Who when where:
An experimental test of the event-indexing model

MIKE RINCK and ULRIKE WEBER
Dresden University of Technology, Dresden, Germany

We tested the event-indexing model proposed by Zwaan, Langston, and Graesser (1995). Participants read narratives containing target sentences that involved situational shifts. Independently of each other, continuity and shifting of the protagonist, time, and location dimensions were varied. In Experiment 1, reading times of the target sentences increased for protagonist shifts and temporal shifts, whereas the effect of spatial shifts was weak. Moreover, an interaction of protagonist shifts and spatial shifts was found. These results were replicated in Experiment 2, which also revealed strong effects of these situational shifts on coherence ratings that participants gave immediately after reading each target sentence. Experiment 3 addressed the interaction of protagonist shifts and spatial shifts, showing that it may be due to the differential involvement of unexplained protagonist motions. These experimental results support the processing load predictions of the event-indexing model and extend previous correlational results.

During the comprehension of discourse—for instance, during reading of narrative texts, such as novels, short stories, or fairy tales—comprehenders build multilevel representations of the information conveyed (e.g., Bower & Rinck, 2001; Gernsbacher, 1990; Glenberg, Kruley, & Langston, 1994; Johnson-Laird, 1983; Kintsch, 1988, 1998; van Dijk & Kintsch, 1983; Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). The level of representation investigated here is the *situation model*, or mental model, of the text. Situation models constitute the level of text representation associated most closely with "deep" understanding and serve to integrate the information stated in a text with general information supplied by the reader's world knowledge. In short, situation models "represent what the text is about, not the text itself" (Glenberg, Meyer, & Lindem, 1987, p. 70). Zwaan and Radvansky have proposed a detailed account of how readers construct situation models during text comprehension. In their general processing framework, they distinguish between three types of situation models: the current, the integrated, and the complete models. They also distinguish four types of processes operating on these models: the construction, updating, and retrieval of situation models and the foregrounding of specific situation model elements. According to their framework, readers *construct* a *current model* of the situation described by a sentence or a clause. As they continue reading, each sentence yields a new current model. The information from all sentences read so far is being integrated into a single model, which is therefore called the *integrated model*. The process of incorporating a new current model into the integrated model is called *updating* of the model. During reading and updating, readers may focus more on some types of information than on others, a process called *foregrounding*. When all sentences have been read, the integrated model is stored in long-term memory as the *complete model*. Later, the complete model or elements of it may be *retrieved* from memory in an attempt to remember what has been read (see Zwaan & Radvansky, 1998, for details of these processes). The present article focuses on the updating of situation models and on the way in which different aspects of the situation are processed during the updating process.

According to all relevant theories of text comprehension, situation models are multidimensional—that is, many different aspects of a situation are represented in them. Among these aspects are causal, temporal, and spatial relations, as well as the protagonist, including his or her goals and emotions (see Zwaan, 1999; Zwaan & Radvansky, 1998). The most detailed theoretical account of the multidimensionality of situation models has been proposed by Zwaan and his colleagues (Zwaan, 1999; Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). They developed the *event-indexing model* to explain how readers construct coherent multidimensional representations of situations. According to the event-indexing model, events and intentional actions of characters are the focal points of situation models created from narratives. During comprehension, readers monitor several different aspects of the situation and update the current model according to (at least) five as-

pects: temporality, spatiality, protagonist(s), causality, and intentionality. This updating process is dynamic and occurs for every event described in the text. When the first event is understood, the event is indexed according to the time frame in which it occurs, the spatial region in which it occurs, the protagonist or protagonists it involves, its causal status with regard to prior events, and its relation to the protagonist's goals. For subsequent events, readers monitor whether they require updating of the index of any of these dimensions. For instance, the temporal index needs to be updated if a clause indicates a time shift when compared with the previous clause (as indicated, e.g., by such phrases as *a few minutes later* or *the next day*). If an event takes place in a different spatial region, the spatial index has to be updated. Similarly, the protagonist index needs to be updated when an event involves a different protagonist. The causal index needs to be updated when an event is causally unrelated to the previous event. Finally, the intentionality index must be updated when an action introduces a new goal structure.

The experiments reported here were designed to test the *processing load hypothesis* derived from the event-indexing model: Updating of the situation model should be more difficult if the current model differs from the integrated model on one or more situational dimensions. That is, the greater the number of indices (protagonist, time, space, causality, or intentions) that have to be updated during comprehension of a clause, the more time it will take to comprehend the clause. To keep the model parsimonious, Zwaan, Langston, and Graesser (1995) assumed that the different indices do not interact and are monitored independently of each other. Some empirical evidence for the processing load hypothesis is already available (for a review, see Zwaan & Radvansky, 1998). However, most of this evidence comes from studies in which shifts and breaks on single dimensions were investigated. For instance, Zwaan (1996) and Zwaan, Madden, and Whitten (2000) showed that time shifts in a narrative increase the reading time of sentences. Time shifts were signaled by the phrases *an hour later* and *a day later*, as compared with the phrase *a moment later*. Reading time for the sentences containing time shift signals were longer than those for sentences signaling temporal continuity.

As compared with the many studies on single dimensions, multidimensional studies of situation model updating are extremely rare. There have been very few experiments in which continuity on more than one index of the situation was manipulated to test the processing load hypothesis. The most comprehensive studies were reported by Zwaan and his colleagues. Zwaan, Magliano, and Graesser (1995) investigated three dimensions—namely, time, space, and causality. They found that breaks in temporal continuity and causal continuity caused reliable increases in reading times for sentences, whereas spatial shifts did not. Zwaan, Radvansky, Hilliard, and Curiel (1998) extended these findings by examining all five dimensions of the event-indexing model. Again, shifts on the temporal and causal dimensions increased

processing difficulty, as did shifts on the intentional and protagonist dimensions. In contrast, spatial shifts led to increases in reading time only when readers had acquired sufficient prior spatial knowledge by first memorizing a map of the building in which the events described in the narratives took place. These studies suggest that not all dimensions of situation models are of equal importance. In particular, the spatial dimension may be monitored less carefully than other dimensions. Therefore, the experiments reported below were designed to address the relative importance of different situational dimensions. Magliano, Miller, and Zwaan (2001) have applied the event-indexing model to the comprehension of narrative films, assessing how shifts in time, shifts in story region, and movements of protagonists affected how participants perceived filmed situations as continuous or changed. Although the study was correlational, the amount of data collected allowed for fully crossed $2 \times 2 \times 2$ analysis of shifts on the temporal, spatial, and protagonist dimensions. Interestingly, Magliano et al. found that some shift combinations occurred much more frequently than others in their natural materials, a finding that we will return to later.

In all of these studies, naturalistic texts or films were employed as experimental materials, and regression analyses were used to determine the effects of situational shifts. However, the use of naturalistic materials increases the ecological validity of the results at the expense of experimental control. Since the regression approach is correlational in nature, it does not allow one to draw the strong causal conclusions that an experimental approach offers. Moreover, regression approaches make it difficult to test potential interactions among situational dimensions and to compare the relative sizes of the effects of the dimensions. Finally, the extremely large amount of materials that Magliano et al. (2001) had to analyze in order to find a full crossing of three situational dimensions is another argument for adopting an experimental approach. Therefore, it would be useful to supplement the existing correlational studies with experiments in which different situational dimensions are varied independently of each other. The experiments reported here were designed to achieve just this.

We are aware of only one attempt to test the importance of several situational dimensions simultaneously and experimentally, reported by Scott Rich and Taylor (2000). In three experiments, they studied shifts on the protagonist, time, and space dimensions, using experimentally constructed narratives. Scott Rich and Taylor found that all three types of shifts affected reading times and probe reaction times, as well as the perceived cohesion of critical sentences and the perceived coherence of the narratives overall. Shifts in protagonist (introduction of a new protagonist not mentioned previously), place (move to a new location), and time (jump of narrative time across the typical duration of the scenario) increased the reading time for shift sentences, as compared with nonshift sentences, as well as decreasing the perceived

coherence and cohesion. Comparing the effects of the three dimensions, Scott Rich and Taylor concluded that the protagonist dimension is most important as an event index. Unfortunately, a number of problems make it very difficult to use the results of Scott Rich and Taylor in an evaluation of the processing load hypothesis of the event-indexing model. First, the authors did not use a full factorial combination of the three dimensions; thus, possible interactions cannot be tested. Second, the coherence ratings were collected at the end of each narrative. Each narrative, however, contained two different shifts. Therefore, it is impossible to determine the relative importance of each single shift on coherence. Third, there is no adequate baseline to compare the shift effects with, because the standard condition used by Scott Rich and Taylor contained another type of shift—namely, an activity shift. Finally and most serious, there was a confound of experimental conditions and materials: The materials used in each experimental condition (standard, protagonist shift, time shift, and location shift) were made up of completely different sentences without any counterbalancing.

In order to perform a strong experimental test of the event-indexing model, one would have to construct experimental narratives with at least two critical features. First, they should allow for a full combination of all the shifts investigated, comprising all combinations from *no shift at all* to *simultaneous shifts on all dimensions*. Second, the narratives should allow for a complete counterbalancing of materials and experimental conditions. Ideally, identical target sentences should be used in all of the different experimental conditions. With the experiments reported here, we attempted to achieve this experimental test of the event-indexing model. With regard to the full combination of possible shifts, we restricted ourselves to the same dimensions as those Scott Rich and Taylor (2000) studied: protagonist, time, and space. On each of these dimensions, critical target sentences could be either continuous or shifting. Full combination of these dimensions yielded a total of eight experimental conditions. Regarding counterbalancing, each narrative contained a single target sentence that was identical in all the experimental conditions. Instead of varying the target sentence, we varied continuity on the protagonist, time, and space dimensions by creating eight different versions of the episode preceding the target sentence. Keeping the target sentence constant in all the experimental conditions also served to reduce error variance and to increase the power of the statistical analyses. After extensive pilot testing, these experimental narratives were used to determine the effects of protagonist, temporal, and spatial shifts on on-line processing (as indicated by target sentence reading times), on perceived coherence (as indicated by coherence ratings of target sentences), and on perceived likelihood of alternative story continuations (as indicated by rankings of alternatives). Our goal was to answer the following questions. Do all types of shifts increase the processing load involved in sentence comprehension, in

comparison with fully continuous sentences? Are the effects of protagonist, temporal, and spatial shifts of comparable size? And are these effects additive, or do the dimensions interact?

EXPERIMENT 1

Experiment 1 was designed to study the effects of protagonist, temporal, and spatial shifts on on-line processing of sentences. The participants read narratives in a sentence-by-sentence self-paced manner, with sentence reading time used as the dependent variable.

Method

Participants. Forty-eight students at Dresden University of Technology, most of them psychology undergraduates, participated in the experiment, either to fulfill a curricular requirement or for compensation with a small monetary payment equivalent to $5. All the participants were native speakers of German.

Materials. We constructed eight experimental narratives, which followed the patterns shown in Appendix A and Appendix B. The first part, which was identical in all versions of a single narrative, introduced the two protagonists,[1] the time and place of the narrative, and the general setting. The introduction also explained why activities were begun by the two protagonists in separate locations. The introduction was followed by the experimental episode, which was varied to create continuity or shifts of the protagonist, the time of the actions, or the location of the actions. For instance, in the sample stories shown in Appendix A and Appendix B, the experimental episode was about the actions of either one or the other protagonist (Paul vs. Frieda), the actions occurred within one or the other temporal frame (morning to noon vs. noon to evening), and they occurred in one or the other location (house vs. garden). The subsequent target sentence was again identical. In this story, it read *In the last daylight, Paul stood in the garden and looked around satisfied.* This sentence contained a protagonist shift if the previous episode had been about Frieda rather than about Paul. It contained a temporal shift if the previous actions had lasted until noon rather than until the evening. And it contained a spatial shift if the previous actions had occurred in the house rather than out in the garden. These three shift types were created independently of each other, yielding eight different versions of the experimental episode, ranging from the fully continuous version shown in Appendix A to the fully discontinuous version in Appendix B. To avoid a confound of situational continuity and surface characteristics, such as word repetition, the last sentence of the episode never contained the critical words used in the target sentence (e.g., *Paul*, *last daylight*, or *garden*; see Appendix A). The target sentence was followed by the final part of the story and a comprehension question. In the final part, the concluding sentences served to wrap up the story line and to explain any shifts that the target sentence might have contained. The eight filler texts were of similar length, but they contained only one protagonist and no shifts. They were used to distract the participants' attention from the uniform structure of the experimental texts.

Pilot testing of the materials. In naturalistic discourse, it would be unusual to use exactly the same wording of the target sentence, regardless of whether it was continuous or discontinuous with the previous discourse. Instead, for instance, one would add such words as *meanwhile* versus *later* to signal temporal continuity versus temporal shift. Therefore, it is important to ensure that readers clearly understand the protagonist, time, and location of the actions described in the target sentences. Moreover, they should not find the stories and the target sentences strange, incoherent, or contradictory. The latter point was clarified in a first series of consecutive in-

formal pilot studies. As a result, only eight experimental stories remained. For these, the continuous versions, as well as all the shift versions, were judged to be well written, sensible, and at least moderately coherent by the participants in the pilot tests. All other stories were discarded or transformed into filler stories. In an additional, formal pilot test, we checked whether the protagonist, temporal, and spatial shifts in each target sentence version were nonambiguous. The participants were asked to indicate the current protagonist, time, and location of the actions described in the narrative, both immediately before and immediately after the target sentence in each story. The results of this test showed that for all but one experimental story, the current situational dimensions were nonambiguous, both for the target sentences and for the sentences preceding them. The single problematic target sentence was revised and successfully retested.

Procedure. A total of eight experimental texts and eight filler texts were presented to the participants. For practice, the first text was always a filler text, followed by experimental texts and filler texts in alternating order. The order of the experimental texts was counterbalanced across participants. The participants were asked to read the texts for comprehension at their normal speed. To ensure mindful reading, a comprehension question was asked at the end of each text. Two participants who made errors on more than 20% of these questions were dropped from the analyses and were replaced by other participants. The stories were read one sentence at a time on the screen of a computer, controlled by the RSVP software (Williams & Tarr, n.d.). Sentence reading times and responses to the comprehension questions were automatically recorded. Across participants, each text occurred equally often in each of the eight experimental conditions, to ensure full combination of conditions and materials. It took the participants about 45 min to read the passages and answer the questions. At the end of the experiment, the participants were informed about the purpose of the experiment.

Design. Full combination of the *protagonist continuity* (continuous, shift), *temporal continuity* (continuous, shift), and *spatial continuity* (continuous, shift) factors yielded a $2 \times 2 \times 2$ design. All the factors were varied within subjects, and each participant encountered one target sentence in each experimental condition. Reading times for the target sentences were recorded as the dependent variable.

Results and Discussion

The results of Experiment 1 are summarized in Table 1, which shows reading times per syllable of the target sentences. These reading times were computed by dividing each individual reading time by the number of syllables of the corresponding target sentence. This way, error variance due to variations in sentence length was reduced. All the analyses were computed twice, once using the 48 participants as a random factor and once using the eight experimental sentences. Below, F_1 values relate to

the by-participants analyses, whereas F_2 values relate to the by-materials analyses. The data in Experiment 2 were treated the same way. All the effect sizes reported are f values according to Cohen (1988).

Reading times. The $2 \times 2 \times 2$ analysis of variance (ANOVA) of the target sentence syllable reading times yielded significant main effects of protagonist shifts and temporal shifts. If the target sentence contained a new protagonist, mean syllable reading times increased from 209 to 232 msec [$F_1(1,47) = 8.12$, $p < .01$, $f = .15$; $F_2(1,7) = 10.33$, $p < .05$]. Similarly, temporal shifts caused a mean increase from 208 to 234 msec [$F_1(1,47) = 6.69$, $p = .01$, $f = .13$; $F_2(1,7) = 8.89$, $p < .05$]. The effect of the spatial dimension was somewhat smaller, and in the by-participants analysis, it was only marginally significant [$F_1(1,47) = 3.02$, $p < .10$, $f = .09$; $F_2(1,7) = 5.6$, $p < .05$]. As compared with spatial continuity (209 msec), spatial shifts caused mean syllable reading times of 232 msec. Although these means are very similar to the ones observed for protagonist shifts and temporal shifts, the corresponding standard deviations are larger, making the difference less reliable. Of the four possible interactions, only the interaction of protagonist and location reached statistical significance [$F_1(1,47) = 6.71$, $p < .05$, $f = .13$; $F_2(1,7) = 8.05$, $p < .05$]: Reading times were short when both dimensions were continuous (186 msec), whereas they were of comparable length when either dimension or both of them involved a shift (protagonist only, 233 msec; location only, 233 msec; both, 231 msec). In an additional, pooled analysis, we found a strictly monotonic increase of reading times with the number of shifts (no shift, 164 msec; one shift, 220 msec; two shifts, 231 msec; three shifts, 248 msec).

These results conform very nicely with conclusions derived from correlational studies using multiple regression analyses (Zwaan, Magliano, & Graesser, 1995; Zwaan et al., 1998): Both protagonist shifts and temporal shifts increased processing load significantly, whereas the effect of spatial shifts was less reliable. In general, the greater the number of situational shifts involved in a target sentence, the longer it took the readers to comprehend the sentence. This effect was strongest for the difference between continuous sentences and sentences containing a single shift (164 vs. 220 msec): Interrupting the situational continuity on a single dimension was sufficient to cause a strong increase in processing load. In addition, the present experiment yielded an interaction of protagonist shifts and spatial shifts that could not be observed in previous correlational studies: Shifting both space and the protagonist did not cause longer reading times than shifting only one dimension. We addressed this new finding in Experiment 3, after attempting to replicate it in Experiment 2.

EXPERIMENT 2

Experiment 2 was designed to replicate and extend the results observed in the first experiment. First, we wanted to replicate the reading time results reported above. Sec-

Table 1
Mean Target Syllable Reading Times
(in Seconds, With Standard Deviations) in Experiment 1

Spatial Information and Temporal Information	Protagonist Information			
	Continuity		Shift	
	M	*SD*	*M*	*SD*
Spatial continuity				
Temporal continuity	164	60	232	113
Temporal shift	207	114	234	121
Spatial shift				
Temporal continuity	222	106	213	98
Temporal shift	245	121	248	150

Table 2
Mean Target Syllable Reading Times
(in Seconds, With Standard Deviations) in Experiment 2

Spatial Information and Temporal Information	Protagonist Information			
	Continuity		Shift	
	M	SD	M	SD
Spatial continuity				
Temporal continuity	155	53	232	104
Temporal shift	211	94	225	110
Spatial shift				
Temporal continuity	218	117	212	93
Temporal shift	238	111	257	103

ond, the reading times should be complemented by an off-line measure of sentence processing. Therefore, the participants in Experiment 2 also judged the perceived coherence of each target sentence immediately after reading it. These judgments reflect conscious processing of situational shifts after encountering them, whereas reading times may reflect conscious as well as unconscious processes during updating of the situation model. It is currently unclear to what extent indexing and updating of situation models are conscious processes, although results by Rinck, Hähnel, and Becker (2001) have suggested that, at least for the temporal dimension, both conscious and unconscious processes may be involved. Moreover, judging the coherence of sentences may involve careful processing of all situational dimensions, whereas readers may choose to monitor some dimensions more carefully than others during reading.

Method

This experiment was very similar to the first one; therefore, only the new features will be described.

Participants. Forty-eight undergraduates at Dresden University of Technology participated in this experiment, either to fulfill a curricular requirement or for monetary compensation (equivalent to $5). All of them were native speakers of German, and none of them had participated in Experiment 1.

Materials, Procedure, and Design. The same narratives as those in Experiment 1 were used. However, in addition to reading the narratives and answering the final comprehension questions, the participants also judged the perceived coherence of target sentences during reading. Immediately after each target sentence, the instruction *Please give an estimate of how closely the last sentence is connected to the previous sentences* (*Bitte schätzen Sie ein, wie eng sich der letzte Satz an die vorhergehenden anschließt*) appeared on the screen. The participants responded to the question by pressing a numerical key between 1 (labeled *not connected at all*) and 7 (labeled *very closely connected*). After they had responded, the next sentence was displayed, and they resumed the self-paced reading task. To keep the participants from focusing their attention on the target sentences, the same question was also asked at random locations in the first half of each experimental narrative. Moreover, the filler stories contained only one question each, positioned at random locations. The design in Experiment 2 was identical to that in Experiment 1, except for the introduction of the coherence ratings as another dependent variable.

Results and Discussion

Reading times. The $2 \times 2 \times 2$ ANOVA of the target sentence syllable reading times shown in Table 2 yielded

results that were very similar to those in the first experiment. First, the main effect of protagonist shift was significant [$F_1(1,47) = 13.3$, $p < .001$, $f = .19$; $F_2(1,7) = 13.42$, $p < .01$]: If the target sentence contained a new protagonist, mean syllable reading times increased from 206 to 232 msec. Similarly, temporal shifts caused a mean increase from 205 to 233 msec [$F_1(1,47) = 8.69$, $p = .01$, $f = .15$; $F_2(1,7) = 10.03$, $p < .05$]. Again, the effect of spatial shifts was smaller and less reliable, with spatially continuous sentences averaging 206 msec and spatial shift sentences 231 msec [$F_1(1,47) = 3.81$, $p < .10$, $f = .10$; $F_2(1,7) = 6.18$, $p < .05$]. As in Experiment 1, the interaction of protagonist shift and location shift was significant [$F_1(1,47) = 5.17$, $p < .05$, $f = .12$; $F_2(1,7) = 7.87$, $p < .05$], yielding the same pattern of reading times: They were short when both dimensions were continuous (183 msec), whereas they were of comparable length when either dimension or both of them involved a shift (protagonist only, 228 msec; location only, 228 msec; both, 235 msec). In addition, the three-way interaction of protagonist, time, and location was significant [$F_1(1,47) = 7.13$, $p = .01$, $f = .14$; $F_2(1,7) = 7.59$, $p < .05$]. This was due to the fact that the fully continuous version was read much more quickly than any version containing one or two shifts and that the fully discontinuous version was read more slowly than the latter (155 vs. 223 vs. 257 msec).

Coherence ratings. The $2 \times 2 \times 2$ ANOVA of the target sentence coherence ratings depicted in Table 3 yielded a very clear and simple pattern of results: All the main effects were large and highly significant, whereas none of the interactions reached statistical significance. If the target sentence contained a new protagonist, the mean perceived coherence decreased from 3.9 to 3.1 [$F_1(1,47) = 26.58$, $p < .001$, $f = .27$; $F_2(1,7) = 28.22$, $p < .001$]. Similarly, temporal shifts caused the mean coherence ratings to decrease from 4.0 to 3.0 [$F_1(1,47) = 29.28$, $p < .001$, $f = .30$; $F_2(1,7) = 32.14$, $p < .001$]. For these ratings, the effect of spatial shifts was also highly significant [$F_1(1,47) = 34.47$, $p < .001$, $f = .31$; $F_2(1,7) = 34.65$, $p < .001$], with mean ratings decreasing from 4.0 to 3.0. Similar to the reading times, a monotonic decrease of coherence ratings with number of shifts was observed, with mean ratings of 5.0, 3.9, 3.0, and 2.3 being given for sentences with zero, one, two, and three shifts, respectively. We also observed a weak negative correlation between coherence ratings and reading times

Table 3
Mean Coherence Ratings of Target Sentences
(With Standard Deviations) in Experiment 2

Spatial Information and Temporal Information	Protagonist Information			
	Continuity		Shift	
	M	SD	M	SD
Spatial continuity				
Temporal continuity	5.0	1.4	3.9	1.7
Temporal shift	3.9	1.7	3.1	1.8
Spatial shift				
Temporal continuity	3.7	1.8	3.0	1.7
Temporal shift	2.8	1.7	2.3	1.3

($r = -.29$, $p < .05$, after n corrected to 48), suggesting that the readers were, at least in some instances, aware of the situational shifts during reading.

Taken together, the reading times observed in this experiment duplicate the results in Experiment 1 almost perfectly, including the interaction of protagonist shift and location shift. This indicates that the effects of protagonist shifts, temporal shifts, and spatial shifts are robust. Moreover, the replication suggests that the additional task given in Experiment 2 (coherence ratings) did not distort processing of the target sentences. In addition, the coherence ratings themselves yielded large effects of shifts on the three dimensions. Protagonist shifts, temporal shifts, and spatial shifts alike decreased the perceived coherence of the target sentences. Moreover, their effects were additive, jointly decreasing the perceived coherence from a maximum of 5.0 (no shift) to a minimum of 2.3 (three shifts). The fact that coherence ratings reflected large main effects but no interactions will be discussed in the General Discussion section.

EXPERIMENT 3

Experiment 3 was designed to study the observed interaction of protagonist shifts and spatial shifts in more detail. In the first two experiments, both shifts together did not create more difficulty than one shift alone. One possible explanation of this finding states that this particular type of *double shift* is very common and, therefore, highly practiced. Many narratives depict the activities of several characters who act independently of each other and in different locations. In this case, shifting the focus of attention from one character to another naturally involves a location shift as well (but not necessarily a temporal shift, because the characters may act simultaneously). As Magliano et al. (2001) have found, some shift combinations are indeed more common than others in natural materials. However, their results are not directly applicable to our experiments, because the protagonist shifts they studied were defined by movements of the protagonist(s) to new locations, not by (re-)introduction of a different protagonist. This difference is important because it relates to an alternative explanation: If there is *neither* a protagonist shift *nor* a location shift in our materials, the two protagonists do not leave their locations. Interestingly, the same is true if there is a protagonist shift *as well as* a location shift. For instance, even if the target sentence refers to Paul in the garden after the previous episode was about Frieda in the house (see Appendix B), neither Paul nor Frieda left his or her location. For pure protagonist shifts and pure location shifts, however, one person must have moved without this being explicitly mentioned. For pure location shifts, if the target sentence refers to Paul in the garden whereas the previous episode described Paul in the house, he must have moved out into the garden somehow. Likewise, for pure protagonist shifts, if the previous episode was about Frieda in the garden, the target sentence about Paul in the garden also implies his movement there. Thus, sudden and unexplained movements of a protagonist are implied independently from situational shifts. As a result, sentences containing shifts on both the protagonist and the location dimensions may seem more plausible than sentences containing only one of these shifts, because the latter, but not the former, imply unexplained protagonist motions. Experiment 3 was designed to study the effects of these protagonist motions.

Method

Participants. Thirty undergraduates at Dresden University of Technology participated in this experiment to fulfill a curricular requirement. All were native speakers of German, and none had participated in the previous experiments.

Materials, Procedure, and Design. The same narratives as those in Experiment 1 were used. However, only the introduction and the episode corresponding to the fully continuous versions were used (see Appendix A). Each narrative was printed on one page of paper, up to (but not including) the target sentence. Instead of the target sentence, the question *How likely is it that the narrative will continue with a sentence about the following:* (*Wie wahrscheinlich ist es, dass der nächste Satz der Geschichte von folgendem handelt:*) was printed below the story. The question was followed by four statements corresponding to the four combinations of protagonist shifts and location shifts. For instance, for the sample story in Appendix A, the four alternatives read *about Paul in the garden*, *about Paul in the house*, *about Frieda in the garden*, and *about Frieda in the house*. The participants judged the four alternatives by rank-ordering them: They were asked to assign values between 1 and 4 to the alternatives (1 to the *most likely* alternative, 4 to the *least likely*). Two sets of materials, containing four narratives each, were created, and each participant received only one set. The participants were tested individually, most of them (26) in a group session, in a seminar by the first author. The remaining participants were tested in individual sessions. It took them about 20 min to complete the experiment, after which they were debriefed. The mean ranks assigned by the participants to the four alternatives were analyzed according to a 2 × 2 × 2 design with materials set as the between-subjects factor and protagonist shift and location shift as the within-subjects factors.

Results and Discussion

The 2 × 2 × 4 ANOVA of the mean ranks yielded significant main effects of both protagonist shifts [$F_1(1,28) = 33.27$, $p < .001$; $F_2(1,7) = 36.07$, $p < .01$] and location shifts [$F_1(1,28) = 8.4$, $p < .01$; $F_2(1,7) = 9.13$, $p < .05$]. Most important, a strong interaction of protagonist shift and location shift was also observed [$F_1(1,28) = 213.11$, $p < .001$; $F_2(1,7) = 244.5$, $p < .001$]: The statements without any shift were considered the most likely continuation ($M = 1.27$, $SD = 0.37$). These statements received higher rankings than did those containing double shifts [$M = 2.22$, $SD = 0.63$; $t(29) = 6.16$, $p < .001$]. These, in turn, were considered more likely than statements containing *either* a protagonist shift only [$M = 3.43$, $SD = 0.50$; $t(29) = 6.99$, $p < .001$] *or* a location shift only [$M = 3.08$, $SD = 0.51$; $t(29) = 4.79$, $p < .001$]. Thus, the two statements containing a protagonist motion were considered less likely than the statement containing two situational shifts but no protagonist motion. All of these effects appeared in both materials sets; therefore, no interaction of

materials set and the effects of interest occurred [all Fs$(1,28) < 2.10$, n.s.]. In sum, these results reflect the importance of sudden unexplained protagonist motions on the likelihood that readers assign to alternative story continuations. Sentences containing a shift on both the protagonist dimension and the spatial dimension of the situation model do not imply such a motion, which might explain why they did not create more processing load than did sentences containing a single shift on one of these dimensions.

GENERAL DISCUSSION

The present findings add to the small but growing body of empirical evidence supporting the event-indexing model proposed by Zwaan and his colleagues (Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan et al., 1998). Unlike earlier studies, in the present one, an experimental design was used to test the processing load hypothesis derived from the event-indexing model. Of the five event dimensions proposed by the event-indexing model, three were studied here—namely, temporality, spatiality, and protagonist. As is predicted by the processing load hypothesis, information containing shifts on any of these dimensions increased the processing load, as compared with information that was continuous with previous information. The increased processing load was reflected by increased reading times for unmarked target sentences contained in narratives. One might argue that our comparison of different situational dimensions is flawed because it is impossible to equate shifts on different dimensions. For instance, why should noon and evening be as different from each other as house and garden are or as Paul and Frieda are (see Appendix B)? Although this argument is certainly true, it is an unlikely explanation of our results, because these agree very well with earlier ones found with natural texts (Zwaan, Magliano, & Graesser, 1995; Zwaan et al., 1998). Most important, both types of studies suggest that protagonist shifts and temporal shifts increase the processing load significantly, whereas the effect of spatial shifts is weaker. Also, in general, the greater the number of situational shifts involved in the processing of a sentence, the longer it took the readers to comprehend the sentence.

In addition to providing converging evidence, the present experiments also extend earlier studies by a number of new findings. First, reading times of target sentences revealed that interrupting the situational continuity on a single dimension was sufficient to cause a strong increase in processing load. Additional shifts increased reading times further, but not as strongly as the first shift, yielding a negatively accelerated increase in reading times. This result has parallels in many other areas of cognition—for instance, in retrieval from long-term memory. Many experiments on the *fan effect* in which recognition tests have been employed have shown that the largest reaction time difference occurs between concepts without any additional associations in memory and

concepts with a single additional association (e.g., Anderson, 1974, 1983). As compared with this difference, the additional effects of a second or third association were small. Second, the reading times also revealed a particular interaction of protagonist shifts and spatial shifts: Both shifts together did not create more difficulty than did one shift alone. The results of Experiment 3 suggest that this interaction may be related to the plausibility of different combinations: Although both shifts together involve more situational updating than do single shifts, only the latter involve implicit protagonist movements. The space $\times$ protagonist interaction was observed in all the experiments, and it indicates that the event-indexing model needs revision with respect to the assumed independence of dimensions. In fact, the recent correlational study by Magliano et al. (2001) has yielded evidence for interactions, too. The independence assumption has been made mainly for theoretical parsimony, since the early correlational studies had not yielded evidence for interactions. Now that this evidence is available, it needs to be accounted for by the model. One way to do so would be by incorporating indices of frequency and/or plausibility for each combination of shifts, so that frequent and highly plausible combinations would yield smaller or even no increases in processing load.

In addition to the effects on reading times measured on line, situational shifts also affected judgments of perceived coherence. These judgments were given after processing of the situational shifts; therefore, they constitute an off-line measure of conscious processing involved in situation model updating. In fact, these judgments yielded even stronger effects of protagonist shifts, temporal shifts, and spatial shifts than did the reading times, indicating that readers are well aware of the situational discontinuities introduced by shifts. Most probably, coherence ratings yielded stronger effects than did reading times since the former are more reliable than the latter, particularly since each participant supplied only one reading time in each experimental condition. The coherence ratings, however, did not reveal evidence of the space $\times$ protagonist interaction observed with reading times. This suggests that off-line measures based on conscious postprocessing of information cannot be an appropriate substitute for on-line measures of processing. The lack of a significant space $\times$ protagonist interaction in coherence ratings illustrates that readers are able to detect shifts on each dimension when asked to think about them. However, the interaction observed with reading times suggests that additional shifts do not necessarily increase processing load correspondingly. Other studies in which both on-line and off-line measures of comprehension have been employed have also yielded slightly divergent results for the two types of dependent variables (e.g., Scott Rich & Taylor, 2000), suggesting that on-line measures may be a better indicator of processing load than off-line measures are.

Neither reading times nor coherence ratings gave any indication of the protagonist dimension superiority that

had been postulated by Scott Rich and Taylor (2000). Instead, both protagonist shifts and temporal shifts yielded stronger effects on reading times than did spatial shifts. We strongly suspect that this difference in the importance of protagonist shifts was caused by the particular type of shifts employed in the experiments. In our narratives, protagonist shifts involved the reintroduction of a character mentioned previously. On the other hand, the narratives used by Scott Rich and Taylor contained protagonist shifts that involved the introduction of an unknown new character. In processing this more drastic shift, readers may have wondered whether they should update the current situation model at all or create a new one instead. Results by Magliano et al. (2001) are also inconsistent with the postulated dominance of the protagonist dimension: They found an interaction between time and protagonist movement indicating that time is monitored independently of protagonist movement, whereas monitoring of movement was affected by shifts in time. Together, these results suggest that the theoretical concept of situational shifts, or at least protagonist shifts, may need refinement (as was already suggested by Zwaan, 1999). Rather than being all or none, protagonist shifts may differ in degree, depending, at least, on whether they involve the reintroduction of a known character or the introduction of a new one. This hypothesis should be tested directly in future experiments. These experiments should also address the other dimensions postulated by the event-indexing model—that is, causality and intentionality.

To summarize, our experimental data are quite consistent with earlier results of correlational studies and extend them in meaningful ways. Naturally, both approaches have specific disadvantages: Most important, correlational techniques do not allow for causal conclusions, whereas experimental techniques tend to employ rather artificial materials. However, the high consistency observed here suggests that the results found with both techniques are not seriously compromised by these disadvantages and that the differing methodologies complement each other well. Therefore, we suggest continuing to employ experimental approaches in addition to correlational ones. Admittedly, the design of experimental materials is difficult, and we strongly agree with the closing comments of Zwaan (1999): "It will require a great deal of imagination, norming studies, and pilot work to come up with stimulus materials in which relationships are manipulated while relationships on other dimensions are kept constant" (p. 108). However, we believe that the results of experimental tests of the event-indexing model—and of course, also other models of text comprehension—will be worth the efforts involved. It is to be hoped that the experiments presented here will provide a small step ahead on the way to achieving this goal.

REFERENCES

ANDERSON, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, **6**, 451-474.

ANDERSON, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

BOWER, G. H., & RINCK, M. (2001). Selecting one among many referents in spatial situation models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 81-98.

COHEN, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

GERNSBACHER, M. A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Erlbaum.

GLENBERG, A. M., KRULEY, P., & LANGSTON, W. E. (1994). Analogical processes in comprehension: Simulation of a mental model. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 609-640). San Diego: Academic Press.

GLENBERG, A. M., MEYER, M., & LINDEM, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory & Language*, **26**, 69-83.

JOHNSON-LAIRD, P. N. (1983). *Mental models.* Cambridge: Cambridge University Press.

KINTSCH, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, **95**, 163-182.

KINTSCH, W. (1998). *Comprehension.* New York: Cambridge University Press.

MAGLIANO, J. P., MILLER, J., & ZWAAN, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology*, **15**, 533-545.

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, & Cognition*, **27**, 67-80.

SCOTT RICH, S., & TAYLOR, H. A. (2000). Not all narrative shifts function equally. *Memory & Cognition*, **28**, 1257-1266.

VAN DIJK, T. A., & KINTSCH, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

WILLIAMS, P., & TARR, M. J. (n.d.). *RSVP: Experimental control software for MacOS* [On line]. Retrieved November 12, 2003, from http://www.cog.brown.edu/~tarr/rsvp/.

ZWAAN, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1196-1207.

ZWAAN, R. A. (1999). Five dimensions of narrative comprehension: The event-indexing model. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 93-110). Mahwah, NJ: Erlbaum.

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., 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, & Cognition*, **21**, 386-397.

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.

NOTE

1. The experimental narratives always contained two main characters; therefore, both of them are called *protagonists*.

APPENDIX A
Excerpt of Sample German Text in Fully Continuous Version, With English Translation

1. *Introduction:*

Seit Paul und seine Frau Frieda Rentner geworden waren, steckten sie den Großteil ihrer Kraft in ihr kleines Haus und den zugehörigen Garten. An einem sonnigen Tag im Mai beschloss das Rentnerehepaar, einen großen Frühjahrsputz zu veranstalten; mit allem, was dazu gehört.

Ever since Paul and his wife Frieda had retired, they put a lot of effort in their little house and their garden. On a sunny day in May, they decided to do a big spring cleaning including everything involved.

2. *Episode with Variation of Protagonist, Time, and Location:*

Frieda begann damit, im Haus aufzuräumen. Paul dagegen widmete sich voller Elan dem Garten, den er zu seinem Arbeitsterritorium erklärt hatte. In der warmen Mittagssonne begann er, sorgfältig und selbstvergessen die letzten welken Blätter des vergangenen Herbstes zusammenzurechen. Dann grub er die Beete um, säte Bohnen, steckte Kartoffeln und pflanzte Stiefmütterchen und Tomaten an. Gegen Abend holte Paul seinen nagelneuen Rasenmäher, den er erst zu Weihnachten bekommen hatte, aus dem Geräteschuppen und kürzte den kleinen Rasen im Garten. Da es bereits dämmerte, musste er sich richtig beeilen, um noch rechtzeitig mit dem Rasenmähen fertig zu werden.

Frieda started to tidy up the house, whereas Paul took care of the garden, which he had declared his territory. In the warm midday sun, he started to carefully clean up last year's withered leaves. Then he dug up the beds, sowed beans and potatoes, and he planted pansies and tomatoes. In the evening, Paul fetched his new lawn mower, which he had gotten for Christmas, from the shed, and he cut the small lawn in the garden. It was twilight already, so he had to hurry in order to finish mowing the lawn in time.

3. *Target Sentence:*

Im letzten Tageslicht stand Paul im Garten und blickte sich zufrieden um.

In the last daylight, Paul stood in the garden and looked around satisfied.

4. *Concluding Sentences*

5. *Comprehension Question:*

Hat Paul im Garten gearbeitet?

Did Paul work in the garden?

Note—Some of the German materials may sound awkward when translated into English. They were judged to be perfectly fine in German, however, as the pilot tests showed.

APPENDIX B
Excerpt of Sample German Text in Fully Discontinuous Version, With English Translation

1. *Introduction:* See Appendix A

2. *Episode with Variation of Protagonist, Time, and Location:*

Paul widmete sich voller Elan dem Garten, den er zu seinem Arbeitsterritorium erklärt hatte. Frieda dagegen begann damit, im Haus aufzuräumen. Da die Morgenluft besonders angenehm und erfrischend war, riss sie erst alle Fenster weit auf und ließ den Frühling bis in die letzten Ecken des Hauses vordringen. Dann kletterte sie voller Abenteuerlust auf den Dachboden, kramte in alten Kisten und wackligen Schränken herum und jagte nach Mäusen. Gegen Mittag kehrte sie im Flur den Winterstaub aus und brachte ihren über alles geliebten und wohlgehüteten Wäscheschrank in Ordnung. Da die Mittagssonne schon recht warm war, unterbrach sie diese Arbeit und schloss erst einmal alle Fenster, damit es im Haus angenehm kühl blieb.

Paul took care of the garden, which he had declared his territory, whereas Frieda started to tidy up the house. The morning air was pleasant and refreshing, so she opened all windows and let spring reach every corner of the house. Then, in a spirit of adventure, she climbed up to the attic. There she searched old boxes and shaky cupboards, and she checked for mice. At noon, she cleaned the winter dust out of the hall and tidied up her beloved cabinet. The noon sun was quite warm already, so she interrupted this work for some time while she closed all windows to keep the house pleasantly cool.

3. *Target Sentence:* See Appendix A

4. *Concluding Sentences and Comprehension Question:* See Appendix A

Note—Some of the German materials may sound awkward when translated into English. They were judged to be perfectly fine in German, however, as the pilot tests showed.