Processing and Representing Temporal Information in Narrative Text

David J. Therriault
Department of Educational Psychology
University of Florida

Gary E. Raney
Department of Psychology
University of Illinois at Chicago

According to current theories in discourse research, readers monitor a series of 5 situational dimensions during narrative comprehension (Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). These dimensions are time (e.g., the order of events), space (e.g., locations), protagonist (e.g., main character actions), causality (e.g., how one event influences another event), and intentionality (e.g., goals). These experiments were designed to further explore how readers process and represent time (duration) in situation-model construction. In 3 experiments, we examined how duration-related inconsistencies influenced processing time and processing strategies. Results indicate that readers routinely monitor the duration of events and detect temporal inconsistencies even when temporal information is implicitly presented. This provides evidence for the representation of duration information in situation models.

An important goal in discourse research is identifying and understanding what elements of text might be represented in memory when we read. Prior research provides evidence that readers create different levels of representation (i.e., surface, textbase, and situation model; Fletcher & Chrysler, 1990; Kintsch, Welsch, Schmalhofer, & Zimny, 1990; van Dijk & Kintsch, 1983). The situation model is described as a representation that includes information about the reader's understanding of the text in addition to the text proper. Thus, a situation model may in-

---

Correspondence should be addressed to David J. Therriault, Department of Educational Psychology, University of Florida, Gainesville, FL 32611–7047. E-mail: therriault@coe.ufl.edu

clude a reader's background knowledge of the domain associated with the text and inferences that are beyond the scope of the text.

The significance of the situation model has not been lost on cognitive psychologists. Indeed, there has been considerable research exploring the nature of situation models since 1983 (see Zwaan & Radvansky, 1998, for a review). One contribution has been the development of the event-indexing model. According to Zwaan, Langston, and Graesser (1995), events are the focal points, or glue, that hold together a situation model. Events are organized in memory based on a series of five specific dimensions: space, causality, protagonist, intentionality, and time.

Research exploring these dimensions provides insight into the complex nature of readers' representations (Zwaan & Radvansky, 1998). With respect to temporal relations, research demonstrates that comprehension processing increases when the discourse structure does not match the order of presented events (e.g., a flashback is employed), when a connective or shift signifies a longer versus shorter time frame (e.g., a moment later vs. an hour later) and when explicit time inconsistencies are introduced in the text (Bestgen & Vonk, 1995; Ohtsuka & Brewer, 1992; Rinck, Gámez, Díaz, & de Vega, 2003; Rinck, Hahnel, & Becker, 2001; Zwaan, 1996). Although we are certainly making progress, there is still much that we do not understand about how readers process and represent time in discourse. In the experiments, we hoped to demonstrate that one element of time, specifically duration, is monitored by readers and represented in memory as part of the situation model.

One difficulty in exploring time can be attributed to its conceptualization. Ostensibly, there is no one agreed-on definition of time, and there appear to be many different senses, aspects, or types of time that humans could potentially monitor. The quantity and description of temporal experiences vary considerably among psychologists (Block, 1990; Fraisse, 1963; Friedman, 1990; Ornstein, 1972, 1997) and linguists (Harnish, 1976; Labov, 1972; Labov & Waletzky, 1967; ter Meulen, 1995). For example, duration, succession, simultaneity, order, temporal perspective (or a sense of past or present), rhythm, impersonal time (socially constructed units of time), and alternative senses of time are just some of the conceptual components explored in temporal research.

This example underscores the complexity of studying time. Discourse researchers who include time as an important element of the situation model often do not acknowledge the multiplicity (i.e., different conceptualizations) of time in their research (Friedman, 1990). To make claims about how readers process, represent, and recall time, researchers need to be clear about what type (i.e., sense) of time is being examined.

Duration is clearly an important, perhaps fundamental, aspect of time. Duration is included (in some form) in all of the temporal research we examined. Duration can be understood as our ability to represent the relative time course of some event and possibly recall it later. Consequently, we believe duration is a suitable starting

point in examining readers' temporal representations. To our knowledge, duration has not been explicitly examined as its own element of discourse that may be represented by readers. However, researchers have used duration to manipulate discourse, such as characters, aspect, or goals (Anderson, Garrod, & Sanford, 1983; Magliano & Schleich, 2000; Singer & Richards, 2005). Duration has also been explored in combination with other situation dimensions (Scott Rich & Taylor, 2000) or other types of time (Rinck et al., 2001, 2003). For the purposes of our research, Rinck et al.'s work is most relevant.

Rinck et al. (2001) examined how readers process inconsistent temporal sequence information. An example of a temporally inconsistent passage from Rinck et al. follows:

> Today, Markus and Claudia would finally meet again.
> Markus's train arrived in Dresden on time at 4:10 p.m., and Claudia's train arrived at 4:30 p.m.
> Markus was very excited when his train stopped at Dresden Central Station just as scheduled.
> He tried to think of what he should say when he met her.
> Many people were crowding around the platform.
> Claudia was already waiting for him when he got of the train with his huge bag. (p. 79)

Their experiments provide evidence that reading time increased when readers encountered inconsistent temporal information. Interestingly, this was the case whether the readers noticed the inconsistencies or not (and effect sizes were quite large). Rinck et al. (2001) concluded that their results confirmed the importance of the temporal dimension of situation models.

Rinck et al. (2001) provided a strong case for the inclusion of temporal information in situation-model construction. However, they were admittedly noncommittal about how readers use different aspects of time when constructing situation models. They stated, "So far, we have been deliberately vague about the exact way temporal information is represented in situation models. … Further research will be needed to identify the factors that determine the way temporal information is represented in situation models" (p. 77). For example, the authors described their time manipulations as blunt. Specific time terms were employed, and the inconsistencies were direct contradictions. It is unclear whether readers would attend to temporal information without the use of specific time terms and direct contradiction. Also, the purpose of Rinck et al.'s experiments was to manipulate temporal order, but other senses of time may have inadvertently influenced readers. In the prior example, the inconsistency must be worked out based on given information much like a high school math problem. In other passages, the inconsistency must be inferred from both order and duration information (i.e., participants are pre-

sented with an inconsistency regarding which–of two–trains arrive first at a station, but the inconsistency can only be inferred if participants attend to the duration of the train trips).

Rinck et al.'s (2001) research provided insight into the importance of temporal information in situation-model development. Additional research is needed to document how readers make use of different types of time, which types of time comprehension are necessary for building a coherent situation model, and which are routinely monitored. The purpose of the research described here is to begin to fill this research gap about how readers process, understand, and represent duration.

We present three experiments that provide insight on how duration is utilized when readers construct situation models. Experiment 1 examined whether readers were sensitive to inconsistencies in the duration of events described in a text using the inconsistency paradigm employed by Albrecht and O'Brien (1995) and Rinck et al. (2001). Experiment 2 expanded on the design of Experiment 1 by measuring readers' reactions to duration-related comprehension problems when duration terms were not explicitly mentioned. Experiment 3 provides information about readers' awareness of event durations and the strategies they used to process the texts.

EXPERIMENT 1

Albrecht and O'Brien (1995) and Rinck et al. (2001) used the inconsistency paradigm to discern whether readers were sensitive to specific types of information in text. We adopted this procedure to explore whether readers were sensitive to inconsistencies regarding the duration of text events.

Experimental passages contained a sentence that was either consistent with the normal duration of a particular activity or inconsistent with the normal duration of a particular activity. The sample passage depicted in Table 1 describes eating in a restaurant, and different time terms were used to describe that activity. In the consistent version, a sentence describes two friends who spend 1 hr in a restaurant (normal amount of time). In the inconsistent versions, two friends spend 7 hr in a restaurant (an unusually long time) or 5 min in a restaurant (an unusually short time).

Using the duration term *1 hr* is generally consistent with what we know about the average time spent in a restaurant. Readers were expected to read the critical sentence fastest in this condition. The other duration terms are less consistent with our conceptualization of the normal amount of time spent at a restaurant given the context of the passage. It is possible, but unlikely, to be in a restaurant for only 5 min or for as long as 7 hr. If readers have trouble (indicated by increased reading times) processing the inconsistent time sentences, this would provide direct evidence that temporal inconsistencies disrupt reading and indirect evidence that

TABLE 1
Sample Passage for Experiment 1 Demonstrating the Three Conditions:
Consistent, Short Inconsistent, and Long Inconsistent

Sally had just finished her last college final in Organic Chemistry and was going home for the summer soon. Sally decided to meet her friend Cleo for lunch that day. Cleo and Sally were old friends from high school but Cleo was attending a different college across town. Sally and Cleo went to a small Thai restaurant that was located between their two schools.

Sally and Cleo spent one hour (consistent) at the restaurant.
Sally and Cleo spent five minutes (inconsistent–short) at the restaurant.
Sally and Cleo spent seven hours (inconsistent–long) at the restaurant.

Then Sally decided to head back. Sally was glad that she had a chance to talk with Cleo. Sally was looking forward to seeing Cleo over the summer break.

Comprehension Questions (true or false)
  1. Sally's exam was in Political Science.
  2. Sally and Cleo spent very little time together in high school.

readers represent the duration of an event described in the text. If reading times do not increase in the inconsistent versions, this suggests that readers do not monitor information about the duration of events in text while reading. We predicted that reading times for the long and short inconsistent duration target sentences would be longer than the consistent duration target sentences.

Method

Passage development and norming.   To provide a stable base of duration times to draw on, when developing the stimuli, mean duration estimations for a list of 25 personal events were collected. Forty-eight participants provided duration estimates for the activities. Duration inconsistencies were developed based on the average event durations from the norming study. Of the 25 activities, 12 were selected for use in experimental passages. For all three versions of a given passage, target sentences had the same number of words. Character length was also roughly equated across conditions (i.e., within two characters). Spillover sentences (i.e., sentences following the target sentences) were identical across the three versions of each passage (see Table 2).

Placing an event in the context of a passage may lead to different estimates of that event's duration than estimates generated by considering the event in isolation; therefore, the durations used in each passage were evaluated within their context using a ranking procedure. Twenty participants completed the ranking procedure. Participants were provided with the context of the passage and decided the appropriateness of each of the possible duration sentences. The order of options in the

TABLE 2
Mean Duration Estimations for Example Activities Used
in Experiment 1 and 2

Activity	Estimate
Tying your shoes	18.4 s
Brushing your teeth	2.4 min
Showering	16.5 min
Commuting to school or work	48 min
Walking a mile	19 min
Closing a door	2.2 s
Talking a nap	1.4 hr
Writing a check	50 s
Getting a haircut	29 min
Average telephone call	27 min
Time you spend sleeping on an average night	6.5 hr
Average vacation duration	1.3 weeks
Preparing a meal	48 min
Eating at a restaurant	38.5 min
Reading a book in one setting	3 hr
Working out	62 min
Getting dressed	9.7 min
Duration of your childhood	12.2 years
Watching TV	2.3 hr
Shopping	3 hr
Visiting the dentist	1.1 hr
Drinking a cup of coffee	17 min
Spending time at the beach	2.8 hr
Playing a board game	1.2 hr
Classes for the week	14 hr

ranking guide was randomized to control for potential order effects. This ranking procedure provided information that was used to ensure that the duration of events described in context matched their duration out of context in terms of which duration was the most natural (consistent) and which durations were least natural (short or long inconsistent). Thus, the ranking procedure provided information regarding the relative plausibility of the duration of events. The results were used to construct experimental passages in which plausibility was controlled (i.e., ensure that a short or long event was not the most plausible event within the context of a passage). The final passage rankings for the conditions were as follows: consistent (1.1), short inconsistent (2.5), and long inconsistent (2.4).

Participants and design.   Seventy-eight students from the University of Illinois at Chicago participated to fulfill an introductory psychology course requirement. The design was a one-factor, within-subjects design with three groups (con-

sistent duration, inconsistent short duration, and inconsistent long duration). The primary dependent measure was reading time per sentence. The second dependent measure, accuracy on comprehension questions, was evaluated to ensure that participants attended to the passages.

Participants read 32 passages (20 filler and 12 experimental). Two additional practice passages (filler passages) were presented at the beginning of the experiment to familiarize participants with the task. Filler passages did not contain time inconsistencies and were included to draw attention away from the duration inconsistencies (as suggested by Rinck et al., 2001). Of the 12 experimental passages, 4 were consistent, 4 were short inconsistent, and 4 were long inconsistent. In the overall design, only 25% of the passages contained an inconsistency. The order of passages was randomized.

Two comprehension questions were administered after reading each passage. One comprehension question was designed to tap surface or textbase understanding and the other tapped situation-level understanding. Neither comprehension question tested readers about information from the critical duration sentences. Table 1 provides example passages and comprehension questions.

Procedure.   All aspects of the experiment were controlled using the program SuperLab (for the PC). Passages were presented one sentence at a time via computer using a noncumulative self-paced reading procedure. All text was displayed in black on a white background using a 15-in. monitor. Text was presented in Times New Roman font and a 12-point size. Passage sentences were centered vertically and left-aligned.

Participants were verbally instructed to read the passages for comprehension such that they could answer subsequent questions about the passages but not to dwell on any particular part of the text. Participants were informed that there were three keys that they would use when completing the experiment. The spacebar was used to move forward in the experiment (i.e., read each sentence of the passages), and the T key and F keys were used to answer true and false questions. These instructions were also presented on-screen. The SuperLab program recorded the time between all key presses and all key types that were pressed.

Results

The primary dependent variable was mean reading time per sentence for target sentences. Reading times for sentences immediately following the target sentences (Target + 1) were also measured to evaluate the possibility of spillover effects. Using upper and lower cutoffs of +2 standard deviations (within condition) for the critical sentences resulted in omitting 9.1% of the target sentence data and 6.2% of the spillover sentence data. One participant did not have data in one condition after outlier removal. The condition mean was inserted for that subject to maintain an

TABLE 3
Mean and Standard Error Reading Times in Milliseconds
for Target Sentences in Experiment 1

	Consistent		Short Inconsistent		Long Inconsistent	
Sentence Type	M	SE	M	SE	M	SE
Target	2,030	54	2,418	59	2,206	71
Target + 1 (spillover)	2,122	64	2,136	55	2,167	62

equal number of participants per condition. Reading times were analyzed by subjects (F_1) and by items (F_2).

Accuracy on comprehension questions was measured to ensure that participants attended to the materials. No participant's accuracy score was at or below 70%. Consequently, all participants were included.

To test the hypothesis that reading inconsistent duration sentences would increase reading times relative to congruent duration sentences, one-factor within-subjects analyses of variance (ANOVAs) were performed to compare mean reading times for the consistent, short inconsistent, and long inconsistent target sentences. Target sentence means are presented in Table 3.

An initial ANOVA indicated significant main effects by participant, $F_1(1, 77) = 9.96$, $MSE = 2,944,356$, $p < .001$, and by items, $F_2(2, 22) = 5.04$, $MSE = 458,898$, $p = .016$. Short inconsistent duration sentences were read 388 ms slower than consistent duration sentences, and this difference was significant by participants, $t_1(77) = 7.10$, $p < .001$, and by items, $t_2(11) = 3.96$, $p < .05$. Long inconsistent duration sentences were read 176 ms slower than consistent duration sentences. This difference was significant by participants, $t_1(77) = 2.53$, $p < .006$, but not by items, $t_2(11) = 1.61$, $p > .05$. There was also a significant difference in reading times for long and short inconsistent duration sentences by participant, $t_1(77) = 3.13$, $p < .001$ (two-tailed) but not by item, $t_2(11) = 1.36$, $p > .05$ (two-tailed), but in the opposite direction predicted. Specifically, reading times were 212 ms longer for the short inconsistent duration sentences than for the long inconsistent duration sentences.

To evaluate the possibility of spillover effects, mean reading times on target + 1 sentences were analyzed. ANOVAs did not indicate any significant differences in reading times, $F_1(1, 77) < 1$; $F_2(2, 22) < 1$, across conditions, indicating no spillover.

Discussion

As predicted, readers had difficulty reading text that introduced duration inconsistencies. This provides evidence that readers monitor the duration of events in narrative text. Increased reading times for sentences with duration inconsistencies did

not spillover to the target + 1 sentences. Interestingly, reading times for long incon-sistent sentences were less than short inconsistent sentences. One possible reason for this outcome is that the effects of duration in Experiment 1 are manifesting themselves in goal achievement. This is possible because changing the duration of activities described in the text may violate the expected completion of the activity. For example, in one passage two individuals were described as getting together at a restaurant to catch up on their lives. In the short inconsistent condition, the individ-uals spend 5 min in the restaurant, in the long inconsistent condition they spend 7 hr, and in the consistent condition they spend 1 hr. The short inconsistent condition might cause a goal-comprehension problem because 5 min is not likely to com-plete the goal of eating in a restaurant and discussing life events. This is not the case for the 7-hr scenario. It would be unusual to spend so much time in a restau-rant, but the goal of eating and talking would still be achieved. If the temporal in-consistency is interpreted in terms of the goal, then the short inconsistent condition should produce the greatest violation and, consequently, the greatest increase in reading time. Thus, the short inconsistent condition can be seen as a temporal problem that produces a goal-achievement problem, whereas in the long inconsis-tent condition the temporal problem does not produce a goal-achievement prob-lem. This possibility is explored in Experiment 2.

Another possible explanation for the obtained pattern of data is the dependence on socially constructed units of time, such as seconds, minutes, or hours. Recall that duration inconsistency was built by varying exact time terms in the target sen-tences. The use of specific time units may not be the most direct manner in which to examine duration. The literature on time understanding suggests that social con-structed time units (e.g., seconds, dates, and so forth) are the last sense of time to develop in young adults and are even difficult for some adults to use (Friedman, 1990). This does not suggest that duration cannot be monitored using such units but that the introduction of specific units might further complicate duration under-standing. Moving away from using specific social-time terms might provide a clearer picture of the effects of duration. Experiment 2 was designed to manipulate duration without the use of specific socially constructed time terms.

EXPERIMENT 2

Experiment 2 was designed to evaluate whether (a) readers evaluate temporal in-formation when time terms are not explicitly stated and (b) whether temporal in-formation interacts with goal achievement. To do this, potential duration-related problems were built into passages by implicitly contrasting the duration of two events in a particular context. Experimental passages contained four sentences of interest: (a) a base event that roughly anchored a time frame, (b) a comparison event that was either congruent with the base event (similar duration) or not con-

gruent (different duration) given the context of the passage, (c) an intervening sentence (intended to capture any spillover effects that extended beyond the immediate processing of the comparison sentence and to separate potential effects from subsequent outcome sentences), and (d) an outcome sentence that provided information regarding the achievement of the goal as well as a context for further evaluation of the base and comparison events.

Because Experiment 2 was designed to evaluate whether readers monitor the duration of events when a text does not include specific time terms, it required a more complicated manipulation of duration than Experiment 1. Consider the following two examples (taken from the Appendix, passages 3 and 2, respectively). In passage 3, the base event is waiting for a TV show to finish so that an upcoming show may be watched, the comparison event is painting all the bathroom walls, and the outcome event is missing a TV show. A TV show typically lasts between 30 min and 1 hr. The comparison event (painting the walls) takes considerably more time than the base event. The base-event duration and comparison-event duration do not match; therefore, they are labeled incongruent. This duration incongruency leads to duration-related comprehension difficulty because it would not be possible for the protagonist to watch his TV show if he engaged in the comparison activity. In this example, the comparison sentence is expected to be the source of the duration-related comprehension problem. The outcome sentence indicates that the show was missed. This outcome is consistent with the duration-comprehension problem because the protagonist misses his TV show. Therefore, it will not be a source of comprehension difficulty (i.e., reading time should not increase in this particular example for the outcome sentence compared to other conditions).

In passage 2, the comparison event is taking out the bathroom trash. The comparison event takes considerably less time than the base event. The base event duration and comparison event duration do not match; therefore, they are labeled incongruent. This incongruency is not expected to lead to a large duration-related comprehension problem because taking out the garbage is not likely to exceed the duration of the base event. Although a large duration-related comprehension problem is not expected for this sentence, the outcome sentence is expected to cause a duration-related comprehension problem. Readers are expected to have difficulty comprehending the fact that the protagonist missed the show given the short duration associated with the comparison event.

If readers represent the duration of described events, then reading times of comparison sentences and outcome sentences that establish duration-related comprehension problems should increase relative to comparison events and outcome sentences that do not create duration-related comprehension problems. The critical point is that comprehension difficulties that occur when reading the comparison sentence or outcome sentence both result from duration-related comprehension problems.

Presenting events in the manner described here created particular challenges. One challenge was to control the effect that goal achievement (i.e., outcome) might have on reading times. Introducing a goal that was not achieved was expected to increase reading times; therefore, goal achievement was included as a factor in this experiment. Half of the experimental passages contained goals that were achieved (e.g., protagonist watches the TV show) and half did not (e.g., protagonist misses the TV show).

A second challenge that needed to be considered was the plausibility of events. Events in the experimental passages might vary in terms of their plausibility (i.e., be more or less likely) given the context of the passage. Reading times for more plausible events might be shorter than less plausible events, irrespective of duration of the events described (Speer & Clifton, 1998). To control for potential plausibility effects, comparison-event sentences were ranked for plausibility by a separate group of participants, and this information was used to help construct stimuli. The plausibility ranking was similar in format to the ranking procedure used in Experiment 1. Twenty-two participants were provided with the context of the passage and decided the appropriateness of each of the possible comparison sentences. The order of comparison sentences in the ranking guide was randomized to control for potential order effects. As expected, the congruent condition received the lowest mean rank of 1.3, indicating these events were most plausible. Mean ranks for the short incongruent sentences and long incongruent conditions were 2.2 and 2.5, respectively.

Primary Hypotheses

Each experimental passage contained four target sentences (indicated by italics in the Appendix), and reading times on each sentence were collected. If reading times increase at the point of expected duration-related comprehension problems, this indicates that the duration of events were represented and evaluated even though no specific duration terms (i.e., seconds, minutes, or hours) were used in the passages. Predictions regarding the pattern of results for each target sentence are provided next. The predictions refer to the sample passages in the Appendix.

Hypotheses Set 1: Comparison sentences. Temporally incongruent long conditions (Appendix, passages 3 and 6) were expected to produce longer reading times for comparison sentences relative to short incongruent and congruent sentences. This was predicted because the comparison activity in these conditions overshoots the expected duration frame, which creates a duration-comprehension problem. That is, the comparison activity (painting all of the walls) takes considerably more time than the duration of the base event (waiting for the TV show).

Temporally incongruent short conditions (e.g., passages 2 and 5) were expected to produce the next longest reading times for comparison sentences because the duration of the comparison event does not match the base event. That is, emptying the garbage does not "fill" the time necessary to wait for the TV show. The short incongruent conditions were not expected to produce the same level of duration-comprehension problem as the long incongruent condition, because it is still possible for the protagonist to watch the show (e.g., the protagonist could have engaged in some other activity that wasn't specified in the text).

Congruent conditions (e.g., passages 1 and 4) were not expected to produce comprehension difficulties and thus were expected to elicit shorter reading times than long and short incongruent conditions.

Hypothesis Set 2: Intervening sentences.   Intervening sentences were included to separate potential effects of the comparison and outcome sentences and to capture any "spillover" resulting from the comparison sentence. Spillover effects were only predicted when comparison-event sentences cause duration-comprehension problems. Temporally incongruent long conditions (e.g., passages 3 and 6) were candidates for spillover. Reading times for these conditions were expected to be longer than the other conditions, which were not expected to differ.

Hypothesis Set 3: Outcome sentences.   A main effect of goal achievement was predicted (not achieving a goal would lead to increased reading time for outcome sentences). An interaction was also predicted between congruency and goal achievement. Specifically, reading times for not-achieved outcome sentences in the short incongruent condition (passage 5) were expected to be greater than outcome sentences in the long incongruent condition (passage 6). The opposite was predicted for goal-achieved outcome sentences; reading times for goal-achieved outcome sentences in the long incongruent condition (passage 3) should be greater than for short incongruent outcome sentences (passage 2). This interaction was expected because the outcome events in these conditions create an unnatural duration situation with respect to the expected duration frame and comparison event. In essence, the reader might wonder why one would miss a show because of taking out the garbage. How would someone catch the show given the long period of time spent painting the bathroom? Resolving these questions should increase processing time.

Method

Participants.   One hundred twenty-six students from the University of Illinois at Chicago participated to fulfill an introductory psychology course requirement.

Design and materials. The design of the experiment was a 3 (duration congruency: congruent, incongruent short, incongruent long) × 2 (goal achievement: goal achieved, goal not achieved) factorial. See the Appendix for sample passages.

The primary dependent measure was reading time for target sentences. The second dependent measure, accuracy on comprehension questions, was evaluated only to ensure that participants attended to the passages. Questions were not asked about information related to the base or comparison sentences in the passages. Two comprehension questions were administered after reading each passage. One comprehension question was designed to tap surface or textbase understanding and the other tapped situation-level understanding.

Participants read a total of 32 passages (20 filler and 12 experimental). Filler passages did not include time inconsistencies and were included to draw attention away from the duration-related problems. Of the 12 experimental passages, 2 passages were presented in each of the six conditions. Only 8 of the 12 experimental passages included some form of a duration-related manipulation. Overall, only 25% of the passages contained a duration incongruency. Comparison sentences were constructed so that sentence lengths would be similar across the congruent, short, and long conditions. Similarly, outcome sentences in which the goal was achieved were matched for length with outcome sentences in which the goal was not achieved. All other critical sentences (intervening and wrap-up) were identical across versions of each passage.

Procedure. The procedure, apparatus, and instructions were identical to those used in Experiment 1. The only change was in the stimuli used.

Results

Mean reading times were analyzed for comparison sentences, intervening sentences, and outcome sentences. Reading times were analyzed by subjects (F_1) and by items (F_2). The same outlier removal procedure was adopted from Experiment 1 with one exception. Extreme outliers (i.e., sentence reading times greater than 20 s) were first removed to avoid artificial inflation of outlier cutoffs, thereby reducing the number of data points labeled outliers. Extreme reading times are well outside the range of normal reading times for the critical sentences. Removing extreme outliers and using the ± 2 standard deviations cutoffs within condition resulted in omitting 8.6% of the comparison-sentence data, 3.8% of the intervening-sentence data, and 8.3% of the outcome-sentence data.

Accuracy on comprehension questions was measured to ensure that participants attended to the materials. No participant's accuracy score was at or below 70%. Consequently, all participants were included in analyses.

To test the prediction that reading incongruent comparison sentences would increase reading times relative to congruent comparison sentences, one-factor within-subjects ANOVAs were performed to compare mean reading times for the congruent, short, and long comparison sentences. The ANOVAs for comparison sentences indicated significant main effects by subjects, $F_1(1, 125) = 12.47$, $MSE = 3{,}783{,}116$, $p < .05$, and items, $F_2(2, 22) = 3.89$, $MSE = 320{,}004$, $p = .036$. Table 4 shows the mean reading times. Mean reading times for each condition were compared using one-tailed t tests because specific directional predictions were made a priori. As expected, long incongruent comparison sentences were read significantly slower (347 ms slower) than consistent duration sentences, $t_1(125) = 5.37$, $p < .001$. Short incongruent comparison sentences were also read slower than consistent duration sentences (140 ms slower), $t_1(125) = 2.04$, $p < .05$. Long incongruent comparison sentences were read 207 ms slower than short incongruent sentences, $t_1(125) = 2.70$, $p < .05$.

To evaluate the possibility of spillover effects, mean reading times on intervening (comparison + 1) sentences were analyzed. One-factor within-subjects ANOVAs indicated significant differences by subject but not by item, $F_1(1, 125) = 4.06$, $MSE = 289{,}624$, $p < .05$; $F_2(2, 22) > 1$. Intervening sentences following long incongruent comparison sentences were read significantly slower (88 ms slower) than consistent duration sentences, $t_1(125) = 2.72$, $p < .05$. Intervening sentences following short incongruent comparison sentences were also read slower than those following consistent duration sentences (76 ms slower), $t_1(125) = 2.12$, $p < .05$. No difference was detected between reading times for intervening sentences following long incongruent and short incongruent comparison sentences, $t_1(125) = .30$, $p > .7$. These analyses indicate that the duration manipulation spilled over to the sentence immediately following the comparison sentence; providing stronger evidence that the duration manipulation had an effect. The analyses could also be interpreted to suggest that readers are representing the duration of events after they have been initially processed.

TABLE 4
Mean and Standard Error Reading Times in Milliseconds
for Sentences in Experiment 2

Sentence Type	Congruent		Short Incongruent		Long Incongruent	
	M	SE	M	SE	M	SE
Comparison	3,120	63	3,260	64	3,467	76
Intervening	1,948	36	2,024	43	2,036	39
Outcome						
Goal achieved	2,550	59	2,584	59	2,530	54
Not achieved	2,558	55	2,666	51	2,656	59

Note.   Goal status cannot be defined prior to reading the outcome sentence.

According to Hypotheses Set 3, there should be a main effect of goal achievement and an interaction between goal achievement and duration congruency. To evaluate these predictions, reading times for goal-achievement sentences were compared using 3 (congruency) × 2 (goal achievement) within-subjects ANOVAs. The ANOVAs indicated a significant main effect of goal achievement by subject, $F_1(1, 124) = 6.54$, $MSE = 1,220,674$, $p < .05$, but not significant by item, $F_2(1, 11) = 3.41$, $MSE = 222,447$, $p = .092$. Overall, reading time for goal achieved outcome sentences (2,555 ms) was 82 ms faster than goal not achieved sentences (2,637 ms). There was no main effect of the duration manipulation on outcome sentences, $F_1(2, 123) < 1$; $F_2(2, 22) < 1$, and no evidence of an interaction between congruency and goal achievement, $F_1(2, 248) < 1$; $F_2(2, 22) < 1$.

Discussion

The primary hypothesis set was confirmed: Readers had more difficulty reading temporally incongruent comparison sentences than temporally congruent sentences. These results provide direct evidence that readers monitor duration because comprehension problems could only occur if the durations of the events described were evaluated and represented in memory. Situation models are often described as containing information that extends beyond the textbase (i.e., information added from knowledge that was not included in the text proper). In the incongruent comparison sentences, readers must have accessed their own knowledge of duration or inferred the durations of events because this information was not provided in the textbase. This is consistent with a situation-model level of analysis. The effects of duration incongruencies also manifested themselves in sentences that follow the incongruency (the intervening and goal sentences). Consequently, duration information must have been retained after reading the comparison sentence.

Also as predicted, long incongruent comparison sentences caused more comprehension difficulty than the short incongruent sentences. This prediction was based on the expected level of the comprehension disruption. The long incongruent condition clearly violates the time frame set by the base event (i.e., it overshoots the time associated with the base event) and creates a duration-related comprehension problem. The short incongruent condition also violates the time frame set by the base event (i.e., it undershoots the time associated with the base event), but because the short-duration events may be completed within the time frame established by the base event, less disruption was expected than for long-duration incongruent events.

Analyses also indicated a significant main effect of goal achievement: Readers took longer to read outcome sentences in which the goal was not achieved than when it was achieved. This was expected and is consistent with intentionality research, which provides evidence that readers spend additional cognitive effort attempting to integrate failed goals (Lutz & Radvansky, 1997; Suh & Trabasso,

1993; Trabasso & Suh, 1993). The effects of goal achievement were also localized to the sentence in which the goal outcome was stated. There was no evidence of spillover in wrap-up sentences.

There was also no evidence for the predicted interaction of congruency and goal achievement. A significant interaction would suggest that duration information influenced subsequent goal processing. Thus, the status of duration information after reading the comparison sentence remains an open question. This experiment was not designed to measure the contents of the final situation model after reading was completed, but this would be a logical next step in the exploration of duration representation.

In summary, Experiment 2 demonstrated that implicit duration incongruencies cause reading difficulties (as measured by increased reading times). The results provide strong evidence that the duration of events was monitored and evaluated in comparison sentences but no evidence that the duration information was retained or used to evaluate nonduration information, such as intentionality in the outcome sentences. The findings also indicate that the methodology used was sensitive enough to measure comprehension disruptions at both temporal and intentional (i.e., goal) levels.

EXPERIMENT 3

Readers could have approached duration-related problems in a variety of ways. For example, readers might have noticed a duration-related comprehension problem (e.g., that painting all of the bathroom walls would take more time than waiting for a TV show to start) but dismissed it as an unintended mistake. Or, readers might have noticed a problem and made some sort of repair or elaborative inference to explain it. Consider the TV/painting example again. Although the protagonist is waiting for a TV show to start, he paints all of the walls in his bathroom. He then comes back to watch the TV show. The reader might have inferred that very little painting was needed because the bathroom was small, or that the protagonist was finishing up a job he started previously. Elaborative inferences would help the reader understand how the protagonist was able to watch the TV show.

Another alternative is that readers might not notice duration-related comprehension problems. Rinck et al. (2001, Experiment 4) had participants read a passage containing a temporal inconsistency and then report whether an inconsistency was present. Only half of the participants explicitly reported the inconsistency. Importantly, reading time increased when participants encountered the inconsistency even if readers did not report the inconsistency. In a follow-up eye-tracking study, Rinck et al. (2003) reported only 36% of participants noticed the inconsistency. Participants who noticed the time problem were more likely to make regressions to the critical sentences and make longer (second-pass) fixations on that sentence.

The purpose of Experiment 3 was to examine whether readers were aware of the implicit duration-related problems and to identify strategies used by readers to process duration-related information. In Experiment 3, participants produced written think-aloud protocols while they read passages identical to those used in Experiment 2. To collect the written protocols, participants typed their thoughts after reading each sentence from a passage. Written responses to the critical sentences (i.e., comparison sentences, intervening sentences, outcome sentences, and wrap-up sentences) were then analyzed. Although verbal protocols have been used more frequently than written protocols, written protocols provide an effective way for participants to report their thinking processes during a task (Schooler, Ohlsson, & Brooks, 1992). Lutz and Radvansky (1997) provided specific evidence that written protocols replicate previous results obtained using verbal protocols.

Predictions About Awareness Levels

Participants were not expected to notice every duration-related inconsistency. This expectation is based on Zabrucky's (1990) finding that students frequently fail to adequately monitor their understanding of texts and on Rinck et al.'s research (2001, 2003, Experiments 4 and 2, respectively). Using Zabrucky and Rinck et al. (2003) as a guide, we expected roughly a third of the participants in Experiment 3 to notice the duration-related problems specified in Experiment 2 (i.e., from comparison sentences in passages 2, 3, 5, and 6 in the Appendix, all incongruent conditions). The decision to examine outcome sentences was based on the hypothesized interaction between the duration manipulation and goal achievement. There was no evidence for this interaction in Experiment 2, but the outcome sentences were analyzed to further explore readers' thought processes. We also expected that sentences that contained duration-related problems would be more likely to elicit protocols that referenced a duration-related comprehension problem than sentences that did not contain duration-related problems.

Awareness was assessed by coding participants' protocols for statements that did or did not reference duration-related comprehension problems. Protocols were scored as indicating awareness if participants reported a duration problem. This report could take many forms. For example, some participants explicitly mentioned time terms when reporting duration-related problems (e.g., "4 CDs mean about 3 hours of music. So she was out a long time"). Other participants made statements or posed questions that suggested there was a duration problem in the text (e.g., "Sounds like a really short day. What happened to tanning?"). Some participants made an elaboration or inference in an attempt to explain the duration problem (e.g., "Wow clothes must be that dirty. I wonder if he knows how to wash right?"). All of these examples indicate some awareness of the duration problem.

Each response was assigned a score of 1 if an actual duration problem or an incorrect inclusion (i.e., the participant reported a duration problem that did not ex-

ist) was reported or a 0 if a duration problem was not reported. Two raters independently evaluated the protocols for all critical sentences. The agreement level between raters' coding was 87% ($\kappa = .70$, indicating substantial agreement). Disagreements were settled by discussion between the raters (as in Cote, Goldman, & Saul, 1998).

Predictions About Specific Strategies Used

There were several strategies that readers could have used when encountering duration-related comprehension problems. Cote et al. (1998), Cote, Goldman, and Saul (1999), Magliano, Trabasso, and Graesser (1999), and Trabasso and Magliano (1996) all provided categorization guides for the types of responses gathered from think-aloud protocols. For the purposes of Experiment 3, protocols were categorized based on the type of reasoning demonstrated by participants using the prior research as a guide. Three general response categories were used. Protocol responses were categorized as (a) *reflective statements*: metacognitive, monitoring, or questioning statements (e.g., "Wow … Just how I predicted it"), (b) *generative statements*: explaining, elaborating, or inferencing statements (e.g., "Maybe Tyler was sleepy because he had no food in his stomach"), or (c) *other statements*: simple paraphrase statements, nonintegrative association statements, or unrelated statements (e.g., "I love lemonade").

For individuals who noticed the duration-related problems, protocol information was expected to aid in describing the types of processes used to understand duration-related problems in the texts. Critical sentences that included a duration-related comprehension problem were expected to elicit more reflective and generative statements than critical sentences that did not include potential comprehension problems. Two independent raters categorized all critical-sentence protocols into the three response categories; agreement level between raters' coding was 81% ($\kappa = .71$).

Method

Participants.   Sixty students from the University of Illinois at Chicago participated to fulfill an introductory psychology course requirement.

Design and materials.   The materials used in Experiment 3 were identical to those of Experiment 2. The design was a 3 (duration congruency: congruent, incongruent short, incongruent long) × 2 (goal achievement: goal achieved, goal not achieved) factorial. The dependent measures for Experiment 3 included the number of think-aloud statements related to the acknowledgment of duration-related comprehension problems and the form that these statements took (e.g., reflective, generative, or other statements).

Each participant read a practice passage and then six experimental passages, one from each condition, in a random order. The practice passage (a filler passage from Experiment 2) was included to familiarize participants with the think-aloud procedure. According to Ericsson and Simon (1993), only one practice passage is necessary because thinking aloud is a natural process and does not require extensive training. The experimental passages presented were counterbalanced across participants to ensure that all passages were presented in each condition an equal number of times. Note that Experiment 3 did not include filler passages and used a smaller set of experimental passages than Experiment 2. Collecting think-aloud data for a large number of passages was expected to fatigue participants and compromise the quality of data collected. Also, if a duration problem was noticed for several passages, we felt this might artificially focus attention on the duration of activities. There were three observations in each condition across the range of all passages.

Procedure.   Passages were presented one sentence at a time via computer using a noncumulative self-paced reading procedure. All text was displayed in black on a white background using a 15-in. monitor. Text was presented in Times New Roman font in a 12-point type size. Sentences were centered vertically but not horizontally. Horizontal starting position was the same for each sentence (left-aligned).

Participants were verbally instructed to read the passages for comprehension and to report their thoughts as they read through the text. The think-aloud instructions used were adopted from Schooler et al. (1992) and Trabasso and Magliano (1996). The experimenter was present throughout the entire experiment to ensure that participants understood the procedure and were on task. Participants' typed responses were recorded for every sentence they read. Reading times could not be recorded with the software used.

Results

Experiment 3 was a two-factor (Congruency and Goal Achievement) within-subjects design. The dependent variable was the number of protocols that fit within each awareness and strategy category. Duration-awareness levels were converted into probability scores (i.e., the probability of noticing the duration problem) for analysis of all target sentences.

To gauge awareness levels for comparison sentences, a one-factor within-subjects ANOVA was performed to compare the awareness-probability means for the congruent, short, and long comparison sentences. Table 5 presents participants' awareness levels for all sentence types. The ANOVA indicated a significant main effect of congruency, $F_1(2, 118) = 8.84$, $MSE = .651$, $p < .001$. Pairwise comparisons indicated that the congruent condition was significantly different from the

TABLE 5
Mean Probability and Standard Error of Reporting Duration Incongruency
for Sentences in Experiment 3

Sentence Type	Congruent		Short Incongruent		Long Incongruent	
	M	SE	M	SE	M	SE
Comparison	.12	.03	.22	.04	.32	.04
Intervening	.03	.02	.05	.02	.04	.02
Outcome						
Goal achieved	.07	.03	.10	.04	.10	.04
Not achieved	.20	.05	.25	.06	.23	.06
Wrap-up						
Goal achieved	.02	.02	.03	.02	.03	.02
Not achieved	.02	.02	.08	.04	.15	.05

Note.   Goal status cannot be defined prior to reading the outcome sentence.

short and long incongruent conditions ($p = .03$ and $p = .001$, respectively). No difference was found between the short and long conditions ($p > .08$). This pattern of awareness mirrors the pattern of reading-time data observed in Experiment 2. That is, conditions that elicited longer reading times were also more likely to elicit duration awareness.

For intervening sentences, the one-factor within-subjects ANOVA indicated that detection rates for each condition were not significantly different, $F_1(2, 118) < 1$. The fact that there were so few awareness responses made after reading the intervening sentences provides further evidence that the effects of the duration manipulation did not spillover to the sentence immediately following the comparison sentence. Processing difficulties were localized to the comparison sentences.

To gauge awareness levels for outcome sentences, awareness-probability means were compared using a 3 (congruency) × 2 (goal achievement) within-subjects ANOVA. There was a significant main effect of goal achievement, $F_1(1, 59) = 13.79$, $MSE = 1.74$, $p < .01$. Overall, the probability of noticing duration incongruencies for goal-not-achieved outcome sentences was .14 greater than for goal-achieved sentences. There was no main effect of congruency on awareness, $F_1(2, 118) < 1$. Mean probabilities were .14 for outcome sentences in the congruent condition, .18 for outcome sentences in the short incongruent condition, and .17 for outcome sentences in the long incongruent condition. There was no interaction between congruency and goal achievement, $F_1(2, 118) = 2.36$, $p = .10$, *ns*.

The outcome + 1 sentences (wrap-up sentences) were also analyzed using a 3 (congruency) × 2 (goal achievement) within-subjects ANOVA. There was a significant main effect of congruency, $F_1(2, 118) = 3.16$, $p < .05$. Mean probabilities of noticing duration incongruency were .02, .06, and .09 for the congruent, short incongruent, and long incongruent comparison sentences, respectively (only the difference between the long incongruent condition and the congruent condition

reached statistical significance, $p = .01$). This overall pattern of awareness mirrors the results obtained in the comparison-sentence analyses, albeit at a greatly reduced level. There was also a significant main effect of goal achievement, $F_1(1, 59) = 6.86$, $p < .05$. Participants were more likely to notice a problem when goals were not achieved. No evidence was obtained for an interaction of congruency by goal achievement, $F_1(2, 118) = 2.37$, $p = .1$, *ns.*

The second set of hypotheses for Experiment 3 contained predictions regarding how readers might attempt to deal with duration-related problems. Sentences containing a duration problem were expected to elicit more reflective and generative think-aloud comments from participants than sentences that did not contain a duration problem. To evaluate this prediction, the frequency of reflective, generative, and other comments were examined for the critical comparison sentences.

Table 6 provides the total number of reflective, generative, and other strategies for congruent, short incongruent, and long incongruent sentences. The distribution of strategies across different target sentence types could not be evaluated using an ANOVA because detection and nondetection of duration problems cannot be directly compared across congruency conditions. Specifically, because there are no duration problems in congruent passages, participants could only incorrectly report (incorrect inclusion) or correctly not report (correct omission) duration problems. In contrast, in the two incongruent conditions there are duration problems in every passage; therefore, participants only could incorrectly not report (incorrect omission) or correctly report (correct inclusion) duration problems. Nonetheless, the distribution of strategies for comparison and outcome sentences provides useful information.

TABLE 6
Frequency of Protocol Strategies Used in Experiment 3 for Comparison
Sentences That Report a Duration Incongruency (Correct Inclusion,
Incorrect Inclusion) or Do Not (Correct Omission, Incorrect Omission)

	Reflective	*Generative*	*Other*
Congruent			
Incorrect inclusion	5	4	5
Correct omission	33	30	42
Total	38	34	47
Short incongruent			
Correct inclusion	10	8	8
Incorrect omission	31	23	40
Total	41	31	48
Long incongruent			
Correct inclusion	10	16	14
Incorrect omission	19	27	34
Total	29	43	48
Total	108	108	143

Three outcomes are worth emphasizing. First, the total number of reflective and generative comments elicited was the same for congruent, short incongruent, and long incongruent sentences (72 for each condition). Second, the total number of adaptive strategies (sum of reflective and generative strategies across conditions = 216) exceeded the number of nonadaptive strategies (sum of other strategies across conditions = 143). One simple interpretation of this outcome is that readers tried to make sense of the passages, and adaptive comments reflect this process. Third, converting strategy counts into proportions (recall that reporting opportunities are unequal between noticing and not noticing duration problems) revealed that a greater proportion of adaptive strategies were reported when duration problems were noticed (44/66) than when they went unnoticed (100/174), .67 and .57, respectively.

Discussion

Experiment 3 provided an evaluation of participants' duration-awareness levels using a think-aloud procedure. As expected, readers reported having some awareness of duration incongruencies. The percentage of individuals that noticed duration incongruencies was 27%, slightly less than the 36% originally reported by Rinck et al. (2003). This is not surprising considering differences in the materials used. The duration incongruencies used in Rinck et al.'s experiment explicitly mentioned time terms, which highlights the temporal inconsistencies. The 27% awareness level reported in our Experiment 3 can be attributed to the implicit nature of the duration manipulations.

The awareness results from Experiment 3 suggest the need to explore what is driving the reading-time results obtained in Experiment 2. Specifically, are the reading-time differences from Experiment 2 produced by only 27% of the participants? To examine this issue, reading times for congruent sentences were compared to the average reading time of incongruent sentences for each subject. To be complete, this analysis was also conducted for Experiment 1. Analyses indicated that more than two thirds of the participants displayed the expected pattern of results in each experiment (i.e., longer reading times for incongruent sentences). This indicates that reading time reflects differences not present in self-reports. The duration-awareness level obtained in Experiment 3 (27%) underestimates differences in reading behavior.

A main effect of goal was observed in the awareness levels for outcome sentences. Participants were more likely to report a duration problem if the goal was not achieved than if it was achieved. One interpretation of this result is that when the goal was not achieved, the additional decrement in coherence in the intentionality dimension highlighted awareness of duration problems encountered earlier in the text. Zwaan, Radvansky, Hilliard, and Curiel (1998) provided empirical evidence to support this interpretation. Zwaan et al. found that when coherence

breaks were introduced in more than one dimension, these breaks affected situation-change judgments more so than when only one dimension was manipulated.

Examination of the strategies used by participants tentatively suggested that participants were more likely to produce reflective or generative comments if they noticed the duration incongruency. Reflective and generative comments were considered to be adaptive and aid comprehension. This supports the conclusion that readers who noticed duration problems were attempting to integrate that information with their developing representations of the text.

Experiment 3 mirrors many of the important results obtained in Experiment 2 but provides an important result that did not occur in Experiment 2. Recall that in Experiment 2, there was no interaction between duration congruency and goal achievement. In Experiment 3, however, participants did make comments about duration problems when reading goal sentences, and the number of duration problems reported was larger when the goal was not achieved than when the goal was achieved. This provides additional evidence that readers represent duration and that this information was accessible after the duration problem was first encountered. The lack of interaction in Experiment 2 could be attributed to the nature of the experiment. The implicit duration manipulation might not have been strong enough to force individuals to reinterpret the goal in terms of duration; therefore, there was little or no increase in reading time. Experiment 3, which used a qualitative method, allowed readers to comment on potential problems even if the problems were not large enough to force a reinterpretation.

In summary, Experiment 3 demonstrated that readers report awareness of implicit duration incongruencies 27% of the time. The overall prediction that incongruency would lead to a larger proportion of reflective statements (i.e., metacognitive, monitoring, questioning) and generative statements (i.e., elaborative, explanatory, inferencing) was not supported. However, inspection of the data tentatively suggests that incongruency led to an increase in the use of reflective and generative statements for comparison sentences if the incongruency was noticed. A small proportion of readers also reported duration problems while reading outcome sentences, suggesting that the representation of duration extends beyond the comparison sentence.

GENERAL DISCUSSION

The research presented here was designed to accomplish two goals. First, we wanted to provide evidence that a specific and fundamental aspect of time (i.e., duration) was included in online situation-model construction. Second, we wanted to develop a methodology (i.e., Experiment 2) that could be used to test future duration-related hypotheses involving other dimensions of time, such as temporal order

and perspective. These senses of time are based on how duration is processed and represented; therefore, duration needs to be considered before exploring them.

The experiments provide evidence that readers routinely include duration information when building a situation model. Experiment 1 demonstrated that readers were sensitive to (i.e., had difficulty reading) inconsistent duration terms used to describe common events. Experiment 2 demonstrated that readers monitored the duration of events and used this information (i.e., compared event durations) even though duration terms were never explicitly mentioned. The results of Experiment 2 cannot be explained as a simple mismatch of duration information derived from the textbase because the sentences do not provide explicit duration information. Experiment 3 demonstrated that readers noticed implicit duration incongruencies 27% of the time and that these incongruencies were more likely to produce reflective or generative statements when noticed. Overall, the pattern of results supports the hypotheses that duration information is monitored by readers and may be an important element of the temporal dimension of the situation model.

The experimental approach used here has benefits over existing research. Nearly all of the studies cited as evidence that readers represent time rely on the use of time-shift manipulation statements such as *a moment later*, *an hour latter*, or *a week later* (Scott Rich & Taylor, 2000; Zwaan et al., 1998). These time shifts signal the need to construct a new situation model or substantially update the existing model and explicitly force the reader to evaluate time information. Consequently, these studies do not provide direct evidence that readers represent the temporal aspects of text when the construction of a new model is not required or cued. In contrast, these three experiments manipulated time within the development of a single situation model and provided evidence that readers monitor time (online) without the use of specific time-shift statements.

The experiments presented here extend our understanding of the role of time in text comprehension and situation-model development. Duration was identified as a fundamental aspect of time and, to our knowledge, was examined explicitly for the first time here. We have shown that readers noticed explicit and implicit inconsistent characterizations of event durations and that these inconsistencies caused comprehension problems. Identifying which aspects of time are important to reading comprehension will help bring us closer to understanding the complexities involved in reading.

REFERENCES

Albrecht, J. E., & O'Brien, E. J. (1995). Goal processing and the maintenance of global coherence. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading*. (pp. 263–278). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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, 35,* 427–440.

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

Block, R. A. (Ed.). (1990). *Cognitive models of time.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cote, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes, 25,* 1–53.

Cote, N., Goldman, S. R., & Saul, E. U. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In S. R. Goldman & H. Oostendorp (Eds.), *The construction of mental representations during reading* (pp. 169–194). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ericsson, K. A., & Simon, H. A., (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.

Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, textbases, and situational models: Recognition memory for three types of textual information. *Discourse Processes, 13,* 175–190.

Fraisse, P. (1963). *The psychology of time.* New York: Harper & Row.

Friedman, W. (1990). *About time.* Cambridge, MA: MIT Press.

Harnish, R. M. (1976). Reading comprehension and readability in educational practice and psychological theory. In T. G. Bever, J. J. Katz, & D. T. Langendoen (Eds.), *An integrated theory of linguistic ability* (pp. 313–391). New York: Crowell.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language, 29,* 133–159.

Labov, W. (1972). *Language in the inner city: Studies in the black English vernacular.* Philadelphia: University of Pennsylvania Press.

Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press.

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

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

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91,* 615–629.

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

Ornstein, R. (1972). *The psychology of consciousness.* San Francisco: Freeman.

Ornstein, R. (1997). *On the experience of time.* Boulder, CO: Westview.

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., Hahnel, 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.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1992). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General, 122,* 166–183.

Scott Rich, S., & Taylor, H. A. (2000). Not all narrative shifts function equally. *Journal of Memory and Cognition, 28,* 1257–1266.

Singer, M., & Richards, E. (2005). Representing complex narrative goal structures: Competing memory based and situational influences. *Discourse Processes, 39,* 189–204.

Speer, S. R., & Clifton, C. (1998). Plausibility and argument structure in sentence comprehension. *Memory & Cognition, 26,* 965–978.

Suh, S., & Trabasso, T. (1993). Inferencing 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. (1995). *Representing time in natural language: The dynamic interpretation of tense and aspect.* Cambridge, MA: MIT Press.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes, 21,* 255–287.

Trabasso, T., & Suh, S. Y. (1993). Understanding text: Achieving explanatory power coherence through on-line inferences and mental operations in working memory. *Discourse Processes, 16,* 3–34.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic.

Zabrucky, K. (1990). Evaluation of understanding in college students: Effects of text structure and reading proficiency. *Journal of Reading Research and Instruction, 29,* 46–54.

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

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: an event-indexing model. *Psychological Science, 6,* 292–297.

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

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998) Constructing multidimensional situation models during reading. *Scientific Studies of Reading, 2,* 199–220.

APPENDIX

The following are sample passages for each condition in Experiment 2 demonstrating the two conditions: Congruency and Goal Achievement. (Target sentences are italicized.)

(1)   Temporally Congruent (comparison activity is appropriate length) / Goal is Achieved

> Brian got up late on Saturday. It had been a hard week working in the office and he worked until 11:00 on Friday night. Today Brian had tentative plans to do some work around the house and watch one of his favorite TV shows. Brian turned on the TV and noticed that another whole show was playing before his favorite would be on. *Brian went into the bathroom and cleaned the sink and tub. He was pleased he was taking initiative. When Brian came back he sat down and enjoyed his show. Brian recalled that his friend Lisa was also a big fan of the show.* Brian decided he would call up Lisa and chat.

Comprehension Questions (true or false)

1. Brian got up late.
2. Brian wanted to talk to Lisa about the show.

(2)   Temporally Incongruent (comparison activity is too short) / Goal is Achieved

Brian got up late on Saturday. It had been a hard week working in the office and he worked until 11:00 on Friday night. Today, Brian had tentative plans to do some work around the house and watch one of his favorite TV shows. Brian turned on the TV and noticed that another whole show was playing before his favorite would be on. *Brian went into the bathroom and emptied the garbage can. He was pleased he was taking initiative. When Brian came back he sat down and enjoyed his show. Brian recalled that his friend Lisa was also a big fan of the show.* Brian decided he would call up Lisa and chat.

Comprehension Questions (true or false)

1. Brian got up late.
2. Brian wanted to talk to Lisa about the show.

(3)   Temporally Incongruent (comparison activity is too long)/ Goal is Achieved

Brian got up late on Saturday. It had been a hard week working in the office and he worked until 11:00 on Friday night. Today, Brian had tentative plans to do some work around the house and watch one of his favorite TV shows. Brian turned on the TV and noticed that another whole show was playing before his favorite would be on. *Brian went into the bathroom and painted all of the walls. He was pleased he was taking initiative. When Brian came back he sat down and enjoyed his show. Brian recalled that his friend Lisa was also a big fan of the show.* Brian decided he would call up Lisa and chat.

Comprehension Questions (true or false)

1. Brian got up late.
2. Brian wanted to talk to Lisa about the show.

(4)   Temporally Congruent (comparison activity is appropriate length)/ Goal is Not Achieved

Brian got up late on Saturday. It had been a hard week working in the office and he worked until 11:00 on Friday night. Today Brian had tentative plans to do some work around the house and watch one of his favorite TV shows. Brian turned on the TV and noticed that another whole show was playing before his favorite would be on. *Brian went into the bathroom and cleaned the sink and tub. He was pleased he was taking initiative. When Brian came back*

he discovered he had missed his show. Brian recalled that his friend Lisa was also a big fan of the show. Brian decided he would call up Lisa and chat.

Comprehension Questions (true or false)

1. Brian got up late.
2. Brian wanted to talk to Lisa about the show.

(5)   Temporally Incongruent (comparison activity is too short)/ Goal is Not Achieved

Brian got up late on Saturday. It had been a hard week working in the office and he worked until 11:00 on Friday night. Today, Brian had tentative plans to do some work around the house and watch one of his favorite TV shows. Brian turned on the TV and noticed that another whole show was playing before his favorite would be on. *Brian went into the bathroom and emptied the garbage can. He was pleased he was taking initiative. When Brian came back he discovered he had missed his show. Brian recalled that his friend Lisa was also a big fan of the show.* Brian decided he would call up Lisa and chat.

Comprehension Questions (true or false)

1. Brian got up late.
2. Brian wanted to talk to Lisa about the show.

(6)   Temporally Incongruent (comparison activity is too long)/ Goal is Not Achieved

Brian got up late on Saturday. It had been a hard week working in the office and he worked until 11:00 on Friday night. Today, Brian had tentative plans to do some work around the house and watch one of his favorite TV shows. Brian turned on the TV and noticed that another whole show was playing before his favorite would be on. *Brian went into the bathroom and painted all of the walls. He was pleased he was taking initiative. When Brian came back he discovered he had missed his show. Brian recalled that his friend Lisa was also a big fan of the show.* Brian decided he would call up Lisa and chat.

Comprehension Questions (true or false)

1. Brian got up late.
2. Brian wanted to talk to Lisa about the show.