Ordering our world: An examination of time in autobiographical memory

John J. Skowronski
Northern Illinois University, De Kalb, IL, USA

W. Richard Walker
Winston-Salem State University, NC, USA

Andrew L. Betz
Progressive Insurance, Mayfield Village, OH, USA

In two studies people judged the order in which two real-world events occurred. Ordering performance was better for events that were recent or widely separated in time. Ordering performance was also consistently related to predicted event memorability and to the amount of processing given to an event during encoding. Ordering performance was not consistently related to the person-typicality, pleasantness, or emotional intensity of the events, and was also not related to whether the two events judged came from the same thematic category. These results suggest that memory for event order is not entirely reconstructed from event content. We suggest that the self-concept may sometimes serve as an implicit timekeeper in autobiographical memory.

Two approaches to the study of memory for time have dominated the autobiographical memory literature. In the first, participants estimate an event’s age (Huttenlocher, Hedges, & Prohaska, 1992) or report how old an event “feels” (Ross & Wilson, 2001). In the second, participants reconstruct an event’s exact date using a blank calendar as a dating aid (Skowronski, Betz, Thompson, & Shannon, 1991; Thompson, Skowronski, Larsen, & Betz, 1996).

Alternative approaches can be used to study time in autobiographical memory. One approach involves examining memory for stimulus sequences (Healy, 1974; Sinclair, Healy, & Bourne, 1997), a method only recently applied to autobiographical memory (Burt, Kemp, Grady, & Conway, 2000; Burt, Watt, Mitchell, & Conway, 1998). A second approach explores memory for time as it is derived from the number of items judged to intervene between two other items (Underwood & Malmi, 1978). This has been infrequently used to study autobiographical memory, although a self-report variant was employed to study memory retrieval strategies (Skowronski, Betz, Thompson, & Larsen, 1995). A third possibility is the judgment-of-recency (JOR) task (Greene, Thapar, & Westerman, 1998; Guttentag & Carroll, 1997; Hacker, 1980; Tzeng & Cotton, 1980; Yntema & Trask, 1963), in which participants judge the order in which stimuli were encountered. Although evidence suggests that people can date personal events by relating them...
to temporal landmarks (Loftus & Marburger, 1983; Skowronski et al., 1995; Shum, 1998), little research has explored JOR performance using real-world or autobiographical events (for exceptions, see Linton, 1975; Underwood, 1977).

The present article describes research exploring JOR performance for real-world memories. This research is prompted by three considerations. First, the JOR paradigm can explore how thematic relations between events might aid event dating. For example, if events are both part of an event script (Schank & Abelson, 1977), they be more easily ordered than events that are not both part of the script. These thematic questions are difficult to explore in the age estimation or calendar-aided dating paradigms (but see Thompson, Skowronski, & Betz, 1993).

Second, multiple methodologies help real-world memory researchers understand general properties of cognition. For example, results from age estimation and event dating tasks often produce converging findings. One such finding shows that recent events yield the most accurate time estimates (Betz & Skowronski, 1997; Friedman & Huttenlocher, 1997; Huttenlocher et al., 1992; Skowronski et al., 1991). Confidence would further increase in the age-accuracy relation if a similar finding emerged from research using the JOR paradigm. However, convergence is not always the rule. For example, Huttenlocher, Hedges, and Bradburn (1990) found that people overuse round numbers (e.g., 1 week) in open-ended age estimates. In comparison, research that asks participants to provide exact dates for events using a calendar has found that people will often know the day of the week on which an event occurred, but will be incorrect about the exact date (Betz & Skowronski, 1997; Larsen, Thompson, & Hansen, 1996; Thompson et al., 1993).

Open-ended age estimates have also provided evidence for a 5-day workweek and 2-day weekend within-week structure (Huttenlocher et al., 1992), while calendar-aided dating studies found evidence for a more complex within-week structure (Betz & Skowronski, 1997). Finally, research using the calendar paradigm has found relations between event characteristics (e.g., pleasantness) and event dating accuracy (Thompson et al., 1996); other methods that demand less precision do not (see Gibbons & Thompson, 2001; Prohaska, N.R. Brown, & Belli, 1998).

A third consideration prompting the use of the JOR paradigm with real-world events is assessment of the generality of laboratory results produced by this method. For example, judgements of event order become easier (more accurate, faster) as the events become more widely separated in time. Surprisingly, there is a paucity of evidence showing that this temporal distance effect can occur for long-term memories and real-world stimuli (but see Fuhrman & Wyer, 1988; Hurst & Volpe, 1982; Linton, 1975; Squire, Chace, & Slater, 1975, 1976; Underwood, 1977). Such generality has implications for many laboratory-derived theories of temporal representation, a fact already noted by Nairne (1991).

OBJECTIVES OF THE RESEARCH DESCRIBED IN THE PRESENT ARTICLE

The research described in the present article first examines whether JOR performance is related to the thematic similarity of the judged events. Two themes were examined: location (home, work, school, and other) and person (self and a specific other person). The expected relation between thematic similarity and JOR performance is unclear. On the one hand, events that have the same theme might possess a logical temporal relation (“first date” comes before “second date”) and may be stored in the same mental construction (N.R. Brown & Schopflocher, 1998a, b). Furthermore, thematically related events may be discussed in relation to each other more often than unrelated events, which may highlight temporal relations between events (“I’m sure I e-mailed the review before I left my office”). Finally, an event may remind one of other, similar events, which may facilitate JOR performance (Kemp, 1999; Winograd & Soloway, 1985). These ideas all suggest that JOR performance should be better when the two events judged are thematically related (also see Wyer, Shoben, Fuhrman, & Bodenhausen, 1985).

Alternatively, thematic similarity might interfere with JOR performance. This can occur if memory representations for events are nondiscriminable. For example, the memories “went to the store to buy a doll” and “went to the store to buy toys” might not be sufficiently distinct in memory, producing confusion about when they occurred. Such confusion may be especially large when the theme is a temporal category, as when events both occurred on the same day of the week. Such temporal cue over-
lap may make it especially difficult to order events.

A second objective of our research was to explore whether JOR performance is predicted by several event characteristics (event pleasantness, extremity, person-atypicality, and event age) and event-related processing (initial mental involvement with an event). The detailed memory traces promoted by such characteristics may provide cues that help place the event in time. Results from several calendar-aided dating studies support this conjecture (see Thompson et al., 1996, Chapter 7). If JOR performance depends on one’s ability to recall an event, then JORs made about better-recalled events (recent, pleasant, extreme, person-atypical, high mental involvement) should be made faster and with greater accuracy than JORs made about poorly recalled events. The current research also tested whether participants’ predictions of event memorability were related to JOR performance. Previous research has found that such judgements are good predictors of later event memory and event dating accuracy (Betz & Skowronski, 1997).

A third objective of our research was to examine whether laboratory-based JOR results generalise to JORs used in real-world events. Two results were of particular interest. The first was the temporal distance effect: JOR performance should be better for events that are widely spaced in time than for events that are closely spaced. The second result of interest was a direction matching effect: JOR performance should be better when the order of the events presented matches the direction of the judgement requested. That is, performance should be better if the JOR can be correctly answered with a “Yes” than if correctly answered with a “No”.

**METHOD**

JOR performance with real-world events was explored in two separate studies. Because there were only minor method differences between the two studies, these methods are most efficiently presented in a common Method section and the results in a common Results section.

**Participants**

In Study 1, 29 females (M age = 19.9) and 20 males (M age = 18.9) completed both phases of the study. In Study 2, 20 females (M age = 19.5) and 11 males (M age = 18.7) completed both phases. The ethnic heritage of all participants was Caucasian, except for one Hispanic female in Study 2. Participants received partial credit toward fulfilment of a research requirement in a psychology course as compensation for completion of the study.

**The diary collection phase**

The methods used in Phase 1 of these studies resemble those from several earlier studies (Skowronski et al., 1991; Thompson, Skowronski, & Lee, 1988). In Study 1 each person kept a diary of their own events and behaviours. The recording period lasted for 9 weeks. On each day of the recording period participants simply wrote a description of an autobiographical event in the designated space on each diary page. These designated spaces were pre-dated.

Participants were told that the recorded events were subject to several constraints. First, events should be unique, in that they should happen only once in an academic term and should have characteristics distinguishing them from similar events. Second, participants were asked to record events that differed in projected memorability, pleasantness, extremity and person-typicality. Third, events should be described in fewer than three lines of text. Finally, as much unique information about each event as possible (within the constraints) should be recorded. Participants were told that their memory for the events would be assessed and that performance in the diary test session would be optimal if their event descriptions were maximally informative.

Participants provided three ratings for each event. One rating assessed the typicality of the event. For example, participants were told that getting a C in a psychology exam could be very typical if they usually got Cs, but could be very atypical if they usually got As. This rating was made on a 7-point scale ranging from −3 for Very Atypical to +3 for Very Typical with Neither Typical nor Atypical as the label at the 0 point. A second rating assessed participants’ perceptions of event pleasantness. This rating was made on a 7-point scale, where −3 was Very Unpleasant and +3 was Very Pleasant, with Neither Pleasant nor Unpleasant at the 0 point. In a third rating participants predicted their future memory for the event. This rating was made on a 1–5 scale.
anchored by Not at All on the low end and Very Well on the high end.

In Study 2 participants received similar instructions but kept two diaries, one to record events in their own life and one to record events in the life of another person. Participants chose this person, whom they saw almost daily. Most were roommates, friends, or family members. The recording period lasted for 8 weeks. Participants in Study 2 made person-typicality and pleasantness ratings, but not the projected memorability rating. Instead, participants rated the extent to which they were mentally engaged in the event while it happened. This rating was made on a 5-point scale, where 1 was Not at all Involved and 5 was Extremely Involved.

The JOR phase

One week after completing their diaries, participants in Study 1 completed a task in which, on each of 200 trials, they judged the order of two events. The events presented on each trial came from the participants' own diaries. The entries were edited to maintain relatively constant length across entries. Item pairs were randomly selected, with replacement, on each trial.

On each trial, one of the events appeared at the top of a computer screen with a cue word (either Sooner or Later) immediately beneath it. The cue word varied randomly across trials and conveyed the direction of the judgement participants were to make. If the word Later appeared, participants were to judge whether the second event occurred after the first. If the word Sooner appeared, participants were to judge whether the second event occurred before the first. After reading the event and the cue word, participants pressed the space bar. This started the response timing routine and caused the second event to appear at the bottom of the computer screen.

If participants were given the Later prompt and thought that the second event occurred after the first, or were given the Sooner prompt and thought that the second event occurred before the first, they were to press the Y key on the keyboard. If participants were given the Later prompt and thought that the second event occurred before the first, or were given the Sooner prompt and thought that the second event occurred after the first, they were to press the N key. When a key was pressed, the response and the latency were automatically recorded and the next trial immediately began.

This continued until all 200 trials were completed.

In a final task, events were presented one at a time in a random order unique to each participant. After reading each event, participants reported the location at which the event occurred: home (by pressing the H key), school (by pressing S), work (by pressing W), or other (by pressing O). A new event appeared immediately after each response until all events had been classified. After completion of this task, participants were debriefed and dismissed.

The JOR phase of Study 2 proceeded in the same manner, with two exceptions. First, in Study 2 a given JOR trial could present two events from the self diary, two events from the other diary, or one event from each diary. The second exception was that the participants in Study 2 were not asked to complete the location classification task.

RESULTS

An overview of the measures and the regression analyses

The dependent measures in Study 1 were the accuracy of each of the 200 order judgements and the time it took to make each judgement. Latency data are often transformed or trimmed prior to analysis (see Fazio, 1990; Ratcliff, 1993). Our analyses examined three separate variants of the latency data (deletion of latencies greater than 3 s.d. from the overall latency mean, natural logarithmic transformation, inverse transformation). Results from the three sets of analyses were highly comparable, hence, this article reports only results from analyses of the trimmed data sets.

These measures were analysed using within-participant regression analyses. Dummy-coded variables corresponding to each participant were entered into all regression models to control for within-subject effects. The pleasantness, person-typicality, and initial mental involvement of each of the two items on each trial were additional variables in the regressions. These were treated as dummy-coded categorical variables so that we could examine the mean accuracy and the mean latency latency in the presence of these effects.
latency of responses at each level of each variable. The direction of the judgement (sooner or later), the congruency between the item order presented and the judgement requested (a correct “yes” = congruent; a correct “no” = incongruent), and whether there was a match between the days of the week on which the two events occurred (match vs no match) were additional dummy-coded categorical variables in the analyses. The analyses included two linear terms reflecting the age of the younger event in the pair and the difference between the ages of the two events in the pair. In addition, the relation between the thematic qualities of events and JOR performance was assessed by using event location match as a theme. Judgements in which both events came from the same location (e.g., home–home) were placed in one category; judgements in which events came from different locations (e.g., home–school) were placed in a second category. For purposes of the analyses the category for each judgement was dummy coded.

Analyses of the Study 2 data included these same variables, with two exceptions. First, a dummy-coded categorical variable reflecting the initial mental involvement associated with each event replaced the predicted memorability variable. Second, a person match variable replaced the location match variable. Judgements were placed into one category if both events on a trial were self-events, into a second category if both were other-events, and into a third category if one event came from each person. Categories were identified in the analyses via dummy coding.

Two additional terms in all analyses reflected linear and quadratic effects of trials. These were included only to control for practice effects, and they predicted JOR performance in all analyses—Study 1: linear latency \( \beta = -0.315, F(1,9285) = 1293.16, p < .0001; \) quadratic latency \( \beta = 0.148, F(1,9285) = 228.03, p < .0001, \) linear accuracy \( \beta = 0.037, F(1,9476) = 14.76, p < .0001; \) quadratic accuracy \( \beta = -0.025, F(1,9476) = 5.35, p < .03, \) \( M \) \( SE = .87; \) Study 2: linear latency \( \beta = -0.327, F(1,5811) = 961.66, p < .0001; \) quadratic latency \( \beta = -0.136, F(1,5811) = 131.52, p < .0001, \) \( M \) \( SE = .65; \) linear accuracy \( \beta = 0.028, F(1,5878) = 5.10, p < 0.03, \) quadratic accuracy \( \beta = -0.043, F(1,5878) = 9.27, p < .01, \) \( M \) \( SE = .94.

Finally, in regression analyses a variable will often predict a criterion when that variable is the sole predictor, but will not predict the criterion when placed in a simultaneous regression analysis with other predictors. This occurred infrequently in our analyses, so we report results only for simultaneous regressions that included all predictors. Hence, significant (or nearly significant) effects that we report take into account relations among all predictors in each study.

**Temporal event characteristics and JOR performance**

**Temporal distance.** Performance on the JOR task should be good at long temporal distances between events and poor at short temporal distances. The data from both studies (see Figure 1) show exactly this pattern. Longer temporal distances were associated with higher accuracies [Study 1: \( \beta = 0.123, F(1,9476) = 117.10, p < .0001, \) \( M \) \( SE = .87; \) Study 2: \( \beta = -0.059, F(1,5878) = 89.82, p < .0001, \) \( M \) \( SE = .94) and shorter latencies [Study 1: \( \beta = 0.900, F(1,9285) = 14.35, p < .001, \) \( M \) \( SE = .71; \) Study 2: \( \beta = -0.059, F(1,5811) = 18.09, p < .0001, \) \( M \) \( SE = .65)].

**Event age.** JOR performance should depend on the age of the events involved in the judgement. Results from both studies confirm that JOR performance was enhanced when the pair included a recent event (see Figure 2). These JORs were associated with higher accuracies [Study 1: \( \beta = -0.333, F(1,9476) = 8.71, p < .01, \) \( M \) \( SE = .87; \) Study 2: \( \beta = -0.063, F(1,5878) = 18.09, p < .0001, \) \( M \) \( SE = .94) and shorter latencies [Study 1: \( \beta = 0.39, F(1,9285) = 14.35, p < .001, \) \( M \) \( SE = .71; \) Study 2: \( \beta = 0.028, F(1,5811) = 4.92, p < .03, \) \( M \) \( SE = .65)].

**Day-of-week.** A third question of interest in these studies was whether JOR performance is altered when events occurred on the same day of the week. Because research shows that calendar-aided dating errors often take the form of “right day, wrong week”, the day of the week may be a part of, or may be reconstructed from, the event memory trace. How this information might affect performance on the JOR task is unclear. Two events that occurred on the same day of the week...
Figure 1. Mean response accuracies (top) and response latencies (bottom) plotted as a function of temporal distance for Studies 1 and 2.
Figure 2. Mean response accuracies (top) and response latencies (bottom) plotted as a function of the age of the youngest event in the pair for Studies 1 and 2.
might make the JOR more or less difficult, depending on whether event similarity sparks stimulus confusion or gives the participant cues about the sequence of events.

The data are weakly consistent with the idea that judgements are more difficult when the events to be judged both occur on the same day of the week. In Study 1, judgments were significantly slowed when events occurred on the same day of the week [same day $M = 5.07$, different day $M = 4.90$, $F(1, 9285) = 5.43$, $p < .02$, $MSE = .71$]. However, increased difficulty was not reflected in the accuracy data [same day $M = .67$, different day $M = .67$, $F(1, 9476) = .05$, $p > .82$, $MSE = .87$]. In Study 2 judgments were both slower and less accurate when both events occurred on the same day of the week, but neither measure was significant [latency $F(1, 5811) = 1.30$, $p > .26$, $MSE = .65$; accuracy $F(1, 5878) = 1.19$, $p > .28$, $MSE = .94$].

**Event psychological characteristics and JOR performance**

*Predicted event memorability.* Because event memory is related to accuracy in event dating, predicted event memorability should also be related to an individual’s ability to place events in time. When a trial included an event that was expected to be well remembered, judgement accuracy was better than when the trial included an event that was expected to be poorly-remembered [Ms by event 1 memory ratings: 1 = .65, 2 = .65; 3 = .68; 4 = .66; 5 = .71; Ms by event 2 memory ratings: 1 = .64; 2 = .66; 3 = .67; 4 = .67; 5 = .70; event 1 $F(4, 9476) = 2.95$, $p < .02$; event 2 $F(4, 9476) = 2.29$, $p < .06$, $MSE = .87$].

However, these effects did not characterise the latencies. The shortest latency occurred when event 2 in the pair received the lowest memory rating [Ms by event 1 memory ratings: 1 = 4.97 s, 2 = 4.98 s; 3 = 5.00 s; 4 = 4.96 s; 5 = 5.01 s; Ms by event 2 memory ratings: 1 = 4.79 s; 2 = 5.04 s; 3 = 5.07 s 4 = 4.97 s; 5 = 5.04 s; event 1 $F(1, 9285) = .09$, $p > .98$, event 2 $F(1, 9285) = 3.77$, $p < .04$, $MSE = .71$]. This anomalous result may reflect the fact that participants opt for guessing when they do not know when events occurred, responding quickly and randomly.

*Event person-typicality.* Person-atypical events should be distinctive in memory. In addition, useful temporal information might be included in the event memory trace of atypical events (see Shum, 1998). Both of these ideas lead to the expectation that performance should be enhanced on JOR trials that include person-atypical events. In accordance with this expectation, judgement accuracy in Study 1 was low when one of the events in the judged pair was associated with high person-typicality [event 1 $M$ accuracy ratings by person-typicality rating: $-3 = .70$, $-2 = .69$, $-1 = .68$, $0 = .67$, $1 = .67$, $2 = .66$, $3 = .65$; event 2 $M$ accuracy ratings by person-typicality rating: $-3 = .70$, $-2 = .66$, $-1 = .67$, $0 = .69$, $1 = .67$, $2 = .64$, $3 = .66$; event 1 $F(6, 9476) = 1.97$, $p < .07$; event 2 $F(6, 9476) = 3.14$, $p < .01$, $MSE = .87$]. However, in Study 2 the relation between person-typicality and judgement accuracy was not significant for either event 1 or event 2, although the pattern of means for event 1 did resemble the pattern of means obtained for the events in Study 1 [event 1 $F(6, 5878) = 1.27$, $p > .26$; event 2 $F(6, 5878) = .68$, $p > .66$, $MSE = .94$]. Taken together, these results weakly suggest that JOR accuracy decreases when one of the events in the pair is an event that is very typical for the target of the events.

In comparison, the response latency data from Study 1 shows that JORs were faster for both highly person-atypical and highly person-typical events [event 1 $M$ response latencies by person-typicality rating: $-3 = 4.88$, $-2 = 4.88$, $-1 = 5.03$, $0 = 5.03$, $1 = 5.21$, $2 = 4.99$, $3 = 4.86$; event 2 $M$ response latencies by person-typicality rating: $-3 = 4.83$, $-2 = 5.06$, $-1 = 5.00$, $0 = 4.95$, $1 = 5.08$, $2 = 5.09$, $3 = 4.88$; event 1 $F(6, 9285) = 2.93$, $p < .01$, event 2 $F(6, 9285) = 2.09$, $p < .06$, $MSE = .71$]. The data from Study 2 did not yield a significant relation between person-typicality and latency for either event, although the pattern of means for event 2 did resemble the patterns observed in Study 1 [event 1 $F(6, 5811) = 1.53$, $p > .16$, event 2 $F(6, 5811) = .76$, $p > .59$, $MSE = .65$]. In the context of the accuracy data, these latencies weakly suggest that JORs that involve highly person-typical events are especially difficult: judgements that involve such highly person-typical events are made especially inaccurately and rapidly.

*Event valence and intensity.* We looked for evidence that JOR performance was related to event valence and intensity. We expected intensity to be positively related to JOR accuracy, but in Study 1 event pleasantness ratings did not predict
JOR accuracy [event 1 $F(6, 9476) = .49, p > .81$; event 2 $F(6, 9476) = 1.23, p > .28, MSE = .87$]. The event pleasantness ratings predicted accuracy for one of the two events in Study 2 [event 1 $F(6, 5878) = 2.55, p < .02$; event 2 $F(6, 5878) = 1.25, p > .27, MSE = .94$]. However, examination of the means indicated that the significant effect was largely driven by an abnormally high accuracy score at the zero point (neither pleasant nor unpleasant).

The pleasantness measure was significantly related to judgement latency for one of the Study 1 events [event 1 $F(6, 9285) = .40, p > .88$, event 2 $F(6, 9285) = 2.28, p < .04, MSE = .71$]. However, the means for this effect indicated that it was caused by rapid responding when events were neutral. Given that this latency decrease was not associated with a corresponding change in response accuracy for event 2 in Study 1, this effect is probably meaningless. This is especially likely given that this latency decrease was not observed for event 1 in Study 1, nor did it occur for either of the events in Study 2. In fact, analyses of the pleasantness ratings associated with each of the events in Study 2 indicated that they did not predict response latency [event 1 $F(6, 5811) = .64, p > .69$, event 2$F(6, 5811) = 1.04, p > .39, MSE = .65$].

**Event processing and JOR performance**

*Initial mental involvement.* In Study 2, we explored whether mental involvement with events during encoding predicted JOR performance. Response accuracy rates were significantly higher when one of the events in the JOR produced high initial mental involvement [$M$ response accuracies by initial mental involvement ratings, event 1: 1 = .64, 2 = .64, 3 = .69, 4 = .67, 5 = .67, event 2: 1 = .61, 2 = .66, 3 = .68, 4 = .67, 5 = .69; event 1 $F(4, 5878) = 2.99, p < .02$, event 2 $F(4, 5878) = 4.17, p < .01, MSE = .94$]. However, the response latency data were not significant for either event [event 1 $F(4, 5811) = .60, p > .66$, event 2 $F(6, 5811) = .42, p > .79, MSE = .65$].

The absence of pleasantness effects in the data in Study 2 cannot be attributed to differing relations between the pleasantness and diary target variables, as might be suggested by some results from the calendar-aided dating paradigm (see Thompson et al., 1996). Results were similar when the data were separately examined by the type of events in the judgement pair (self–self, other–other, self–other). In fact, in Study 2 such consistency across these three pair types generally characterised all results from Study 2.

**Thematic relation between events and JOR performance**

*Location.* In Study 1 we assessed whether JOR performance was related to the location consistency between events (same location versus different locations). Because such relations can both produce temporal confusion or can provide cues about event orders, it was unclear whether a thematic location relationship between events would aid or hurt JOR performance. Neither possibility was supported by the data. Location matches or mismatches were not significantly related to JOR accuracy [$F(1, 9476) = 1.01, p > .31, MSE = .87$; match $M = .66$, mismatch $M = .67$]. Although response latencies were shorter when the two events came from the same location (4.93 s) than when the locations differed (5.04 s), $F(1, 9285) = 3.31, p < .07, MSE = .71$, in the absence of a clear accuracy effect the meaning of this result is ambiguous. It is unclear if this decreased latency reflects easier JORs when the two events came from the same location, or if it reflects more difficult JORs (as when an individual recognises that they won’t be able to make an accurate judgement so they respond randomly and rapidly).

*Subject of the event.* Participants in Study 2 kept two diaries, one for self-events and one for the events of another person. Hence, Study 2 provided another opportunity to examine the relation between themes (as reflected in the match between the person described by the events on each trial: self–self, other–other, or self–other) and JOR performance. However, the thematic variable in Study 2 did not significantly predict JOR performance [judgement accuracy, $F(2, 5878) = .01, p > .98, MSE = .94$; judgement latency, $F(2, 5811) = .19, p > .83, MSE = .65$].

**Judgement task characteristics and JOR performance**

*Judgement direction.* The fact that people often tell stories in chronological order might be related to JOR performance. Because of the tendency to tell stories from beginning to end, it seemed plausible that JOR performance might be better when participants responded to the later cue word (which required participants to assess whether the second event came later than the first) than to the sooner cue word (which required
participants to assess whether the second event came before the first). This forward order hypothesis is consistent with memory production latency data reported by Anderson and Conway (1993). Indeed, JOR accuracy was higher in response to the later cue word (Study 1 $M = .68$; Study 2 $M = .67$) than to the sooner cue word (Study 1 $M = .66$, Study 2 $M = .66$) in both studies, although this difference was significant only in Study 1 [$F(1,9476) = 4.71, p < .03, MSE = .87$; Study 2 $F(1,5878) = 1.29, p > .26, MSE = .94$]. The latency measure yielded non-significant results in both studies [Study 1 $F(1,9285) = .01, p > .90$, $MSE = .71$; Study 2 $F(1,5811) = .36, p > .54, MSE = .65$].

Congruency between events presented and judgement requested. Laboratory research suggests that JOR performance is better when the order of events presented is congruent with the direction requested by the cue word (resulting in a “yes” response) than when it is not congruent (resulting in a “no” response). There was substantial evidence for this effect in both the accuracy and latency data. In both studies, responses were more accurate in the congruent condition (Study 1 $M = .69$; Study 2 $M = .69$) than in the incongruent condition [(Study 1 $M = .65$; Study 2 $M = .64$), Study 1 $F(1,9476) = 28.36, p < .0001$, $MSE = .87$; Study 2 $F(1,5878) = 14.17, p < .001$, $MSE = .94$]. The response latencies were also significantly faster in the congruent condition (4.88 s) of Study 1 than in the incongruent condition (5.08 s) [$F(1,9285) = 16.67, p < .0001$, $MSE = .71$]. This latency pattern also emerged in Study 2 (congruent $M = 4.54$, incongruent $M = 4.60$ s), but it did not reach statistical significance [$F(1,5811) = 1.21, p > .27$, $MSE = .65$].

**DISCUSSION**

The two studies reported in this article use a JOR paradigm, a methodology that has been underused in the autobiographical memory literature, to explore how people know when events occurred. These studies can provide convergent validity with other autobiographical memory studies, but can also provide insight into which findings might be bound to specific paradigms. In addition, these studies make connections between the autobiographical memory literature and the extensive literature that developed to explore questions about short-term memory for serial lists.

Our results can be interpreted through the lenses provided by these goals. For example, consider the temporal distance effect: The greater the temporal distance between two items, the better the JOR performance. Our studies extend the generality of this result by showing that this effect occurs with both autobiographical events and with real-world events that describe others. Similar convergence comes from our finding that JOR performance decreases as the age of the youngest event in the pair increases. This finding is conceptually similar to the finding from calendar-aided event dating studies showing that dating accuracy decreases with event age. Similarly, our JOR data show that two of the variables that are related to dating accuracy (initial mental involvement, projected event memorability) are also related to JOR performance.

Perhaps the most important convergence was not apparent until we reconsidered results from the calendar-aided dating accuracy studies. This reconsideration was prompted by the temporal distance effect. In the present studies this effect was independent of all other effects included in the regression analyses. These included relatively potent predictors of the quality of event memory, such as event age, the initial memory rating, and the initial mental involvement rating. To the extent that these variables account for the content of event memory, these results can be interpreted as reflecting a degree of independence between memory content and knowledge of an event’s relative location in time (see McElree & Doser, 1993). We realised that similar findings had previously emerged in our calendar-aided dating studies. One objective in those studies (as summarised in Thompson et al., 1996) was to examine the relation between predictors related to memory and dating accuracy. Event age was included as a control variable in all analyses, and was always a significant predictor of dating accuracy, even when entered into simultaneous regression analyses with other variables. In retrospect, this result can also be interpreted as providing evidence that event age predicts dating accuracy, even when event memory is statistically accounted for. Obviously, the potency of claims about the partial independence of “knowing what” and “knowing when” in both sets of studies depends on the extent to which event memory is adequately controlled for in the regression analyses.

However, the possible independence between “knowing what” and “knowing when” is supported by the results of other research. A frequent
finding in the autobiographical memory literature is that richer cues lead to better recall (Wagenaar, 1986). In pursuing a similar idea, Brewer (1988) used a cueing paradigm to explore which features of a real-world event cued recall of other event features. Brewer found that recall of an event’s date was poorly cued by other event features (such as what happened). Further, the event date itself served as a poor cue for those event features. Hence, in Brewer’s study “memory for what” seemed to be largely independent of “memory for when”. A similar conclusion comes from list learning studies: The ability to recall the position of an item in a serial list is often largely independent of an individual’s ability to recall the item. This phenomenon occurs so regularly that knowledge of the position of an item in a list was thought by some to be automatically encoded (Hasher & Zacks, 1979). Although more recent research suggests that position memory does not meet the classic criteria for automaticity (Naveh-Benjamin, 1990), it is clear that knowledge about the position of an item in a list is often learned incidentally and is not strongly related to memory for the item.

The possible independence of “knowing what” and “knowing when” is a relevant issue given the current emphasis on the reconstruction of temporal information from event memory (Larsen & Conway, 1997; Friedman, 1993). Indeed, many of our own explorations of the relation between psychological characteristics of events and dating accuracy were derived from a belief that temporal estimation can be influenced by a memory’s content. We continue to subscribe to that belief. However, there appears to be more to memory for time than explicit reconstruction of time from event memory: People might sometimes have access to knowledge of an event’s location in time in a way that transcends the factual information contained in the memory.

One possibility is that an event might fit a thematic structure that specifies the order in which events occur in relation to each other (such as a script: Schank & Abelson, 1977). Alternatively, events sharing a theme may be discussed with other people, causing events to become linked in a temporal order independent of memory content (Guenther & Linton, 1975; Schank & Abelson, 1995). A third possibility comes from perturbation theory (e.g., Lee & Estes, 1977), which suggests that events are linked to higher-order “control elements” that can serve as guides to the location of an event in time (see Nairne, 1991). Such control elements might be the approximate time that an item appeared in a serial list (early, middle, or late) or a period of one’s own life (high-school, college). Conway’s recent theorizing (Anderson & Conway, 1997; Conway & Pleydell-Pearce, 2000; Conway & Rubin, 1993) has postulated exactly the latter structure, providing a potential framework for understanding the temporal distance and age effects described in the present article (also see Huttenlocher, Hedges, & Prohaska, 1988).

However, we suggest an additional possibility: The self might serve as a source of implicit knowledge about time. The process that we propose is straightforward. Glenberg, Bradley, Kraus, and Renzaglia (1983) suggested that temporal judgements can be based on the extent to which contextual information encoded at information exposure overlaps with the current context. Assume that the “self” (broadly construed, in a connectionist sense, as a pattern of activation in memory reflecting our experiences and knowledge) can similarly serve as a learning context. When an individual retrieves an event, the self-context associated with that memory may also be activated. A comparison of the recalled self-context with the state of the current self can provide a sense of how old the event is, independently of event content. If the self-context of the memory is close to the current self-concept, the memory will be judged as recent. If the self-context of the memory is different from the current self-concept, the memory will be judged as old. Similarly, comparison of the relative discrepancies between each of two events and the current self can provide clues to the relative ages of the events.

Wilson and Ross (2001) have already obtained data supporting these conjectures. They found that negative events “feel older” than positive events, ostensibly because those events conflict with the positive current self-concept that people typically have of themselves. Other existing data also provide supportive evidence for the mechanism that we describe. For example, reminiscing may temporarily activate the self-concept from an earlier life period, causing it to temporarily influence the current self-concept. If this were to occur, events that occurred during those times would seem as if they had occurred relatively recently. Interestingly, in an early description of autobiographical memory, Hall (1899) describes exactly these effects.

One new test of the “self as implicit time-keeper” idea might come from a comparison of
the slope for the temporal distance effect obtained from a JOR task completed by teenagers (or even pre-teens) to that obtained from adults. The self-concept of teenagers should be rapidly changing in response to the physical, cognitive, social, and emotional changes in their lives, whereas adults should have a relatively stable self-concept (see Mortimer, Finch, & Kumka, 1982). One implication of these different rates of change in the self-concept is that events of equal age should seem older to teenagers than to adults. Certainly, this matches a frequent lament: As fifth-graders, the fourth-grade class that we last year seems as if it happened aeons ago; as college professors, the class that we taught last year seems as if it happened yesterday. A similar result might emerge when comparing JORs from adults with relative life stability to JORs from adults who experience major life changes. Such changes should produce substantial change in the self-state, so that events that occurred prior to the changes should seem older to a person than should events of equivalent age for a person whose life was relatively stable.

These ideas are consistent with studies suggesting that implicit characteristics of memory, such as accessibility, affect temporal judgements (see G.D.A. Brown, Preece, & Hulme, 2000; N.R. Brown, 1990; N.R. Brown, Rips, & Shevell, 1985). Such effects may be most likely on temporal judgement tasks (such as JOR) that do not demand precision, allowing people to make judgements by relying on “gut feelings”. Certainly, reliance on mindless, implicit processes might help to explain why variables such as person-typically, pleasantness, and emotional intensity were not good predictors of JOR performance. These variables may aid estimation accuracy in calendar-aided dating tasks by increasing the quality of event memory. However, memory quality may play only a minor role in the temporal judgement processes used in the JOR task. If this is true, then it makes sense that memory-enhancing variables are unrelated to JOR performance. These speculations also follow from Smith’s (1988) outstanding hypothesis, which explained the weak and inconsistent nature of context effects in memory (Fernandez & Glenberg, 1985).

This argument can also explain the lack of a relation between the two thematic variables (location in Study 1 and person in Study 2) and JOR performance. However, other explanations are certainly possible. For example, location may have failed predict JOR performance because it is not central to the meaning of an event. However, the same cannot be said of the self–other variable explored in Study 2. We admit to considerable surprise at the finding that performance on the JOR task in Study 2 was not related to the match between the people who performed the behaviours. However, this null effect fits with our speculation that the self can serve as an implicit timekeeper, regardless of where the event occurred or whom the event was about.

This implicit mechanism is, admittedly, speculative. Furthermore, even if the mechanism is valid, it is certainly not the only mechanism that affects our knowledge of when events occurred. Temporal knowledge is retrieved and reconstructed in a variety of ways from a variety of sources in a variety of circumstances. Nonetheless, we think that an implicit mechanism, rooted in the self, can help to understand some of the paradoxes of temporal knowledge. In future research we will pursue this idea, as well as other cognitive and social mechanisms that allow us to know when events occurred, both in our own lives and in the world around us.

Manuscript received 9 February 2001
Manuscript accepted 1 February 2002

REFERENCES


