occurred, specifying the time of day (noting a.m. or p.m.), day of week (Mon.-Sun.), date in the month (1st-31st), and the year." For those events in which time zones were of relevance, the subjects were asked to give the English time.

There followed a list comprising generic descriptions of events similar to those used in Experiment 1, as follows: "A Latin American President is assassinated; A water speed record is broken in England; An astronaut walks on the moon; A member of the royal family is assassinated; A major underwater discovery is made; A film star is shot in New York; A member of the royal family is married; A royal baby is born; The Pope is shot; A Middle East politician is assassinated."

Scoring. Time estimates were converted to deviation scores in the same way that error scores were computed in Experiment 1. Here, the reference times were the actual times of the corresponding events in Experiment 1. Year estimates before 1963 and after 1983 were coded as missing, since some subjects gave estimates as extreme as the nineteenth and twenty-first century.

Results and Discussion

Table 5 presents the deviation scores for each of the hypothetical events. The results show that, for many types of events, subjects select hours and months closer to the real time than would occur by chance. Six of the hour choices were significantly closer than by chance, and three of the month choices met this criterion. Day of the month, day of the week, and apparently year showed no concordance with actual time. The deviation scores for month, day of the month, day of the week, and hour were subjected to a repeated-measures ANOVA with event and scale as crossed, within-subject factors. Event [$F(9,225) = 3.68, p < .001$], scale [$F(3,75) = 14.46, p < .0001$], and their interaction [$F(27,675) = 5.50, p < .0001$] were all highly significant factors. Nonmonotonicity of scale was shown by a planned comparison for quadratic trend [$t(75) = 3.36, p < .01$].

Another analysis showed that not only *could* general time knowledge have contributed to the accuracy of month and hour estimates in Experiment 1, but that it apparently *did*. Mean month error scores for the different events in Experiment 1 were correlated with the corresponding mean deviation scores in Experiment 2 [Pearson $r(8) = .78, p < .01$]. A similar test for hour means also showed a substantial correlation [$r(8) = .92, p < .01$]. This concordance is difficult to attribute to factors other than the influence of knowledge about the usual times of events. Day of the month [$r(8) = -.57$] and day of the week [$r(8) = .24$] correlations were not significant.

The question of main interest is whether the scale effects in accuracy in Experiment 1 can be attributed solely to the sort of constraining process shown here. To test this possibility, separate between-group *t* tests were com-

Table 5
Mean Deviation of Time Estimates in Experiment 2 for Each Scale for Each Hypothetical Event

Event	Absolute Deviation Year	Proportion of Chance Deviation			
		Month	Day of Month	Day of Week	Hour
1 (President assassinated)	12.76	.90	1.63	1.33	1.07
2 (Water speed record)	10.89	1.50	.98	.83	.68***
3 (Astronaut on moon)	6.17	.86	1.46	.97	1.16
4 (Royal assassination)	2.00	.77*	1.89	1.24	.50***
5 (Film star shot)	4.40	.96	1.12	1.12	1.24
6 (Pope is shot)	3.94	.67**	.87	1.26	.70**
7 (Royal wedding)	1.86	.64***	1.72	1.10	.25***
8 (Mideast politician assassinated)	3.88	.92	.95	.94	.56***
9 (Royal baby)	2.29	.86	1.34	.88	1.11
10 (Undersea discovery)	4.36	1.20	1.08	.81	.71**
Mean	5.26	.93	1.30	1.05	.80

Note—Significance levels are for one-tailed *t* tests with the null hypothesis that entries are equal to or greater than chance deviations. "Year" excludes estimates before 1963 and after 1983. **p* < .05. ***p* < .01. ****p* < .001.

puted, comparing Experiment 1 and Experiment 2 mean estimates for month, day of the month, day of the week, and hour for each event. In Table 2, those means are italicized that previously had been found to be significantly less than chance and that, in this analysis, were found to be significantly less than the corresponding deviation scores in Experiment 2. Of the seven month means that were significantly below chance, five were also significantly less than the Experiment 2 means (*ps* < .03, one-tailed). Of the seven hour means that were significantly below chance, three were also significantly less than Experiment 2 means (*ps* < .03). Of the three day-of-the-month and day-of-the-week means that were significantly below chance, one (day of the week for Event 1) was also significantly less than the corresponding mean in Experiment 2 (*p* < .001).

This pattern still seems to show the nonmonotonicity of scale found in the deviation and confidence scores in Experiment 1. It also shows a number of specific violations of a scale decrement pattern. First, for Event 1, day of the week and hour are more accurate than would be expected given month estimates at chance level and a year deviation of 1.75. Second, Event 2 showed an accurate month estimate in spite of a year deviation of 2.76. Event 4 also had an unexpectedly accurate month estimate, but this may have differed from the Experiment 2 estimate in part because the boat was mentioned in the event description in Experiment 1 but not in Experiment 2. Third, for Events 6 and 8, hour is more accurate than any of the longer scales. Finally, for Event 10, month is more accurate than would be predicted from a mean year deviation of .29.

The results of Experiment 2 suggest that knowledge about when events usually occur increases the accuracy of performance on a date memory task. They also indicate that this knowledge is more likely to help on some time scales than on others. However, even if these two phenomena contributed to the accuracy pattern in Experiment 1, they cannot fully account for the unexpected ac-

curacy of many of the individual judgments. Memory also contributes to the violations of a scale decrement pattern.

GENERAL DISCUSSION

The pattern of scale differences in accuracy and confidence and the frequent reports of recall cues make it seem unlikely that we rely on a direct age-of-memory code in dating past events. Models based on a time-ordered memory format or the regular decay of traces cannot easily account for the findings that month and hour are sometimes identified more accurately than longer time scales. This effect is analogous to Hintzman et al.'s (1973) finding that subjects who forget in which of a series of lists a word was presented nonetheless tend to recall its within-list position. Their results present similar problems for age-of-memory models. Age of the trace probably provides useful time information for shorter delays than those tested here. A number of the verbal learning studies show "recency effects" for accuracy of time judgments when recall is within a few minutes of the end of the list (e.g., Hintzman et al., 1973; Tzeng, 1976), but these effects are typically absent for longer delays (e.g., Hintzman et al., 1973; Tzeng et al., 1979). The apparent paradox that more recent events are more accurately dated in the long-term memory studies, including this one, could be attributed to forgetting contextual associates of the events from which time of occurrence can be inferred.

The scale patterns of this study do not rule out the possibility that memory traces are assigned an absolute date tag at the time of encoding, since one could argue that information from some scales is simply not encoded in the first place. But the fact that events vary in scale effects, as revealed by the event $\times$ scale interactions, renders this model incomplete at best. An absolute date-tagging model also seems inconsistent with the facilitating effect of general time knowledge revealed in the Experiment 1-Experiment 2 comparisons and with subjects' reports of inferential processes in this and previous studies. In addi-

tion, Linton (1975) and Thompson (1982) have shown that exact dates are lost fairly quickly, whereas approximate dates are available for a longer period of time. It is not clear why, under a date-tagging model, exact dates should decay to approximate dates. Finally, on purely logical grounds, it seems that widespread date-tagging would require the frequent use of an elaborate tag-assigning mechanism and would present enormous storage demands. Of course, the dates of some important events (e.g., weddings, births) seem to be stable associates, but it is the exceptional quality of these few cases that seems to make them frequent cues for the dating of other events.

Another account of the representation of time information stresses the automatic coding of the relative temporal order of events (Tzeng & Cotton, 1980; Tzeng et al., 1979). This model was designed to explain recency judgments for pairs of stimuli, and it makes a number of correct predictions about performance on these tasks that are difficult for alternative models to explain. Theoretically, a model based on the temporal order of a large number of event pairs could be extended to explain memory for the time of events, as long as some of the events have time associates. Furthermore, introspective reports in the present and previous dating studies indicate that subjects often use preceding or succeeding events to constrain the time of the target event. However, order codes probably do not offer a complete description of the information used in dating. It is difficult for an order model to explain the violations of the scale decrement pattern, the constraining effects of general time knowledge shown in the Experiment 1-Experiment 2 correlations and in Experiment 2 accuracy, or the many recall cue reports that refer to experiences contiguous with the event itself. All of these findings are more easily explained by a reconstructive process.

The essential feature of a reconstructive model is that two sorts of information are combined in dating events, information directly associated with the event and general information. The latter includes knowledge about natural temporal patterns, activity routines, and diverse other autobiographical and social information. The process is assumed to begin with a search of the associates of the event. Occasionally, time labels may be directly associated with an event, such as when the time of an event was rehearsed or previously recalled. Presumably, however, time labels are not usually directly available, and general contextual associates must be evaluated to determine which are temporally informative. Location, activities, percepts, and contiguity to other events are all possible examples that were reported by our subjects. But these cues have temporal value only when supplemented by additional information about when one lived in a particular location, at what times of day one usually engages in particular activities, and so forth.

Other than the introspective data, the results of this study do not provide direct support for the application of inference to remembered contextual associates. There are, however, two lines of evidence for the application of in-

ference to event descriptions and some anecdotal evidence for its application to memory. Time estimates for fictitious events in Experiment 2 showed significant concordance both with real times and with the estimates in Experiment 1. The latter indicates that when subjects are asked to recall the time of some event, they will attempt to constrain the time using general knowledge. Another indication of the importance of inference comes from apparent distortions in dating. The greater-than-chance error on the month estimate for Event 2 in both Experiments 1 and 2 indicates that few subjects judged it likely that a water speed record would be attempted in midwinter. Similarly, the greater-than-chance hour estimates for Event 5 in Experiments 1 and 2 may be influenced by subjects' assumption that a shooting would probably take place during the daytime. A third example of distortion suggests an inference applied to remembered contextual associates. In pilot studies, a number of our (British) subjects incorrectly placed the Kennedy assassination in a summer month, noting that light clothes and an open car were signs of summer. Taken together, these findings illustrate the importance of inferential processes in dating events and show that proper controls are necessary if one wishes to distinguish memory in the narrow sense from inference.

Scale variation in the accuracy of dating seems more consistent with a reconstructive model than with the other models considered, although it is not decisive support for the former. Under the reconstructive model, scale variation in dating would be attributed to the likelihood that time information relevant to a particular scale is associated with the event, is recognized as having temporal value, and is correctly interpreted. If different types of information are associated with different events, one would expect scale effects to vary from event to event. This was shown by the significant interactions between scale and event. But the main effects for scale also indicate that some scales (day of the month and day of the week) are consistently inaccurate, whereas others (month and hour) are unexpectedly accurate. This is probably due to the extent to which associates (or information in the event description) can be used to logically constrain the time on different scales.

Three factors seem to be likely candidates for explaining the degree of constraint. First, the elements of a given scale can be rich or poor in associates. We examined this variation in a small follow-up study in which 30 subjects were asked to produce associations for elements of different scales. Each subject was assigned a randomly selected year from 1963 to 1982 (the range of years sampled in the main study), a month, a day of the month (e.g., "the 3rd of the month"), a day of the week, and an hour. The scales were presented in random order. Subjects were asked to list "In a word or phrase the things that come to mind for each of the following," and to give as many associations as possible during a 60-sec interval for each item. We simply counted the number of independent descriptions, events, or activities, varying from word to

phrase length, for each stimulus. A repeated-measures ANOVA with scale as a within-subject factor showed scale to be a highly significant effect [$F(4,116) = 25.25$, $p < .0001$]. A least significant difference test ($p < .01$) showed that month produced the most associates (5.23), followed by day of the week and time of day (3.97 and 3.87, respectively), which were followed by year and day of the month (2.50 and 2.37, respectively). Clearly, the elements of some scales are richer in associates than the elements of others. But the lack of complete concordance between the scale effect here and that in the main study requires additional factors to explain scale effects in dating.

One factor not reflected in the association scores is the degree of differentiation of elements on a scale. For example, the days of the week could be rich in associations, but several days might share many of the same associations. The latter would be consistent with studies of temporal orientation (Koriat & Fischhoff, 1974; Koriat, Fischhoff, & Razel, 1976; Shanon, 1979) that suggest that weekdays are relatively indistinct. It is notable that all of the events in the present study occurred on weekdays. Finally, event associates will be useful in constraining the time of an event only if they overlap with the distinctive features of a particular element of a scale. For example, an activity that is unique to Tuesday morning might be of little help in remembering the day of the week of an event that occurred on a Tuesday evening. On the other hand, cues like location of residence or age of one's children, which are useful cues to year, pervade long periods of time and are therefore likely to overlap with many events.

The three factors mentioned here may offer a partial account of mean differences in the accuracy, confidence, and cue number for the different scales. But there may also be principled ways in which different groups of people are able to exploit different scales. For example, day of the month may be especially salient to someone working in a payroll department or day of the week to a traveling salesperson with a weekly circuit. Memory for the time of events on scales longer than a day may prove especially difficult for long-term patients whose lives have little temporal structure on these scales.

REFERENCES

- BADDELEY, A. D., LEWIS, V., & NIMMO-SMITH, I. (1978). When did you last . . . ? In M. M. Gruneberg & R. N. Sykes (Eds.), *Practical aspects of memory*. London: Academic Press.
- FERGUSON, R. P., & MARTIN, P. (1983). Long-term temporal estimation in humans. *Perception & Psychophysics*, **33**, 585-592.
- GUENTHER, R. K., & LINTON, M. (1975). Mechanisms of temporal coding. *Journal of Experimental Psychology: Human Learning and Memory*, **104**, 182-187.
- HINTZMAN, D. L., BLOCK, R. A., & SUMMERS, J. J. (1973). Contextual associations and memory for serial position. *Journal of Experimental Psychology*, **97**, 220-229.
- KORIAT, A., & FISCHHOFF, B. (1974). What day is today? An inquiry into the process of time orientation. *Memory & Cognition*, **2**, 201-205.
- KORIAT, A., FISCHHOFF, B., & RAZEL, O. (1976). An inquiry into the process of temporal orientation. *Acta Psychologica*, **40**, 57-73.
- LIEURY, A., AIELLO, B., LEPREUX, D., & MELLET, M. (1980). Le rôle de repères dans la récupération et la datation des souvenirs. *Année Psychologique*, **80**, 149-167.
- LIEURY, A., CAPLAIN, P., JACQUET, A., & JOLIVET, C. (1979). La contraction du temps dans la datation des souvenirs anciens. *Année Psychologique*, **79**, 7-22.
- LINDSAY, P. H., & NORMAN, D. A. (1977). *Human information processing: An introduction to psychology* (2nd ed.). New York: Academic Press.
- LINTON, M. (1975). Memory for real world events. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition*. San Francisco: Freeman.
- LOFTUS, E. F., & MARBURGER, W. (1983). Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. *Memory & Cognition*, **11**, 114-120.
- SHANON, B. (1979). Yesterday, today and tomorrow. *Acta Psychologica*, **43**, 469-476.
- THOMPSON, C. P. (1982). Memory for unique personal events: The roommate study. *Memory & Cognition*, **10**, 324-332.
- TZENG, O. J. L. (1976). A precedence effect in the processing of verbal information. *American Journal of Psychology*, **89**, 577-599.
- TZENG, O. J. L., & COTTON, B. (1980). A study-phase retrieval model of temporal coding. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 705-716.
- TZENG, O. J. L., LEE, A. T., & WETZEL, C. D. (1979). Temporal coding in verbal information processing. *Journal of Experimental Psychology: Human Learning and Memory*, **5**, 52-64.
- UNDERWOOD, B. J. (1977). *Temporal codes for memories: Issues and problems*. Hillsdale, NJ: Erlbaum.

(Manuscript received October 3, 1984;
revision accepted for publication March 15, 1985.)