Journal of Experimental Psychology:
Learning, Memory, and Cognition
2006, Vol. 32, No. 5, 1031–1044

0278-7393/06/$12.00 DOI: 10.1037/0278-7393.32.5.1031

Comprehending Narratives Containing Flashbacks: Evidence for Temporally Organized Representations

Berry Claus and Stephanie Kelter
Technical University of Berlin

This study investigated the representations that readers construct for narratives describing a sequence of events. Participants read narratives describing 4 successive events in chronological order (Event 1, Event 2, Event 3, Event 4 [E1, E2, E3, E4] Experiment 1) or in nonchronological order with E1 being mentioned in a flashback (E2, E3, E1, E4; Experiments 2–4). The information about the duration of E2 was manipulated, and the mental accessibility of E1 was tested at the end of a passage. All 4 experiments showed that E1 was less accessible if the text implied that it occurred a relatively long time ago in the described world compared with when it occurred a shorter time ago. This result suggests that readers construct a temporally organized representation even if the text structure does not suggest such an organization.

Keywords: text comprehension, narratives, temporal organization, flashbacks

Time plays a central role in how people experience the world. It is therefore not surprising that language provides a rich means of conveying temporal information. In recent years, the question of how temporal information is processed during sentence or text comprehension has gained increasing interest in psycholinguistic research. This research has addressed various linguistic devices for conveying temporal relations and properties. Some studies have investigated the processing of tense, grammatical aspect, or lexical aspect (Carreiras, Carriedo, Alonso, & Fernández, 1997; Gennari, 2004; Madden & Zwaan, 2003; Magliano & Schleich, 2000; Morrow, 1985, 1990; Proctor, Dickey, & Rips, 2004; Todorova, Straub, Badecker, & Frank, 2000). Other studies have been concerned with the processing of information conveyed by temporal connectives or temporal and durative adverbials (Anderson, Garrod, & Sanford, 1983; Bestgen & Vonk, 1995, 2000; de Vega, Robertson, Glenberg, Kaschak, & Rinck, 2004; Hoeks, Stowe, & Wunderlink, 2004; Kelter, Kaup, & Claus, 2004; Levine & Klin, 2001; Mandler, 1986; Münte, Schiltz, & Kutas, 1998; Rinck & Bower, 2000; Rinck & Weber, 2003; Speer & Zacks, 2005; Zwaan, 1996; Zwaan, Madden, & Whitten, 2000). Finally, some studies have investigated the interaction of linguistically conveyed temporal information and background knowledge (Anderson et al., 1983; Graesser, Kassler, Kreuz, & McLain-Allen, 1998; Hagmayer & Waldmann, 2002; Rapp & Taylor, 2004; Rinck, Gámez, Díaz, & de Vega, 2003; Rinck, Hähnel, & Becker, 2001).

One main result of this research is that at any moment during narrative processing, the states of affairs that obtain at the reference time of a currently processed sentence are especially easy to mentally access. States of affairs that obtained in the described world prior to that time are less accessible, no matter whether the relevant temporal information has to be derived from tense, aspect, temporal adverbials, the text structure, or general knowledge (Anderson et al., 1983; Bestgen & Vonk, 1995; Carreiras et al., 1997; Magliano & Schleich, 2000; Rinck & Bower, 2000; Rinck & Weber, 2003; Zwaan, 1996; Zwaan et al., 2000). This finding can be considered a piece of evidence that comprehenders impose a temporal structure on the described world that corresponds to how they experience the world. The reference time is interpreted as the "now" in the described world, and just as in direct experience, the situation existing at this time is mentally highlighted.

In real life, people experience the now point as moving forward continuously. Empirical findings suggest that when reading a narrative, people similarly expect the now point to move forward continuously in the described world. When a sentence implies a temporal discontinuity due to a shift of the reference time to a much later interval in the narrated world, then reading times are prolonged compared with when the reference time of the sentence moves on to the immediately following time interval (Bestgen & Vonk, 2000; Rinck & Weber, 2003; Speer & Zacks, 2005; Zwaan, 1996; Zwaan, Magliano, & Graesser, 1995; see also Hyönä, 1995). In the case of a temporal discontinuity, readers seem to terminate the hitherto constructed representation and initiate a new one (Bestgen & Vonk, 2000; Gernsbacher, 1997; Zwaan, 1996).

In contrast, if a narrative describes consecutive events without a temporal discontinuity, then readers apparently construct a coherent representation that captures the passage of time in the narrated world as a continuous variable (Kelter et al., 2004; Rinck & Bower, 2000). For example, Kelter and colleagues (2004) pre-

Berry Claus and Stephanie Kelter, Institute of Psychology, Technical University of Berlin, Berlin, Germany.

This research was supported by German Research Foundation Grants Ke 292/5-1 and 292/5-3 to Stephanie Kelter. We are grateful to Tom Wenzl, Barbara Kaup, and Alan Garvie for helpful comments on earlier versions of this article. We thank Philippe Büttner, Guido Liebe, Peter Pollmanns, Alexander Richter, and Alexander Specht for assistance in collecting the data. We also thank Philippe Büttner for programming the experimental software used in Experiment 3.

Correspondence concerning this article should be addressed to Berry Claus, who is now at Allgemeine Psychologie I, University of Potsdam, P.O. Box 60 15 53, 14415, Potsdam, Germany. E-mail: clausb@rz.uni-potsdam.de

sented readers with narratives describing several consecutive events in chronological order and, at the end of a narrative, tested the mental accessibility of the first event of the sequence. The temporal distance between the point in time at which this event occurred in the described world and the narrative now point at the time of testing was either relatively long or short, depending on what the text implied as to the duration of the second event in the sequence. Participants needed more time to access the target event when it occurred a long time ago in the described world compared with when it occurred only a short time ago. This *temporal-distance effect* suggests that the participants had constructed a temporally analogous representation of the narrated world.

What is still unclear, however, is how much importance comprehenders attach to a temporal organization of their representations. In all the aforementioned studies, the events were described in chronological order. This may have fostered the construction of a temporally organized representation. A more critical test would therefore be narratives containing flashbacks. A *flashback* mentions a state of affairs that obtained in the described world prior to the situation referred to in the sentence that preceded the flashback. For example, Sentence 2 is a flashback, conveying information about an event that occurred earlier than the situation described in Sentence 1:

1. Sorrowfully, John went into his kitchen and poured himself a large glass of whiskey.

2. His girlfriend had just left him.

If it is true that comprehenders attach great importance to a temporal organization of their representations, then it is to be expected that they integrate the past state of affairs mentioned in a flashback at the chronologically appropriate location into their representation. We refer to this hypothesis as the *chronological hypothesis*.

However, there are good reasons to question the validity of the chronological hypothesis in addition to the fact that rearranging the incoming text information may cost considerable effort. A flashback does not occur arbitrarily in a text but is usually closely related to the situation described immediately before. This means that the information of a flashback could be integrated into the hitherto constructed representation by just focusing on local coherence. More specifically, discourse–linguistic analyses have shown that a flashback provides background information pertaining to the causes of the situation that the preceding sentence referred to, its peculiarities, the protagonist's goals, or other aspects (e.g., Lascarides & Asher, 1993a, 1993b; Polanyi & van den Berg, 1996; ter Meulen, 2000). For example, in the text given above, the flashback provides an explanation of why John is full of sorrow. Thus, when considering this function of flashbacks, one would assume that comprehenders attach the flashback information as a supplement to the representation of the situation described immediately before. This hypothesis seems even more plausible when taking into account that a flashback sentence, though providing information about a state of affairs at an earlier time interval, does not shift the reference time to this earlier interval but retains the reference time of the preceding sentence (e.g., Lascarides & Asher, 1993a; ter Meulen, 2000; Webber, 1988; see also "perspective time" in Kamp, van Genabith, & Reyle, in press). We refer to this hypothesis as the *background hypothesis*.

There are only few studies in text comprehension research that were concerned with flashbacks. Baker (1978) and Ohtsuka and Brewer (1992) found that readers committed more errors in verifying or reporting the chronological order of described events when a text contained a flashback compared with when a text described the events in chronological order. This finding is in good agreement with the background hypothesis. However, it is also conceivable that the increased error rate was due to the fact that rearranging a representation is a rather demanding process, which may sometimes simply fail. There is indeed some evidence that flashbacks are more difficult to process than chronological descriptions (Baker, 1978; Kelter & Claus, 2005; Mandler & Goodman, 1982).[1] One could argue that there is another plausible interpretation of the difficulties in processing flashbacks, one that does not refer to a rearrangement of the representation. Zwaan and Radvansky (1998) proposed that the increase in processing load is due to the fact that nonchronological descriptions violate the default *iconicity assumption* of comprehenders. On the basis of this assumption, comprehenders expect the narrated order of events to match the events' chronological order (Hopper, 1979). However, it should be noted that in the case of proper flashbacks, the issue of a violation of the iconicity assumption is theoretically not a simple matter. As was mentioned earlier, flashbacks, though describing an event that occurred earlier, do not shift the reference time. We do not get deeper into this issue as it goes well beyond the scope of the present article. Moreover, recent results of Kelter and Claus suggest that the prolonged processing times for flashbacks are at least not solely due to a violation of comprehenders' expectations.

Carreiras and colleagues (1997) reported an experiment that is interesting with respect to the representation of flashback information, although the study was not explicitly concerned with flashbacks but rather with the processing of tense and aspect. Using a probe-recognition task, the authors tested the accessibility of a concept (e.g., economist) that was previously mentioned either in a sentence referring to the protagonist's past or in a comparable sentence describing the current situation (e.g., *Sometime in the past she worked as an economist for an international company* vs. *Now she works as an economist for an international company*). The target concept was found to be less accessible if it was previously mentioned in the sentence referring to the protagonist's past. At first glance, this result seems to support the chronological hypothesis. However, it can also be explained in terms of the background hypothesis when one makes the reasonable additional assumption that states of affairs that were communicated as background information are mentally less salient and hence less accessible than states of affairs belonging to a situation of the narrative main line.

[1] Studies of the processing of complex sentences with *before* or *after* have yielded mixed results as to the effect of order of mention (Hoeks et al., 2004; Mandler, 1986; Münte et al., 1998). However, we do not dwell on these results because the relation between the two clauses of complex sentences is not comparable to the relation between a flashback and its preceding text. The function of the subordinate clause in a complex temporal sentence is not to provide background information about the event described in the main clause but rather to denote the temporal region within which this event occurred (cf. Kamp et al., in press; Moens & Steedman, 1988).

Taken together, the available literature does not allow for clear conclusions as to the representation of the information provided by flashbacks.

The present study was designed to investigate the validity of the chronological hypothesis and the background hypothesis. More specifically, we examined whether the accessibility of an event that was mentioned in a flashback depends on the temporal distance between the point in time at which the event mentioned in the flashback occurred in the described world and the narrative now point. To understand the logic of the experimental paradigm used in the present study, let us consider a narrative mentioning four events that occurred immediately one after the other in the described world (Event 1–Event 2–Event 3–Event 4 [E1–E2–E3–E4]). In the narrative, the four events are mentioned in the order E2, E3, E1, E4, that is, E1 is mentioned in a flashback after the description of E3. The duration of E2 is stated as being either relatively short or long (e.g., the two protagonists argue for 5 min vs. for 3 hr about an issue). At the end of the narrative, the mental accessibility of an element of the flashbacked target event E1 is tested. The question of interest is whether the time needed for accessing this element depends on the stated duration of E2. If the chronological hypothesis is true, then readers represent the flashbacked target event E1 at the temporally appropriate location, that is, before the duration event E2. Thus, the flashbacked target event E1 is temporally more remote from the current narrative now point (E4) when a long duration of E2 was stated, compared with when a short duration of the duration event E2 was stated (see Figure 1). Given the effect of temporal distance on accessibility found in the aforementioned studies using chronological descriptions (Kelter et al., 2004; Rinck & Bower, 2000), the chronological hypothesis predicts that readers need more time to access an element of the flashbacked target event E1 if the duration of the duration event E2 was long than if it was short. In contrast, the background hypothesis does not predict this difference (see Figure 2). This hypothesis implies that readers attach the information about E1 (the flashback information) to the representation of the event described immediately before, that is, to the representation of event E3. Hence, there is no reason to expect that the time needed for accessing an element of the flashbacked target event E1 is affected by the duration of the duration event E2.

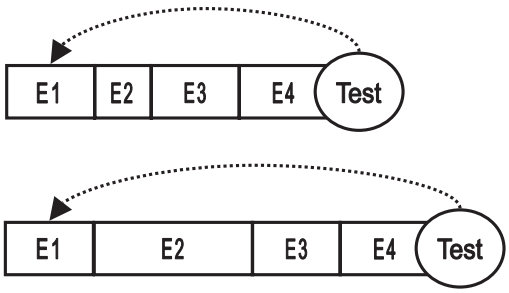

Figure 1. Structure of the mental representation constructed for the four events (Event 1–Event 4 [E1–E4]) according to the chronological hypothesis. The event mentioned in the flashback (E1) is integrated at its appropriate chronological position into the representation. Top: Representation constructed from a text that ascribed a relatively short duration to E2. Bottom: Representation constructed from a text that ascribed a relatively long duration to E2.

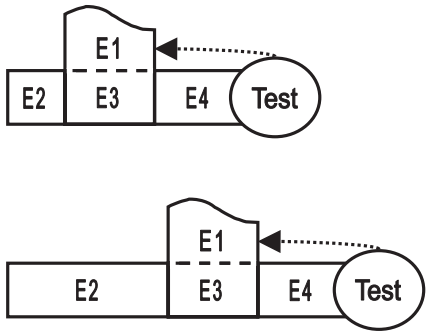

Figure 2. Structure of the mental representation constructed for the four events (Event 1–Event 4 [E1–E4]) according to the background hypothesis. The information about E1, mentioned in the flashback after E3, is attached to the representation of E3. Top: Representation constructed from a text that ascribed a relatively short duration to E2. Bottom: Representation constructed from a text that ascribed a relatively long duration to E2.

A drawback of this methodological approach is that a failure to find a temporal-distance effect cannot unambiguously be interpreted. It may either be due to the fact that readers represent the flashback information together with the immediately preceding information about E3, as the background hypothesis implies, or simply due to the fact that the described scenarios were inadequate to reveal a temporal-distance effect. For this reason, we first conducted an experiment using chronological versions of the passages in order to test the suitability of the materials. After having established a temporal-distance effect with the chronological passages, we examined the effect for nonchronological passages in Experiments 2, 3, and 4 as outlined before.

The accessibility of the target elements was tested either by measuring reading times for a sentence that contained an anaphoric expression referring to the target element (Experiments 1, 2, and 4) or by means of a probe-recognition task (Experiment 3). The passages used in the experiments were in German, and the participants were native German speakers.

Experiment 1

Method

Participants. Forty students at the Technical University of Berlin participated in the experiment. They either received a monetary reimbursement or participated to fulfill undergraduate requirements.

Materials. There were 24 experimental passages and 27 filler passages. All passages had a title, were 8 to 13 sentences long, and were written in the present tense. The present tense is used in German if a story, report, or anecdote is to be made especially vivid. The experimental passages were constructed according to the following scheme (see the sample passage in Table 1). After a short introduction, four events (E1, E2, E3, E4) were described that occurred consecutively in the narrated world. The events were mentioned in chronological order. The description of E1 (the target event) contained a sentence in which the target entity was mentioned. The target entity was a distinct, short-lived incident or action of the protagonist. After the target entity was mentioned, the termination of the target event E1 was described. The following sentence introduced E2 (the duration event), which, strictly speaking, was an activity or a process rather than an event. The duration of the event E2 was explicitly stated in the next sentence (the duration sentence). For each passage, there were two versions of this sentence, differing in an adverbial specifying the duration

Table 1
Sample Passage From Experiment 1, Translated From German

Event	Passage
(Title)	On vacation together
(Setting)	Heike and Frank are on vacation together in Southern France.
E1	They are sitting in a bistro and are very much in love.
	Frank even promises Heike to give up smoking.
	Then they go to the boardwalk because they want to enjoy the sunset.
E2	But instead they start to argue.
	For five minutes/For three hours they quarrel about Frank's mother [duration sentence].
E3	Now they're both really angry.
E4	They walk back to their hotel without saying a word [pretest sentence].
	Frank regrets his promise to give up smoking [anaphoric sentence].
	Because of the stress, he would like to smoke a cigarette.

Note. E1 = target event; E2 = duration event; E3 = filler event; E4 = final event.

of the event E2. The specific numeric values of the durative adverbials were selected for each passage individually on the basis of duration estimates collected from a separate group of 40 participants. These participants were presented with a booklet containing 40 items, 24 of which concerned the E2 activities or processes. In each item, the protagonist(s) and the current situation were first characterized briefly, and then the activity or process in question was described in a full sentence without giving any information about its duration. The participants were asked to specify (a) the typical duration of the event; (b) the longest, but still plausible, duration; and (c) the shortest, but still plausible, duration. The values corresponding to the 33rd percentile of the short-duration estimates were used in the short-duration sentence versions of the main experiment, and the values corresponding to the 67th percentile of the long-duration estimates were used in the long-duration sentence versions. The duration was stated in minutes or hours, depending on what sounded more natural. If necessary, the values were rounded (nine cases). Across the passages, the short duration ranged from 5 min to 1 hr and the long duration ranged from 1 hr to 8 hr. It was not possible to match the length of the duration sentences in the two versions. The short-duration versions were longer (mean number of syllables = 13.22, $SEM = 0.61$) than the long-duration versions (mean number of syllables = 11.35, $SEM = 0.62$), $t(22) = 8.84$, $p < .01$. The sentence following the duration sentence described E3 (filler event). The next sentence was the pretest sentence. It indicated the beginning of E4 (final event). This event was elaborated in the anaphoric sentence, in which anaphoric reference was made to the target entity. The anaphoric expression contained the same noun or verb that was previously used for denoting the target entity, or it contained a nominalization of the verb that was central in the previous mention of the target entity. The anaphoric sentence was followed by a sentence that completed the story.

The 27 filler passages were similar to the experimental passages with respect to topics, style, and length. Twenty-four of them served a different purpose, not related to the issue of temporal information processing. After each experimental and filler passage, a verification statement was presented to encourage the participants to read carefully.

Design and procedure. The 24 experimental passages were randomly assigned to two sets, A and B, comprising 12 passages each. Half of the participants received the passages of Set A in the short-duration version and the passages of Set B in the long-duration version. The other participants received the passages in the complementary versions. Thus, the two duration versions were assigned to participant groups and passage sets according to a 2 (group) × 2 (set) × 2 (duration) Latin square. Experimental and filler passages were presented to the participants in mixed random orders, except that each sequence started with three filler passages serving as warm-ups.

The passages were presented sentence by sentence on a computer screen in 14 point Palatino font. The participants advanced through the passages by pressing the space bar. The time interval between two consecutive presses of the space bar was defined as the reading time for the respective sentence. Upon the participant's pressing the space bar after reading the final sentence of a passage, a short warning signal appeared on the screen, and then the verification statement was presented. Participants indicated their positive or negative response by pressing the *l* key or *d* key, respectively.

Participants were tested individually. They were instructed to read the passages carefully at their normal reading pace. They were informed that they would be asked to verify a statement after each passage. It was not mentioned that reading times were being measured. To familiarize themselves with the procedure, participants completed two practice passages before starting the experiment.

Results and Discussion

Analyses were performed on the reading times for the anaphoric sentences, which were of main interest, and for the pretest sentences in order to assess potential spillover effects from the experimental manipulation. We also analyzed the reading times for the duration sentences for reasons to be pointed out later on. However, this analysis involved only 23 sentences per participant, because for one passage, the reading times for the duration sentence were not recorded because of a programming error. For all three types of sentences (pretest, anaphoric, duration), reading times shorter than 500 ms or longer than 10,000 ms were discarded. Outliers in the remaining data were determined separately for the three sentence types and by taking into account differences not only among participants but also among the individual sentences of each type. More specifically, for each individual sentence, the median of the reading times of the 40 participants was subtracted from each reading time for this sentence, providing a "difference score" for each reading time. Reading times were considered as outliers if their corresponding difference score deviated by more than 2.5 standard deviations from the participant's mean difference score for the respective sentence type. Altogether, 4.1%, 3.1%, and 3.4% of the reading times for the pretest, anaphoric, and duration sentences, respectively, were eliminated. In the experiments reported in this article, F_1 refers to tests against an error term based on participant variability. An alpha level of .05 was used for all statistical tests. We also present the results of the tests against an error term based on item variability (indicated by F_2), as some readers may want to see them. However, it should be borne in mind that, as we used counterbalanced designs, analyses by items are not relevant for testing the significance of the treatment effect (cf. Raaijmakers, Schrijnemakers, & Gremmen, 1999).

The mean reading times for the pretest sentence and the anaphoric sentence in the two duration conditions are displayed in Table 2. The 95% confidence intervals (CIs) given in parentheses are the CIs associated with the contrast between the two duration conditions. Following Masson and Loftus (2003), the CIs were computed from the within-subject mean-square-error term of a 2 (group) × 2 (duration) mixed analysis of variance (ANOVA) for the respective sentence type, with group being adopted from the before-mentioned Latin square. To test the statistical significance, the data for both sentence types were analyzed together. More

Table 2
Mean Sentence Reading Times (in Milliseconds) as a Function of the Duration of Event 2 in Experiment 1

Sentence	Short duration	Long duration	Difference
Pretest	1,948	1,926	−22 (95% CI = ±73)
Anaphoric	2,139	2,215	+76 (95% CI = ±57)

Note. CI = confidence interval.

specifically, for the analysis by participants, data were submitted to a 2 (group) × 2 (sentence type: pretest vs. anaphoric) × 2 (duration: short vs. long) ANOVA with group as the only between-subjects variable. The analysis by items was analogous, with set instead of group. Group and set were included in order to reduce error variance (cf. Pollatsek & Well, 1995). Lacking theoretical relevance, however, their effects are not reported.

The analyses yielded a significant main effect of sentence type, $F_1(1, 38) = 55.83$, $MSE = 41{,}155$, $p < .01$; $F_2(1, 22) = 4.98$, $MSE = 280{,}424$, $p = .04$, which is of little interest however, as the anaphoric and pretest sentences were not matched for length and complexity. The main effect of duration was not significant, $F_1(1, 38) = 1.42$, $MSE = 21{,}076$, $p > .20$; $F_2 < 1$. The Sentence × Duration interaction was significant; $F_1(1, 38) = 4.56$, $MSE = 20{,}849$, $p = .04$; $F_2(1, 22) = 3.56$, $MSE = 11{,}177$, $p = .07$. Planned comparisons showed that the reading times for the pretest sentences did not significantly differ in the two duration conditions (Fs < 1). However, for the anaphoric sentences, reading times were significantly longer in the long-duration condition than in the short-duration condition, $F_1(1, 38) = 7.22$, $MSE = 16{,}041$, $p = .01$; $F_2(1, 22) = 3.23$, $MSE = 14{,}478$, $p = .09$.

This result corresponds to our expectations. Reading times for the anaphoric sentence increased with the temporal distance between the incident referred to by the anaphor and the now point in the narrated world. However, one could suspect that this difference was due to a real-time difference between the two conditions. Possibly, the reading times for the duration sentence were longer in the long-duration version than in the short-duration version. This would mean that more real time elapsed between encoding the target entity and encountering the anaphoric sentence in the long-duration condition compared with the short-duration condition. It is true that with respect to sentence length, the two duration-sentence versions differed in a conservative direction, with the long-duration versions being shorter than the short-duration versions (see the *Materials* section), but when proceeding from the simulation view of language comprehension (cf. Barsalou, 1999; Barsalou, Simmons, Barbey, & Wilson, 2003; Glenberg & Kaschak, 2002; Zwaan, 2004), one may expect that the reading times for duration sentences stating a long duration are longer than those for sentences stating a short duration. Simulating a relatively long-lasting event should take more time than simulating a short event, all else being equal. An analysis of the reading times for the duration sentence ruled out the possibility that the results reflect a real-time difference. There was no substantial difference between the reading times in the long-duration condition ($M = 2{,}074$ ms) and the short-duration condition ($M = 2{,}092$ ms; $M_{\text{difference}} = -18$ ms, 95% CI = ± 93 ms, Fs < 1).

Note that the insignificant result for the duration sentence does not challenge the simulation view per se. It would be reasonable to assume that reading times are determined by both sentence length and the relative duration of the event to be mentally simulated. With regard to the simulation view, it is therefore of some interest to consider the reading times adjusted for sentence length. Accordingly, we computed a regression equation, predicting sentence reading time from number of syllables for each participant, using his or her reading times for all 239 sentences of the filler passages. On the basis of the regression equations, the reading-time residuals for the duration sentences were determined and then analyzed in the same way as the raw reading times.[2] Residual reading times were considerably longer in the long-duration condition ($M = -186$ ms) than in the short-duration condition ($M = -391$ ms), $M_{\text{difference}} = 205$ ms, 95% CI = ± 95 ms; $F_1(1, 38) = 19.26$, $MSE = 43{,}776$, $p < .01$; $F_2(1, 21) = 28.06$, $MSE = 17{,}172$, $p < .01$, thus supporting the simulation view.

Taken together, the results replicate those from the study by Kelter et al. (2004; see also Rinck & Bower, 2000). Reading times for the sentence referring anaphorically to the target event E1 were longer when the intervening duration event E2 was of a long duration compared with when it was of a short duration, suggesting that the accessibility of the target event E1 decreased with increasing temporal distance to the current reference time at the time of testing. This finding supports the assumption that the representations constructed by the readers captured the passage of time in the narrated world.

As to the specific objective of Experiment 1 in the present study, we conclude that the materials qualified as appropriate for our purposes. The scenarios and anaphoric sentences are suitable to reveal an effect of temporal distance on the accessibility of event E1.

Experiment 2

The goal of Experiment 2 was to determine the structure of the representation that readers construct when the information about the target event E1 is given in a flashback sentence. Participants were presented with nonchronological versions of the passages from Experiment 1 that mentioned the four events in the order E2, E3, E1, E4. The chronological hypothesis implies that the readers integrate the information about the flashbacked target event E1 at the chronologically appropriate position into their mental representation of the described world, that is, before the duration event E2 (see Figure 1). Thus, Experiment 2 should yield the same result as Experiment 1, that is, the reading times for the anaphoric sentence referring to the flashbacked target event E1 should be longer when the duration event E2 is of a relatively long rather

[2] Some readers may wonder why we did not use reading time per syllable as the dependent variable. The reason is that a quotient of reading time and sentence length is a biased estimate, as processes uncorrelated with sentence length are ignored (cf. Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). Syllable reading times are overestimated, and most important, the size of the overestimate is a function of the number of syllables of a sentence. Following the suggestion of one of the reviewers, we nevertheless analyzed the reading times per syllable for the duration sentences. The analyses yielded highly significant effects of duration in all four experiments.

than a short duration. In contrast, the background hypothesis does not predict such a temporal-distance effect. This hypothesis implies that the readers attach the information about the flashbacked target event E1 as background information to the filler event E3 (see Figure 2). Thus, resolving the anaphor in the anaphoric sentence should not be affected by the duration ascribed to the duration event E2.

Method

Participants. Forty students at the Technical University of Berlin took part in the experiment. They were either paid for their participation or took part in the experiment to fulfill undergraduate requirements.

Materials. There were 24 experimental passages and 27 filler passages. The experimental passages were nonchronological versions of the experimental passages used in Experiment 1, describing the four relevant events in the order E2, E3, E1, E4. A sample passage is given in Table 3. After the introductory sentences, the duration event E2 was described in two sentences. The second E2 sentence was the duration sentence, stating the same long or short duration as in Experiment 1. Again the long-duration sentence versions were shorter (mean number of syllables = 11.6, *SEM* = 0.59) than the short-duration versions (mean number of syllables = 13.5, *SEM* = 0.58), $t(23) = 9.26$, $p < .01$. The next sentence described the filler event E3. This was followed by two flashback sentences describing the flashbacked target event E1. The first flashback sentence contained the information that the flashbacked target event E1 preceded the duration event E2, with E2 being mentioned in a temporal prepositional phrase or subordinate clause (e.g., *Vor dem Streit* [*before their quarrel*] or *Bevor er mit der Arbeit begonnen hatte* [*before he had started working*], respectively). The second flashback sentence elaborated the description and mentioned the target entity, which was the same as in Experiment 1. The flashback sentences were in the past perfect, which indicated the temporal precedence of the event E1. The flashback sentences were followed by the pretest sentence, which resumed the narrative main line and described the beginning of the final event E4. The next sentence was the anaphoric sentence that made reference to the target entity. The descriptions of the individual events were as similar as possible to the corresponding descriptions in Experiment 1. However, the reorganization of the text information made some changes in the wording necessary. In particular, the end of the introductory section and the first sentence describing the duration event E2 were reformulated in such a way that the first E2 sentence "jumped" to the point in time at which the duration event E2 started. In addition, the pretest sentence had to be reformulated in order to indicate that it reverted to the narrative main line. As a consequence, the pretest sentences were slightly longer in this experiment than in Experiment 1. However, of course, the anaphoric sentences were always identical to those in Experiment 1.

The filler passages were similar to the experimental passages with regard to topics, style, and length. Twenty-four filler passages served a different purpose, not related to the issue of temporal information processing. As in Experiment 1, there was a statement for each passage that was to be verified by the participants.

Design and procedure. The design and procedure were identical to those in Experiment 1.

Table 3
Sample Passage From Experiment 2, Translated From German

Event	Passage
(Title)	On vacation together
(Setting)	Heike and Frank are on vacation together in Southern France. In the evening, they sit on a bench at the boardwalk and want to enjoy the sunset.
E2	But instead they start to argue.
	For five minutes/For three hours they quarrel about Frank's mother [duration sentence].
E3	Now they're both really angry.
E1	Before their quarrel, they had been sitting in a bistro and had been very much in love.
	Frank had even promised Heike to give up smoking.
E4	Now they walk back to their hotel without saying a word [pretest sentence].
	Frank regrets his promise to give up smoking [anaphoric sentence].
	Because of the stress, he would like to smoke a cigarette.

Note. E2 = duration event; E3 = filler event; E1 = flashbacked target event; E4 = final event.

Results and Discussion

Outliers were determined in the same way as in Experiment 1 (pretest sentence, 3.3%; anaphoric sentence, 2.9%; duration sentence, 2.3%). The participants' mean reading times for the pretest sentences and anaphoric sentences are displayed in Table 4. The main effect of sentence was not significant (Fs < 1), nor was the main effect of duration, $F_1(1, 38) = 1.04$, $MSE = 30{,}030$, $p > .20$; $F_2(1, 22) = 1.28$, $MSE = 23{,}932$, $p > .20$. The Sentence Type × Duration interaction was not significant in the analysis by participants, $F_1(1, 38) = 3.29$, $MSE = 17{,}971$, $p = .08$; $F_2(1, 22) = 4.35$, $MSE = 11{,}734$, $p = .049$. However, planned comparisons showed that the reading times for the pretest sentences did not differ significantly in the two duration conditions (Fs < 1), whereas the reading times for the anaphoric sentences were reliably longer in the long-duration condition than they were in the short-duration condition, $F_1(1, 38) = 4.52$, $MSE = 19{,}471$, $p = .04$; $F_2(1, 22) = 4.78$, $MSE = 16{,}819$, $p = .04$.

As in Experiment 1, we also analyzed the reading times for the duration sentence. To be sure, this analysis was not necessary with regard to the interpretation of the observed effect for the anaphoric sentence. In the present experiment, the duration sentence did not intervene between the introduction of the target entity and the anaphoric sentence. However, the reading times for the duration sentence are still interesting with respect to the simulation view of language comprehension. Raw reading times were slightly longer in the long-duration condition ($M = 2{,}164$ ms) compared with the short-duration condition ($M = 2{,}148$ ms; $M_{\text{difference}} = 16$ ms, 95% CI = ± 87 ms, Fs < 1). The analysis of the residuals from the regression on the number of syllables revealed a considerable difference between the two conditions ($M_{\text{difference}} = 250$ ms, 95% CI = ± 94 ms), with the residuals being longer in the long-duration condition ($M = -14$ ms) than in the short-duration condition ($M = -264$ ms), $F_1(1, 38) = 29.27$, $MSE = 42{,}711$, $p < .01$; $F_2(1, 22) = 22.63$, $MSE = 31{,}826$, $p < .01$.

The results of this experiment correspond to those of Experiment 1. The stated duration of the duration event E2 affected the reading times for the anaphoric sentence. The reading times were longer if the duration event E2 was of a relatively long duration, which implied that the flashbacked target event E1, containing the target entity, occurred a longer time ago in the described world, compared with when the duration event E2 was of a short duration, which implied that the flashbacked target event E1 occurred a shorter time ago. The important point here is that in the present experiment, the target entity was mentioned in a flashback after the duration sentence, which stated either a short or a long duration of

Table 4
Mean Sentence Reading Times (in Milliseconds) as a Function of the Duration of Event 2 in Experiment 2

Sentence	Short duration	Long duration	Difference
Pretest	2,200	2,189	−11 (95% CI = ±76)
Anaphoric	2,162	2,229	+67 (95% CI = ±63)

Note. CI = confidence interval.

the duration event E2. The fact that the temporal distance effect still occurred—and was even of a similar magnitude as in Experiment 1—suggests that the readers integrated the flashback information into their representation of the described world at the chronologically appropriate position. Thus, the result is consistent with the prediction of the chronological hypothesis. By contrast, the background hypothesis did not predict the observed effect of temporal distance.

However, one may question whether it is justified at all to draw conclusions from the temporal-distance effect as to the structure of the representation that readers built. When one adopts a probabilistic view on language processing (cf. Seidenberg & MacDonald, 1999), there is an alternative account of the effect in terms of comprehenders' probabilistic knowledge of the consequences of temporal expressions in narratives. There may be a statistical regularity of narratives concerning the time line in the described world, such that the more remote events are in the described world, the less frequent are anaphoric references to the events. If so, then the effect observed for the anaphoric sentence may simply reflect that the readers used their knowledge of this regularity. They may have been more prepared to encounter an anaphoric reference to the target entity when the flashbacked target event E1 was described as having occurred a shorter rather than a longer time ago. To investigate this account of the observed reading-time difference, we conducted a control experiment.[3]

Control Experiment

The control experiment consisted of a paper-and-pencil rating study on the experimental passages of Experiment 2. Readers were presented with the experimental passages up to and including the pretest sentence, either in the short-duration or the long-duration version, and rated the likelihood that the following sentence would anaphorically refer to the target entity (e.g., promise to give up smoking). If the probabilistic account of the temporal-distance effect observed in Experiment 2 is correct, then anaphoric reference to the target entity should be rated as being less likely for the long-duration version of the passages than for the short-duration versions. This outcome would not be incompatible with the chronological hypothesis. However, it would render the hypothesis superfluous, as the probabilistic account would provide a more parsimonious explanation of the temporal-distance effect observed in Experiment 2.

Should the likelihood ratings not differ in the two duration conditions, then this would not necessarily rule out the probabilistic account, as this could be due to the method's not being sufficiently sensitive. For this reason, the rating study included additional passages that served to check the sensitivity of the method. In the sensitivity-control passages, the success of the protagonist's goal was manipulated (e.g., the protagonist, being on a short trip in Lisbon, wants to get accommodation and either succeeds or fails in acquiring a room in one of the hotels listed in her guidebook), and subsequently an entity was mentioned that inferably was an alternative means to achieve the goal (e.g., the protagonist sees a youth hostel, when sitting in a sidewalk café). On the basis of existing empirical findings on the effects of manipulating goal success (Lutz & Radvansky, 1997; Suh & Trabasso, 1993), it can be expected that reference to the goal-related entity is rated to be less likely with the satisfied goal than with the unsatisfied goal.

Method

Participants. Thirty-two participants from the Technical University of Berlin and from the University of Potsdam took part in the experiment. They were either paid for their participation or participated to fulfill undergraduate requirements.

Materials. There were 24 experimental passages and 12 sensitivity-control passages, each in two versions. The 24 experimental passages were identical to the 24 experimental passages of Experiment 2 except that the anaphoric sentence and the final sentence were omitted. The 12 sensitivity-control passages were similar to the experimental passages with respect to topics, style, and length. The two versions of the sensitivity-control passages differed with respect to whether an initially mentioned goal of the protagonist was satisfied or not (e.g., the protagonist, who wants to have chic clothes for an award ceremony, either succeeds or does not succeed in buying something shortly before closing time). Later on, a target entity was mentioned that according to world knowledge would be an alternative means to achieve the goal (e.g., a trouser suit hanging in the bedroom of the protagonist's twin). The structure of the sensitivity-control passages corresponded to that of the experimental passages in that there were two sentences intervening between the manipulated sentence and the sentence introducing the target entity. For both the experimental and sensitivity-control passages, the target entity was always from the penultimate sentence of the presented part of the passage.

Below each of the 36 passages were three expressions naming three different entities (characters, objects, states of affairs, or events) that were mentioned in the respective passage. One of the three entities was the target entity of the respective passage, that is, the entity involved in the flashbacked target event E1 with the experimental passages or the goal-related entity with the sensitivity-control passages. The other two entities served as distractors. One of the distractors was always an entity that was mentioned midpassage. The other one was an entity mentioned either in the beginning or toward the end. The target entity and the two distractors were presented in randomized orders across passages.

Design and procedure. The sensitivity-control passages were randomly assigned to two sets. For the experimental passages, the same two sets as in Experiment 2 were used. The participants were randomly assigned to two groups. For both the experimental and the sensitivity-control passages, the two versions were assigned to the groups and sets according to a 2 (group) × 2 (set) × 2 (version) Latin square. The experimental and sensitivity-control passages were presented to the participants in mixed orders.

The passages were presented in a booklet with each passage printed on a separate page. Participants were instructed to read each passage carefully and to then consider each of the entities to be rated separately. They were asked to rate the likelihood that the following omitted sentence would refer

[3] We thank Maryellen MacDonald for drawing our attention to this alternative explanation and for her suggestion to conduct a rating study.

to the entity using a 7-point scale ranging from 1 (*very unlikely*) to 7 (*very likely*).

Results and Discussion

Analyses were performed on the ratings for the target entities, separately for the sensitivity-control and experimental passages. For the sensitivity-control passages, ratings significantly differed in the two conditions in the expected direction. Reference to the goal-related target entity was rated to be less likely in the satisfied-goal condition ($M = 5.19$) compared with the unsatisfied-goal condition ($M = 6.49$), $M_{\text{difference}} = 1.30$, 95% CI = $\pm$ 0.39; $F_1(1, 30) = 44.60$, $MSE = 0.61$, $p < .01$; $F_2(1, 10) = 27.98$, $MSE = 0.36$, $p < .01$. This indicates that the method was sufficiently sensitive to reveal an effect of comprehenders' expectations. For the experimental passages, ratings of the likelihood of reference to the target entity did not differ significantly between the short-duration condition ($M = 4.69$) and the long-duration condition ($M = 4.57$), $M_{\text{difference}} = 0.12$, 95% CI = $\pm$ 0.18; $F_1(1, 30) = 1.78$, $MSE = 0.12$, $p = .19$; $F_2(1, 22) = 1.14$, $MSE = 0.14$, $p = .30$.

The results do not corroborate the view that the reading-time difference for the anaphoric sentence observed in Experiment 2 was solely due to a difference between the two duration conditions with regard to the degree of which an anaphoric reference to the target entity was expected. To further check this conclusion, we reanalyzed the reading times for the anaphoric sentence, restricting the analysis to the data pertaining to those 11 experimental passages for which the likelihood ratings were higher in the long-duration condition than in the short-duration condition. The result for this particular set of passages corresponded to the result obtained for the entire set of passages. Reading times for the anaphoric sentence were still found to be significantly longer in the long-duration condition ($M = 2{,}279$ ms) than in the short-duration condition ($M = 2{,}152$ ms), $M_{\text{difference}} = 127$ ms, 95% CI = $\pm$ 103 ms, $F_1(1, 38) = 6.30$, $MSE = 50{,}809$, $p = .02$; $F_2(1, 9) = 8.08$, $MSE = 17{,}326$, $p = .02$. Thus, it seems safe to conclude that the reading-time difference for the anaphoric sentence was due to an effect of temporal distance rather than reflecting the degree to which an anaphoric reference to the target entity was expected.

Experiment 3

Experiment 3 addressed an additional concern regarding the conclusion that the temporal-distance effect on the accessibility of flashback information supports the chronological hypothesis. It is true that this effect is difficult to account for on the basis of the background hypothesis. According to this hypothesis, the readers integrated the flashback information as background information for the filler event E3. The accessibility of the flashback information should therefore not have been affected by the duration of the duration event E2 (see Figure 2). However, one may argue that testing the accessibility of the flashbacked target event E1 by means of an anaphoric sentence may not be the best way to evaluate the background hypothesis. In the anaphoric sentence, the flashbacked target event E1 was referred to as if it were a normal previously mentioned past event. This may have led the readers to take the flashbacked target event as an event in its own right, even if they had previously interpreted the information about the flashbacked target event as being merely background information about the filler event E3. Thus, maybe the anaphoric sentence itself caused the readers to integrate the information about the flashbacked target event E1 at the chronologically appropriate position into their representation. Integrating the information at its appropriate chronological position may then have taken more time if the position was temporally further away from the now point compared with when it was closer to it. Thus, the anaphoric sentence may have produced the effect itself by triggering a reorganization of the information.

In Experiment 3, readers were presented with nonchronological event descriptions (E2, E3, E1, E4) as in Experiment 2, but this time the accessibility of the flashbacked target event E1 was tested by means of a probe-recognition task. As the entity named by a probe word need not be integrated into the hitherto constructed representation, it is unlikely that a probe-recognition task gives rise to a chronological reorganization of the representation. Thus, if a temporal-distance effect on the probe-recognition latencies is found, then this would indicate that the flashbacked target event E1 was integrated at its appropriate chronological position before the probe was presented, presumably during the processing of the flashback.

An additional modification was that the flashbacks consisted of only one sentence instead of two. We reasoned that readers, when encountering a flashback, may initially tend to relate the information to the situation described in the preceding sentence in order to achieve local coherence. That is, they may interpret the first sentence of a flashback as giving background information. Only when encountering a second flashback sentence, dwelling on the past event, may they interpret the flashback as describing an event in its own right and integrate the flashback information at its chronologically appropriate position into their representation. Thus, it is conceivable that in Experiment 2 it was only the second flashback sentence that gave rise to the effect. Using single-sentence flashbacks should therefore give the background hypothesis a better chance.

Method

Participants. Forty students at the Technical University of Berlin took part in the experiment. They either were paid for their participation or took part in the experiment to fulfill undergraduate requirements. Two additional participants were replaced because the accuracy of their probe-recognition responses on the experimental items was not significantly better than chance (binomial test, eight or more errors, $p > .05$, one-tailed) and 5 additional participants were replaced because of very long recognition latencies (mean correct recognition latencies greater than 3,000 ms).

Materials. Twenty-four new experimental passages were constructed. The passages had a similar structure to those in Experiment 2, describing four consecutive events in the order E2, E3, E1, E4 (see the sample passage in Table 5). The values for the durative adverbials specifying the duration of the duration event E2 were determined in the same way as in Experiment 1, using the duration estimates from a separate group of 20 participants. The long-duration sentence versions were shorter (mean number of syllables = 12.5, $SEM = 0.74$) than the short-duration versions (mean number of syllables = 14.5, $SEM = 0.70$), $t(23) = 9.26$, $p < .01$. Unlike Experiment 2, the flashback consisted only of one sentence. Starting with a temporal prepositional phrase or subordinate clause, this sentence stated that the flashbacked target event E1 occurred before the duration event E2 in the described world. This sentence also contained the target word, which denoted an object, incident, or other kind of entity involved in E1 (e.g.,

Table 5
Sample Passage From Experiment 3, Translated From German

Event	Passage
(Title)	After a hard working day
(Setting)	Joachim is a building worker. Today he carried building materials up to the third story the whole day long.
E2	Coming home, he lies down on his bed, exhausted. He sleeps for three quarters of an hour/for ten hours [duration sentence].
E3	Then he awakes and notices that his two budgies are chirping cheerfully.
E1	Before lying down he had given them some food.
E4	Now Joachim stands up and goes into the bathroom. He washes his face with cold water [pretest sentence]. FOOD [probe]

Note. E2 = duration event; E3 = filler event; E1 = flashbacked target event; E4 = final event.

Becher [*mug*], *Telefonat* [*phone call*], *Futter* [*food*]). After the flashback sentence, the narrative main line was resumed and the final event E4 was described in two sentences. The second E4-sentence was the pretest sentence. It was followed by the probe word. All probe words were nouns with one to four syllables and 5 to 10 characters.

There were 30 filler passages. For 6 filler passages, the probe word was a noun that had been mentioned in the respective passage. For the other 24 filler passages, the probe word had not been mentioned. Thus, all in all, there were 56% positive probes and 44% negative probes. As in the previous experiments, a verification statement was formulated for each passage.

Design and procedure. The design and procedure were identical to those in the previous experiments except that when a participant pressed the space bar after having read the pretest sentence, the probe word was presented. It was presented in uppercase letters in the middle of the computer screen. Participants indicated their positive or negative response by pressing the *l* or *d* key, respectively. This was followed by the presentation of the verification statement, to which participants again responded by pressing either the *l* or *d* key. In the instruction, participants were asked to respond to the probe words as quickly and accurately as possible.

Results and Discussion

Mean percentages of probe-recognition errors in the experimental trials were 11.0% and 11.3% in the short-duration and long-duration condition, respectively. The difference between the two conditions ($M_{\text{difference}}$ = 0.3%, 95% CI = ± 4.2%) was not statistically significant ($Fs < 1$).

Analyses of the response latencies were based on the correct responses only. Latencies longer than 5,000 ms were not included in the analyses. As the latencies considerably decreased during the sequence of trials, determining outliers in the same way as in the previous experiments would have resulted in classing much more data from the first part of the experimental session as outliers than from the second part. To avoid this bias, we determined outliers separately for the first and second half of the trial sequence. Latencies that deviated more than 2.5 standard deviations from the mean in the respective half were discarded. Altogether, 5.2% of the latencies of correct probe responses were eliminated. For the pretest sentence and the duration sentence, outliers were determined in the same way as in the previous experiments (3.4% and 3.2%, respectively).

Mean reading times for the pretest sentence and mean latencies of correct probe responses are displayed in Table 6. The reading times for the pretest sentence and the latencies of the correct probe responses were analyzed separately. Reading times for the pretest sentence were significantly longer in the short-duration condition than in the long-duration condition, $F_1(1, 38) = 6.07$, MSE = 15,024, $p = .02$; $F_2(1, 22) = 7.35$, MSE = 7,864, $p = .01$. We have no explanation for this unexpected result. Note, however, that the difference is in the conservative direction. Most important, the probe-recognition latencies were reliably longer in the long-duration condition than in the short-duration condition, $F_1(1, 38) = 4.12$, MSE = 22,021, $p = .049$; $F_2(1, 22) = 3.11$, MSE = 22,690, $p = .09$.

Reading times for the duration sentence were analyzed in the same way as in Experiments 1 and 2. Raw reading times did not differ significantly between the long-duration condition (M = 2,155 ms) and the short-duration condition (M = 2,166 ms, $M_{\text{difference}}$ = –11 ms, 95% CI = ± 79 ms, Fs < 1). Residuals from the regression on the number of syllables were considerably longer in the long-duration condition (M = –53 ms) than in the short-duration condition (M = –292 ms), $M_{\text{difference}}$ = 239 ms, 95% CI = ± 70 ms; $F_1(1, 38) = 47.74$, MSE = 23,940, $p < .01$; $F_2(1, 22) = 20.09$, MSE = 30,782, $p < .01$.

The result for the probe response latencies parallels the result for the anaphoric sentence of Experiment 2. Participants needed more time to recognize the target entity if the flashbacked target event E1 occurred a longer time ago in the described world than if it occurred a shorter time ago. This result does not substantiate the suspicion that the temporal-distance effect observed in Experiment 2 was produced by the anaphoric sentence, prompting a reorganization of the hitherto constructed representation. Instead, the result suggests that the readers had inserted the flashbacked target event E1 into its appropriate position in the chronology of events already when processing the flashback. Moreover, the present experiment, using passages with flashbacks consisting of only one sentence, indicates that the readers had rearranged the text information already when confronted by the first sentence of a flashback.

Experiment 4

The purpose of this experiment was to address a concern as to the interpretation of the results in terms of an effect of temporal distance. Although it is true that the numerical values for the duration statements concerning the duration event E2 were selected on the basis of duration estimates, the long-duration events may still have deviated conceptually more from the respective event prototype compared with the short-duration events. As a

Table 6
Mean Reading Times for the Pretest Sentence and Mean Probe-Recognition Latencies (in Milliseconds) as a Function of the Duration of Event 2 in Experiment 3

Stimulus	Short duration	Long duration	Difference
Pretest sentence	1,830	1,762	−68 (95% CI = ±55)
Probe	1,960	2,027	+67 (95% CI = ±67)

Note. CI = confidence interval.

consequence, the long-duration versions of the narratives may have been less plausible, and the readers may have found it more difficult to construct and retain a representation of the narratives in the long-duration condition compared with the short-duration condition. We therefore considered it desirable to test the accessibility of the flashbacked target event E1 using three instead of two duration levels of the duration event E2, with the additional level being the typical duration of the duration event E2. Accessibility was tested by means of anaphoric sentences. The main point of interest was the length of the reading times in the typical-duration condition relative to the length of the reading times in the short- and long-duration conditions. If the reading times in the typical-duration condition turned out to be shorter than those in the two other conditions, then this would support the plausibility argument. In contrast, if the reading times monotonically increased from the short-duration condition to the typical-duration condition to the long-duration condition, then this would substantiate the claim that temporal distance is the crucial factor.

Method

Participants. Sixty students at the Technical University of Berlin took part in the experiment. They either were paid for their participation or took part in the experiment to fulfill undergraduate requirements.

Materials. There were 36 experimental passages and 15 filler passages. The structure of the experimental passages was the same as in Experiment 2. In fact, 24 experimental passages were taken from Experiment 2, with minor stylistic adjustments. For each passage, there were three versions of the duration sentence (short, typical, long). For the passages that were taken from Experiment 2, the numerical values for the short- and long-duration versions were identical to those in Experiment 2, and the values for the typical-duration versions were the median of the typical-duration estimates of the group of raters mentioned in the *Method* section of Experiment 1. For the new passages, the numerical values in the duration sentences were selected in the same way as those for the old passages (see the *Method* section of Experiment 1), using the duration estimates collected from a group of 15 participants. For the sample passage in Table 3, the three versions of the duration sentence read as follows: "For five minutes they quarrel about Frank's mother" (short-duration version), "For one hour they quarrel about Frank's mother" (typical-duration version), and "For three hours they quarrel about Frank's mother" (long-duration version). It was not possible to match the length of the duration sentences in the three versions. The long-duration sentence versions were on average shorter (mean number of syllables = 12.1, $SEM = 0.42$) than the typical-duration versions (mean number of syllables = 13.0, $SEM = 0.45$) and the short-duration versions (mean number of syllables = 13.8, $SEM = 0.44$), $F(2,70) = 31.51$, $MSE = 0.8220$, $p < .01$. As in the previous experiments, a statement was formulated for each passage that was to be verified by the participants after reading the passage.

Design and procedure. The experimental passages were randomly assigned to three sets, and the participants were randomly assigned to three groups. The three text versions were assigned to the groups and sets according to a 3 (group) × 3 (set) × 3 (version) Latin square. Experimental and filler passages were presented to the participants in mixed random orders. The procedure was identical to that of Experiments 1 and 2.

Results and Discussion

As sentence reading times considerably decreased in the course of the experimental session, outliers were determined separately for the first and second half of the experimental session. Within each half, outlying sentence reading times were determined in the same way as in the previous experiments, which resulted in the elimination of 2.6%, 3.2%, and 3.3% of the data for the pretest, anaphoric, and duration sentence, respectively.

The mean reading times for the pretest and anaphoric sentences in the three duration conditions are displayed in Table 7. The analysis yielded a significant main effect of sentence in the analysis by participants, $F_1(1, 57) = 24.69$, $MSE = 35{,}787$, $p < .01$; $F_2(1, 33) = 2.05$, $MSE = 250{,}122$, $p = .16$. The main effect of duration was not significant (Fs < 1). The Sentence × Duration interaction was significant, $F_1(2, 114) = 3.57$, $MSE = 15{,}168$, $p = .03$; $F_2(2, 66) = 3.35$, $MSE = 9{,}471$, $p = .04$. Planned comparisons showed that the reading times for the pretest sentences did not differ significantly in the three duration conditions (Fs < 1). For the anaphoric sentences, the effect of duration was significant, $F_1(2, 114) = 4.08$, $MSE = 12{,}005$, $p = .02$; $F_2(2, 66) = 2.95$, $MSE = 11{,}650$, $p = .06$. As can be seen from Table 7, there was a monotonic increase of the sentence reading times from the short-duration condition to the typical-duration condition to the long-duration condition. The difference between the short-duration and the long-duration condition proved significant, $F_1(1, 57) = 8.31$, $MSE = 11{,}001$, $p < .01$; $F_2(1, 33) = 4.62$, $MSE = 11{,}889$, $p = .04$, as did the difference between the short-duration and the typical-duration condition, $F_1(1, 57) = 5.09$, $MSE = 9{,}635$, $p = .03$; $F_2(1, 33) = 4.40$, $MSE = 10{,}870$, $p = .04$. The difference between the typical-duration and long-duration condition was not significant (Fs < 1).

As to the duration sentence, the raw reading times did not differ significantly in the three conditions (short duration, $M = 1{,}917$ ms; typical duration, $M = 1{,}872$ ms; long duration, $M = 1{,}861$ ms), $F_1(2, 114) = 2.46$, $MSE = 21{,}117$, $p = .09$; $F_2(2, 66) = 1.06$, $MSE = 24{,}806$, $p > .30$. Mean residual reading times from the regression on the number of syllables were –206 ms, –168 ms, and –79 ms in the short-, typical-, and long-duration condition, respectively, $F_1(2, 114) = 12.90$, $MSE = 19{,}793$, $p < .01$; $F_2(2, 66) =$

Table 7
Mean Sentence Reading Times (in Milliseconds) as a Function of the Duration of Event 2 in Experiment 4

	Duration			Difference	
Sentence	Short	Typical	Long	Typical−short	Long−typical
Pretest	2,038	2,018	2,011	−20 (95% CI = ±58)	−7 (95%CI = ±53)
Anaphoric	1,891	1,932	1,947	+41 (95% CI = ±36)	+15 (95% CI = ±45)

Note. CI = confidence interval.

5.94, $MSE = 25{,}794$, $p < .01$. That is, residuals increased from the short-duration to the typical-duration condition by 38 ms (95% CI = ± 54 ms), and from the typical-duration to the long-duration condition by 89 ms (95% CI = ± 55 ms).

The pattern of the reading times for the anaphoric sentence challenges the view that the results of the previous experiments were due to a difference in plausibility between the short-duration and the long-duration condition. If plausibility had been the crucial factor, then in the present experiment, the reading times should have been shortest in the typical-duration condition. Yet, the reading times monotonically increased from the short-duration condition to the typical-duration condition to the long-duration condition. This is exactly what is to be expected if accessibility depends on temporal distance. Thus, the result of Experiment 4 substantiates the conclusion that readers integrate flashback information into their representation at its appropriate chronological position.

General Discussion

This study investigated the structure of the representations that readers construct for narratives containing flashbacks. We considered two hypotheses as to how comprehenders deal with the information of a flashback According to the chronological hypothesis, comprehenders integrate the flashback information at its appropriate chronological position into their representation. In contrast, the background hypothesis states that comprehenders interpret a flashback as providing background information about the situation described just before and, accordingly, attach the flashback information to the representation of this situation. The results of our study support the chronological hypothesis. The mental accessibility of an event that was previously mentioned in a flashback was found to depend on when, according to the text, this event occurred in the described world. More specifically, the event's accessibility decreased with increasing temporal distance between the point in time at which the event occurred in the described world and the narrative now point at the time of testing. This temporal-distance effect indicates that the readers mentally integrated the event mentioned in the flashback at its proper place in the event chronology. This finding is remarkable, as nothing in the experimental setting called for this way of integrating the flashback information. The flashback sentences were coherent with their respective preceding sentence, and neither the instruction nor the comprehension test following each text placed special emphasis on the processing of temporal text information. Moreover, when one considers that the chronological integration of flashback information is likely to require extra effort (Kelter & Claus, 2005; Mandler & Goodman, 1982), this finding suggests that comprehenders attach considerable importance to the temporal dimension of the described world. This is consistent with the empirical evidence from several studies indicating that readers carefully attend to temporal information in narratives (e.g., Carreiras et al., 1997; Magliano & Schleich, 2000; Rinck & Bower, 2000; Rinck & Weber, 2003; Rinck et al., 2001; Zwaan, 1996; Zwaan et al., 1995, 2000).

With regard to principles of communication, the results of the present study appear surprising. They seem to imply that flashbacks violate the cooperative principle. A narrator using a flashback describes a sequence of events in an order that differs from the order in which comprehenders tend to represent events. Considering the frequency of flashbacks in oral and written narratives, it is unlikely that flashbacks are generally the result of careless language production planning by the narrator. What, therefore, might be the reason that narrators sometimes deviate from a strictly chronological description of events? Recall that linguistic analyses have revealed that a flashback usually has a close thematic relationship to its preceding sentence. This finding is, of course, not questioned by the results of the present study. Thus, one could entertain a variant of the background hypothesis as to the production of flashbacks, assuming that for the narrator, the thematic link of the flashback to the previous sentence is of primary importance. More specifically, one could assume that a flashback is produced when the narrator considers a state of affairs not pertinent to the narrative main line but critical for the understanding of one of the main events of the narrative. In other words, the narrator mentions this state of affairs in the narrative only because of its relationship to the respective main event and, accordingly, communicates the state of affairs immediately after that main event rather than at its proper chronological place. By using a flashback, the narrator provides the information at exactly the point in the text where the comprehenders are meant to use it. Hence, the production of a flashback conforms to the cooperative principle. The question is whether the comprehenders recognize the narrator's intent. Note that the results of the present study do not rule out that they do so. It would be perfectly compatible with our results that comprehenders, when encountering a flashback sentence, initially update the representation of the just described situation according to the new information and then store the event mentioned in the flashback at the chronologically appropriate location in the representation of the entire narrative. Thus, although our results support the chronological hypothesis, the background hypothesis cannot yet be ruled out as a complementary (rather than contradictory) hypothesis. In any case, the results of the present study clearly show that the structure of a text is not a blueprint for the structure of the representation that comprehenders construct for a narrative. Comprehenders mentally organize described events according to temporal criteria, even if the text does not mention the events in chronological order.

As we noted earlier, one main finding from text comprehension research is that at any moment during reading a narrative, the protagonist's current situation is easier to mentally access than are situations that occurred in the past (Anderson et al., 1983; Bestgen & Vonk, 1995; Carreiras et al., 1997; Madden & Zwaan, 2003; Magliano & Schleich, 2000; Rinck & Bower, 2000; Rinck & Weber, 2003; Zwaan, 1996; Zwaan et al., 1995, 2000). This indicates that in the representation a distinction is made between the present and the past of the described world. Our study, together with other studies in which a temporal-distance effect was found (Kelter et al., 2004; Rinck & Bower, 2000), extends this finding. Demonstrating that the accessibility of past events depends on how long ago they occurred in the described world, it provides evidence for a temporal differentiation within the past of the described world. Moreover, as the experimental conditions in our study did not differ with respect to the number of events occurring between the critical event and the narrative now point, it can be concluded that this differentiation is not merely at an ordinal level (i.e., in terms of the order of events) but concerns the length of time periods.

This raises the question as to how the length of time periods is coded in the representation of a narrated world. One possibility is that it is coded extrinsically, for example, in propositions (for the extrinsically–intrinsically distinction, see Palmer, 1978). Accordingly, one would assume that the representation of a narrated world contains propositions encoding the duration of the described events as well as propositions encoding the events' temporal distances from the narrative now point, derivable from the information about the duration of the respective subsequent events. However, this does not provide an explanation of the observed temporal-distance effect. Why should it take less time to access an event if a distance proposition contains the argument, for example, "five minutes" compared with when it contains the argument "three hours"? It would be necessary to in addition assume that the access processes operating on the distance propositions are faster or slower depending on the metric information contained in these propositions. This assumption would be completely ad hoc and, moreover, an alien element in a propositional theory.

The second possibility is that the length of time periods is coded intrinsically, that is, by a property of the representation for the narrated world itself. This possibility is particularly plausible when one adopts the simulation view of language comprehension (e.g., Barsalou, 1999; Barsalou et al., 2003; Glenberg & Kaschak, 2002; Zwaan, 2004). According to this view, people understand the description of a situation by mentally simulating the experience of this situation. Thus, when reading a description of successive events, they construct a coherent dynamic representation as they do when experiencing an evolving sequence of events (cf. Kelter et al., 2004). To date, there is still wide divergence among the theories of mental time measurement in real-life experience (cf. Bradshaw & Szabadi, 1997; Helfrich, 2003; Matell & Meck, 2000). With regard to the simulation view, the multiple-time-scale theory (Staddon, 2005; Staddon & Higa, 1999) appears especially attractive, as it assumes that the passage of time is intrinsically coded in event representations. Roughly speaking, the central idea of the multiple-time-scale theory is that an event is stored in multiple memory traces, which have different decay rates, so that the information as to how long ago the event occurred is coded in the strengths of its memory traces. Future research is needed to clarify whether and how the multiple-time-scale theory can be adopted for language comprehension.

An especially interesting implication of the simulation view of language comprehension concerns the time needed to process a sentence describing an event. According to this view, the reading time for a sentence depends not only on the length of the sentence and other variables affecting the reading process (in the narrow sense) but also on the time needed to simulate the experience of the described event. Thus, reading times for sentences describing events with a relatively long duration should be longer than reading times for comparable sentences describing events with a shorter duration. Our results for the duration sentences are consistent with this prediction. After removing the effect of sentence length, the reading times for the sentence describing the duration event E2 were found to be longer when the sentence stated that this event lasted a relatively long time compared with when it stated that it lasted a relatively short time. Matlock (2004) observed a similar effect with a sentence-verification task in which the target sentences were related to passages implying either a short or a long duration. These results are promising with regard to carrying out more precise tests of this implication of the simulation view.

Research on the processing and representation of temporal information in text comprehension is still at an initial stage, but the existing studies, including the present one, have already clearly shown that comprehenders carefully attend to the temporal dimension of the narrated world. Why is the temporal dimension so important? This question can presumably not be answered by text comprehension research. Rather, the role that the temporal dimension plays for cognition in general has to be taken into consideration. People experience the world as dynamic. They take the continuous progression of time for granted and represent the variability on other dimensions as a function of time (Navon, 1978; see also Miller & Johnson-Laird, 1976). In other words, the stream of experience is encoded in terms of a time-course function, in which time constitutes the x variable and the nontemporal dimensions constitute y variables. In order to act in a dynamic world, it is necessary to extrapolate the future. This can only be done if regularities in the time-course function are detected during experience and stored in memory. Research on the representational momentum has shown that when viewing a dynamic stimulus pattern, people virtually automatically detect monotonic trends or periodic changes on its nontemporal dimensions and anticipate its further development (e.g., Freyd, 1987; Thornton & Hubbard, 2002). Furthermore, regularities consisting of the repeated occurrence of a particular section of the time-course function at different intervals provide the basis for individuating events (Avrahami & Kareev, 1994). Similarly, if two or more events frequently occur one after the other with the intervals between them being fairly constant, then these events are stored in memory as a sequence with a particular temporal structure (cf. van der Meer, Beyer, Heinze, & Badel, 2002). The detection and knowledge of regularities in the time-course function are intimately related to the construal of causation (e.g., Buehner & May, 2003; Hubbard & Ruppel, 2002). People feel they understand a situation as long as it develops corresponding to familiar regularities of the time-course function, and deviations from these prompt a search for special causes. Thus, the mental representation of the time course of a dynamic situation is a prerequisite for understanding. This may also hold for text comprehension. Comprehenders need a temporally organized representation to understand the narrated world.

References

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology, 35A,* 427–440.

Avrahami, J., & Kareev, Y. (1994). The emergence of events. *Cognition, 53,* 239–261.

Baker, L. (1978). Processing temporal relationships in simple stories: Effects of input sequence. *Journal of Verbal Learning and Verbal Behavior, 17,* 559–572.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577–660.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7,* 84–91.

Bestgen, Y., & Vonk, W. (1995). The role of temporal segmentation markers in discourse processing. *Discourse Processes, 19,* 385–406.

Bestgen, Y., & Vonk, W. (2000). Temporal adverbials as segmentation markers in discourse comprehension. *Journal of Memory and Language, 42,* 74–87.

Bradshaw, C. M., & Szabadi, E. (Eds.). (1997). *Time and behaviour: Psychological and neurobehavioural analyses.* Amsterdam: Elsevier.

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *Quarterly Journal of Experimental Psychology, 56A,* 865–890.

Carreiras, M., Carriedo, N., Alonso, M. A., & Fernández, A. (1997). The role of verb tense and verb aspect in the foregrounding of information during reading. *Memory & Cognition, 25,* 438–446.

de Vega, M., Robertson, D. A., Glenberg, A. M., Kaschak, M. P., & Rinck, M. (2004). On doing two things at once: Temporal constraints on actions in language comprehension. *Memory & Cognition, 32,* 1033–1043.

Ferreira, F., & Clifton, C., Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348–368.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review, 94,* 427–438.

Gennari, S. P. (2004). Temporal references and temporal relations in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 877–890.

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes, 23,* 265–304.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9,* 558–565.

Graesser, A. C., Kassler, M. A., Kreuz, R. J., & McLain-Allen, B. (1998). Verification of statements about story worlds that deviate from normal conceptions of time: What is true about Einstein's dreams? *Cognitive Psychology, 35,* 246–301.

Hagmayer, Y., & Waldmann, M. R. (2002). How temporal assumptions influence causal judgments. *Memory & Cognition, 30,* 1128–1137.

Helfrich, H. (Ed.). (2003). *Time and mind: Vol. 2. Information processing perspectives.* Cambridge, MA: Hogrefe & Huber.

Hoeks, J. C. J., Stowe, L. A., & Wunderlink, C. (2004). Time is of the essence: Processing temporal connectives during reading. In K. D. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 577–582). Mahwah, NJ: Erlbaum.

Hopper, P. J. (1979). Aspect and foregrounding in discourse. In T. Givón (Ed.), *Syntax and semantics: Vol. 12. Discourse and syntax* (pp. 213–241). New York: Academic Press.

Hubbard, T. L., & Ruppel, S. E. (2002). A possible role of naïve impetus in Michotte's "launching effect": Evidence from representational momentum. *Visual Cognition, 9,* 153–176.

Hyönä, J. (1995). An eye movement analysis of topic-shift effect during repeated reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1365–1373.

Kamp, H., van Genabith, J., & Reyle, U. (in press). Discourse representation theory. In D. Gabbay & F. Guenthner (Eds.), *Handbook of philosophical logic.* Dordrecht, The Netherlands: Kluwer Academic.

Kelter, S., & Claus, B. (2005). How do readers deal with flashbacks in narratives? In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the Twenty-seventh Annual Conference of the Cognitive Science Society* (pp. 1126–1131). Mahwah, NJ: Erlbaum.

Kelter, S., Kaup, B., & Claus, B. (2004). Representing a described sequence of events: A dynamic view of narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 451–464.

Lascarides, A., & Asher, N. (1993a). A semantics and pragmatics for the pluperfect. *Proceedings of the European Chapter of the Association of Computational Linguistics, 6,* 250–259.

Lascarides, A., & Asher, N. (1993b). Temporal interpretation, discourse relations, and commonsense entailment. *Linguistics and Philosophy, 16,* 437–493.

Levine, W. H., & Klin, C. M. (2001). Tracking of spatial information in narratives. *Memory & Cognition, 29,* 327–335.

Lutz, M. F., & Radvansky, G. A. (1997). The fate of completed goal information in narrative comprehension. *Journal of Memory and Language, 36,* 293–310.

Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory & Cognition, 31,* 663–672.

Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models. *Discourse Processes, 29,* 83–112.

Mandler, J. M. (1986). On the comprehension of temporal order. *Language and Cognitive Processes, 1,* 309–320.

Mandler, J. M., & Goodman, M. S. (1982). On the psychological validity of story structure. *Journal of Verbal Learning and Verbal Behavior, 21,* 507–523.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57,* 203–220.

Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *BioEssays, 22,* 94–103.

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition, 32,* 1389–1400.

Miller, G. A., & Johnson-Laird, P. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics, 14,* 15–28.

Morrow, D. G. (1985). Prepositions and verb aspect in narrative understanding. *Journal of Memory and Language, 24,* 390–404.

Morrow, D. G. (1990). Spatial models, prepositions, and verb-aspect markers. *Discourse Processes, 13,* 441–469.

Münte, T. F., Schiltz, K., & Kutas, M. (1998, September 3). When temporal terms belie conceptual order. *Nature, 395,* 71–73.

Navon, D. (1978). On a conceptual hierarchy of time, space, and other dimensions. *Cognition, 6,* 223–228.

Ohtsuka, K., & Brewer, W. F. (1992). Discourse organization in the comprehension of temporal order in narrative texts. *Discourse Processes, 15,* 317–336.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 259–303). Hillsdale, NJ: Erlbaum.

Polanyi, L., & van den Berg, M. H. (1996). Discourse structure and discourse interpretation. In P. Dekker & M. Stokhof (Eds.), *Proceedings of the Tenth Amsterdam Colloquium* (pp. 113–131). Amsterdam: University of Amsterdam, ILLC/Department of Philosophy.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 785–794.

Proctor, A. S., Dickey, M. W., & Rips, L. J. (2004). *The time-course and cost of telicity inferences.* In K. D. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 1107–1112). Mahwah, NJ: Erlbaum.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "The language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41,* 416–426.

Rapp, D. N., & Taylor, H. A. (2004). Interactive dimensions in the construction of mental representations for text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 988–1001.

Rinck, M., & Bower, G. H. (2000). Temporal and spatial distance in situation models. *Memory & Cognition, 28,* 1310–1320.

Rinck, M., Gámez, E., Díaz, J. M., & de Vega, M. (2003). Processing of

temporal information: Evidence from eye movements. *Memory & Cognition, 31,* 77–86.

Rinck, M., Hähnel, A., & Becker, G. (2001). Using temporal information to construct, update, and retrieve situation models of narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 67–80.

Rinck, M., & Weber, U. (2003). Who when where: An experimental test of the event-indexing model. *Memory & Cognition, 31,* 1284–1292.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science, 23,* 569–588.

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language, 53,* 125–140.

Staddon, J. E. R. (2005). Interval timing: Memory, not a clock. *Trends in Cognitive Sciences, 9,* 312–314.

Staddon, J. E. R., & Higa, J. J. (1999). Time and memory: Towards a pacemaker-free theory of interval timing. *Journal of the Experimental Analysis of Behavior, 71,* 215–251.

Suh, S., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language, 32,* 279–300.

ter Meulen, A. G. B. (2000). Chronoscopes: The dynamic representation of facts and events. In J. Higginbotham, F. Pianesi, & A. C. Varzi (Eds.), *Speaking of events* (pp. 151–168). New York: Oxford University Press.

Thornton, I. M., & Hubbard, T. L. (Eds.). (2002). *Representational momentum: New findings, new directions.* Hove, England: Psychology Press.

Todorova, M., Straub, K., Badecker, W., & Frank, R. (2000). Aspectual coercion and the online computation of sentential aspect. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 523–528). Mahwah, NJ: Erlbaum.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33,* 285–318.

van der Meer, E., Beyer, R., Heinze, B., & Badel, I. (2002). Temporal order relations in language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 770–779.

Webber, B. L. (1988). Tense as discourse anaphor. *Computational Linguistics, 14,* 61–73.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1196–1207.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation (*Vol. 44, pp. 35–62). New York: Academic Press.

Zwaan, R. A., Madden, C. J., & Whitten, S. N. (2000). The presence of an event in the narrated situation affects its availability to the comprehender. *Memory & Cognition, 28,* 1022–1028.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 386–397.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185.

Received October 19, 2005
Revision received March 15, 2006
Accepted April 11, 2006 ■