Segmentation in Reading and Film Comprehension

Jeffrey M. Zacks, **Nicole K. Speer**, **Jeremy R. Reynolds**
Washington University

Abstract

When reading a story or watching a film, comprehenders construct a series of representations in order to understand the events depicted. Discourse comprehension theories and a recent theory of perceptual event segmentation both suggest that comprehenders monitor situational features such as characters' goals in order to update these representations at natural boundaries in activity. However, the converging predictions of these theories had not previously been tested directly. Two studies provided evidence that changes in situational features such as characters, their locations, interactions with objects, and goals are related to the segmentation of events in both narrative texts and films. A third study indicated that clauses with event boundaries are read more slowly than other clauses, and that changes in situational features partially mediate this relationship. A final study suggested that the predictability of incoming information influences reading rate and possibly event segmentation. Taken together, these results suggest that processing situational changes during comprehension is an important determinant of how one segments ongoing activity into events, and that this segmentation is related to the control of processing during reading.

Keywords

Situation models; Event perception; Text comprehension

The world is presented to us as a stream of continuous information, but we are able to perceive this information as a series of discrete units (Newtson, 1973; Zacks & Tversky, 2001). This process is called *event structure perception*, and the discrete units, or *events*, are defined as segments of time at particular locations that are perceived by observers to have beginnings and endings (Zacks & Tversky, 2001). The everyday-language term 'event' refers to things that may be very brief (a lightning strike) or very long (the birth of a solar system), but the psychological events we are concerned with here have durations on a human scale, spanning a few seconds to tens of minutes (Barker & Wright, 1954; Dickman, 1963).

Representative examples include a person opening a door, a couple driving to the store, and a baseball game.

Such events have specific structures, and involve particular objects, characters, and goals. Knowledge about the structure of events is essential to function in the world, and is used to fill in missing information, predict what is going to happen in the future, and plan actions (Grafman, 1995; Hommel, 2006; Abelson, 1981; Zwaan & Radvansky, 1998). These *event schemata* (also referred to as 'scripts' or 'structured event complexes') provide a framework for incoming information, such that the active representation of the current event (the *event model*) is made up of information about the current state of the world, as well as information about similar, previously encountered states (Zacks, Speer, Swallow, Braver, & Reynolds, 2007).

Event Structure Perception

The perceived structure of events can be measured explicitly by simply asking people to watch movies of everyday events and identify the points where they believe one meaningful unit of activity ends and another begins (Newtson, 1973). Although these instructions are somewhat vague, there is good agreement both across and within individuals as to where one unit of activity ends and another begins (Newtson, 1976; Speer, Swallow, & Zacks, 2003; Zacks, Tversky, & Iyer, 2001b), and observers are able to adjust their grain of segmentation in order to identify larger or smaller units of activity.

Event structure can also be measured implicitly using physiological measures, such as neuroimaging. In one study (Zacks et al., 2001a), observers watched movies of everyday events while they were scanned with functional Magnetic Resonance Imaging (fMRI). After watching the movies, the same observers segmented the movies into individual events. Although participants did not know about the segmentation task during the original viewing of the movies, a network of regions in the brain increased in activity around the points in time participants later identified as boundaries between units of activity. These results indicate that observers are sensitive to event boundaries even when they are not explicitly attending to event structure.

Although it is clear that observers can and do perceive the structure of events, it is not yet clear what leads to the perception of this structure. *Event Segmentation Theory* (EST) provides a computationally and neurophysiologically explicit account of event structure perception (Reynolds, Zacks, & Braver, 2007; Zacks et al., 2007). According to EST, observers form models in working memory called *event models* that guide the perception of incoming information. These event models facilitate predictions as to what will occur in the environment, and as long as incoming information is consistent with these predictions the current event model remains active. When predictions fail, the current event model is updated and the system becomes momentarily more sensitive to incoming information. Once the event model is updated, prediction error falls and the system falls into a new stable state. The cascade of a transient increase in error, event model updating, and resettling is experienced as an event boundary.

EST proposes that event boundaries tend to occur when features in the environment are changing, because changes are generally less predictable than stasis. Two broad classes of feature can be distinguished: *perceptual* features such as movement, color, and sound timbre, and conceptual features such as characters in a story, characters' goals, and causes. There is growing evidence that changes in perceptual features are associated with event boundaries. Observers are more likely to identify an event boundary at points with large changes in movement (Newtson, Engquist, & Bois, 1977), and brain areas involved in the perception of motion are activated at event boundaries (Speer et al., 2003; Zacks et al., 2001a). More recently, studies using simple animations have provided quantitative evidence that precisely measured movement features, including acceleration, distance, and speed, are excellent predictors of where observers segment activity (Hard, Tversky, & Lang, 2006; Zacks, 2004).

The evidence that conceptual changes influence event segmentation is more limited. Baldwin and colleagues found that when they presented infants with movies of goal-directed activities, the infants were more surprised to see a pause inserted in the middle of a goal-directed activity than a pause inserted after the completion of the activity (Baldwin, Baird, Saylor, & Clark, 2001). Converging methods support the conclusion that the infants' perception was sensitive to these goal achievements (Saylor, Baldwin, Baird, & LaBounty, 2007). Studies of adults also suggest that goals are correlated with the perception of event boundaries (Baird & Baldwin, 2001). When observers watched a movie and were asked to identify the locations of periodic beeps that were heard during the movie, they were more likely to remember those beeps when they occurred at points where goals were completed. Such findings suggest that at least one conceptual change (goal state) is correlated with event segmentation. However, as these authors note, goal changes often co-occur with changes in actors' movements.

Language Understanding and Discourse Comprehension

Natural language allows speakers and writers to make explicit some conceptual dimensions that are implicit in live-action events or movies. In language, one can easily describe changes in dimensions including goals ("Barney decided to stop at the grocery store."), causes ("The melting ice cream stained Delia's car seat."), and temporal relations ("The rain started just as Kim's bus arrived.") Language has numerous sentence-level devices for representing relations amongst events. These include phrase and clause structure, verb-chaining (e.g., Pawley & Lane, 1998), and verb tense and aspect (e.g., Pustejovsky, 1991; Moens & Steedman, 1988). The structure of these devices can provide important information about humans' conceptual representation of events. However, language-specific features of event representations are less directly applicable to understanding the perception of event structure. Language also has devices for representing event structure at the discourse level. The use of discourse-level information by readers to understand event structure is more directly informative with regard to perception.

Current theories of discourse comprehension propose that readers track multiple dimensions of the situation being described by the text in an active *situation model* (van Dijk & Kintsch, 1983). Two particularly relevant accounts are the *Structure Building Framework*

(Gernsbacher, 1990) and the *Event Indexing Model* (Zwaan, 1999; Zwaan, Langston, & Graesser, 1995a; Zwaan & Radvansky, 1998; Radvansky, Zwaan, Federico, & Franklin, 1998). Whereas the event perception literature has focused specifically on actors' goals and intentions as cues to event structure, researchers in discourse comprehension have examined a broader range of conceptual cues, such as changes in characters or their interactions with objects, that may also play an important role in the perception of events. These accounts claim that readers track situational dimensions including characters, their locations, goals, and the causal and temporal relationships present in the story (Gernsbacher, 1990; Zwaan, 1999; Zwaan & Radvansky, 1998; Zwaan, Radvansky, Hilliard, & Curiel, 1998). Readers are thought to use these dimensions to construct a succession of situation models that are bound together based on the degree of overlap on these dimensions. When a reader encounters a break in any of these dimensions, the current situation model is updated (Glenberg, Meyer, & Lindem, 1987; Rinck & Bower, 2000; Zwaan, 1996), and the degree of overlap on the different dimensions will determine the degree to which readers update the current model and integrate similar models (Magliano, Zwaan, & Graesser, 1999; Zwaan et al., 1998; Gernsbacher, 1990). Similar proposals have been made for the comprehension of film (Magliano, Miller, & Zwaan, 2001) and in virtual reality (Copeland, Magliano, & Radvansky, 2006; Radvansky & Copeland, 2006).

Here, we propose that terms "event model" in EST and "situation model" in the Structure Building Framework and the Event Indexing model describe functionally equivalent mental representations. (Zwaan and Radvansky, 1998, distinguish between a reader's current situation model and an integrated situation model that results from forming a succession of current situation models; here we refer to the current situation model.) Although previous research in event perception has focused more on perceptual features, and previous research in discourse comprehension has focused more on conceptual features, we believe this reflects accidents of methodology rather than substantive theoretical disagreement. The experiments described here were designed to test whether the specific mechanisms proposed by EST applied to discourse comprehension, thereby extending discourse comprehension theories, and to test whether the conceptual changes proposed by the Event Indexing Model applied to the perception of live-action events, thereby extending EST.

Overview of Experiments

The experiments reported here used extended narratives to explore the relations between conceptual changes and event segmentation. Changes were identified based on the Event Indexing Model (Zwaan et al., 1995a; Zwaan et al., 1998), which was developed to account for reading comprehension. Narrative texts offer an attractive feature for studying conceptual changes: They remove potentially confounding effects of low-level features such as actor and object motion and configurations of objects. Therefore, extended narrative texts were the stimuli used in Experiments 1, 3, and 4. However, generalization from narrative texts to online perception is not a guaranteed proposition. To provide converging evidence using materials closer to live event perception, Experiment 2 used narrative film stimuli. Experiments 1 and 2 directly measured the relations between changes in conceptual features and event segmentation, in texts and movies respectively. In both cases, changes in the narrative situation predicted where observers perceived event boundaries. Experiments 3

and 4 tested two proposals of EST: Experiment 3 used reading time to investigate the proposal that processing of situational changes produces transient increases in some aspects of processing, whereas experiment 4 tested the theory's proposal that event boundaries tend to occur when activity is less predictable. As predicted, reading was slower for event boundaries than for non-boundaries in Experiment 3, and story elements that were rated less predictable in Experiment 4 were associated with event boundaries and with longer reading times. These predicted effects—and some more surprising interactions—have implications for theories of narrative comprehension and event understanding, which we elaborate at the end of the paper.

Experiment 1

In the first experiment, participants heard or read a series of continuous narratives describing everyday events. These narratives were taken from the book, *One Boy's Day* (Barker & Wright, 1951), and consisted of highly detailed descriptions of the everyday activities of a seven-year old boy. Unlike typical narratives, these stories do not have any jumps in time (e.g., "The next day…"), and are faithful descriptions of directly observed activities in the real world.

The narratives were presented in three different modalities; this was done to provide as informative as possible a comparison between reading comprehension and on-line event perception. Separate groups of participants received each presentation mode. Participants in the *auditory* group listened over headphones to a trained narrator's reading of the stories. Their experience was closest to event perception: Information arrived at a pace determined by the environment, and previous information was not reviewable. The *visual-continuous* group read the narratives printed on paper, at their own pace. Their experience was typical of reading for comprehension; readers could proceed at their own pace with the ability to review previous sentences. The *visual-sequential* group read the narratives presented one clause at a time on a computer screen, at a pace fixed by the experimental software. Their experience was intermediate between that of the auditory and visual-continuous groups— they had some ability to review information within a clause, but could not review previous clauses and the pacing of information was determined by the software. While hearing or reading the narratives, participants were asked to segment the narratives by identifying the points where they believed one meaningful unit of activity ended and another began.

To determine which variables influence readers' explicit perception of the structure of the narratives, each clause in the narratives was coded along six dimensions, closely based on the Event Indexing Model (Zwaan et al., 1995a; Zwaan et al., 1998): references to story time, changes in the causal relationships between narrated events, changes in characters, and changes in characters' locations, interactions with objects, and goals. If participants use these conceptual cues to structure activity during comprehension, then participants should identify changes in these dimensions as explicit boundaries between units of narrated activity.

Each participant segmented a set of narratives twice—once to mark the largest units they found meaningful (*coarse segmentation*) and once to mark the smallest units they found

meaningful (*fine segmentation*). In addition to measuring the effects of conceptual changes on coarse and fine event segmentation, this procedure also provided a measure of the degree to which participants encode narratives in terms of the relations between parts and sub-parts. In the perception of live-action movies, participants segment activity hierarchically, grouping small units of activity into larger wholes (Zacks et al., 2001b). We hypothesized that the same pattern would hold for narratives.

Method

Participants.—Ninety-one undergraduates (ages 18–24, 41 men) participated in Experiment 1 for course credit or a cash stipend. Nineteen additional participants were excluded due to failure to follow the task instructions (e.g., looking away from the screen during story presentation; n = 11), lack of a full data set (n = 6), or computer malfunctions (n = 2).

Materials.—Four excerpts from the book, *One Boy's Day* (Barker & Wright, 1951), were used as the narratives in Experiments 1, 3 and 4. *One Boy's Day* is a record of the real-life activities of a seven year-old boy (referred to by the pseudonym "Raymond Birch") during a 12-hour period on April 26, 1949. The scenes used in the current study described Raymond getting up and eating breakfast ("Waking up"), playing with his friends on the school ground ("Play before school"), performing an English lesson in school ("Class work"), and participating in a music lesson ("Music lesson"). The original narratives occasionally mentioned Raymond's interactions with the observers who recorded his activities; these references were edited out and the scenes were shortened where necessary to keep the length of the narratives below 1,500 words (Waking up, 1368 words; Play before school, 1104 words; Class work, 1182 words; Music lesson, 1404 words).

Participants segmented the narratives in one of three different presentation conditions. For participants in the *auditory* group (n = 32), the narratives were presented over headphones. The auditory narratives were read by a female narrator and digitally recorded. For participants in the *visual-continuous* group (n = 32), the narratives were presented single-spaced on 8 1/2 × 11 paper, and although they retained their original punctuation, they did not contain paragraph marks. For participants in the *visual-sequential* group (n = 27), the narratives were presented one clause at a time on Dell computers (Dell Computer, Round Rock, TX) running SuperLab 2.0 (Cedrus Corporation, San Pedro, CA). Clauses were defined as a verb with its argument structure. Complement clauses, subordinate clauses, and relative clauses that were dominated by a larger inflectional phrase were grouped with the larger unit. "Waking up" consisted of 192 clauses; "Play before school," 178 clauses; "Class work," 172 clauses; "Music lesson," 215 clauses. Each clause appeared centered on a 15" computer monitor in 16 point font. The full narrative set, with clause boundaries marked, is available at http://dcl.wustl.edu/DCL/stimuli.html.

Design and Procedure.—Participants in the visual-continuous group were instructed to read and segment the narratives at their own pace, and participants in the auditory group heard the narratives at the pace at which they were recorded. The presentation rate for the clauses in the visual-sequential condition was determined by computing the average

per-syllable reading time in Experiment 3 across all clauses, participants, and conditions, and multiplying the number of syllables in each clause by this average value (217 ms). The clauses were separated by a 500 ms inter-clause interval to provide enough time for participants to segment the narratives. (The resulting mean presentation rate was 3.35 words/s, SD = 0.70. This was comparable to the reading rate of the narrator for the auditory presentation, who read with a mean rate of 3.23 words/s, SD = .77, and paused for an average of .91 s between each clause, SD = .71.)

After providing informed consent, participants were given a practice narrative that described Raymond casting a fishing line with his father. Participants in the coarse segmentation condition were instructed to identify the largest units of activity that seemed natural and meaningful while reading this practice narrative; participants in the fine segmentation condition were instructed to identify the smallest units of activity that seemed natural and meaningful while reading the practice narrative. All participants were informed that there was no right or wrong way to do this task, and that the experimenters were simply interested in where the participants thought these meaningful units of activity ended and began. The means by which participants indicated their event boundaries depended on the modality in which the stimuli were presented. Participants in the auditory and visual-sequential groups pressed a button on the computer keyboard whenever they perceived one unit ending and another beginning. Participants in the visual-continuous group drew a line between two words to indicate where they perceived the event boundaries. Participants were required to identify at least three (coarse segmentation) or six (fine segmentation) units of activity in the practice narrative before proceeding with the experiment. If participants had not reached this criterion after the first time through the practice narrative, the task was explained again, and they were asked to repeat the practice task, identifying a few more units than they had identified previously (eight participants repeated the initial practice task once to reach criterion).

Once it was clear that participants understood the task, the task continued with the four experimental narratives. The order of the narratives was counterbalanced across participants. Participants then performed the segmentation task for a second time, segmenting at the grain other than the one initially tested. They were again given practice before segmenting the target narratives. During this practice session, participants who were initially in the coarse segmentation condition were required to identify more units of activity than they had identified during the first practice session (and vice versa). After performing the practice task at the new segmentation grain, participants were given the same four narratives they had read or heard previously, and were instructed to segment those narratives at the new grain of segmentation.

Scoring.—For all participants, the narratives were scored such that a given clause was coded as containing an event boundary if the participant had identified at least one boundary during that clause. For participants in the visual-continuous condition, event boundaries identified at sentence boundaries (e.g., periods) were associated with the clause following the sentence boundary (Speer & Zacks, 2005).

Each clause was rated on six situation model dimensions by the first and second authors. We assessed whether or not a given clause contained a change in any of the six dimensions, leading to a binary (yes/no) coding scheme for each dimension. Although there were no temporal changes, each clause was coded for the presence or absence of a *temporal reference* (e.g., "immediately" or "slowly"). Each clause was coded for the presence or absence of a *spatial change*, which consisted of changes in the narrative location, such as moving from one room in a house to another, or changes in the locations of characters, such as moving from one side of a room to another. *Object changes* were coded any time a character changed his or her interaction with an object (e.g., Raymond picking up a candy Easter egg). *Character changes* were coded whenever the subject of a clause changed (e.g., if Raymond were the subject of clause n, and Susan were the subject of clause n+1, clause n+1 would be coded as having a character change). A clause was coded as having a *causal change* whenever the activity described in the clause was not directly caused by an activity described in the previous clause (e.g., a character initiating a new action). *Goal changes* were coded whenever a character began a new, goal-directed activity, such as when a character initiated speaking. Average inter-rater reliability across the six dimensions was .77, as measured by Cohen's Kappa. Discrepancies in the ratings were resolved by discussion. Figure 1 gives a coded example excerpt; the full narratives and codes can be found online at http://dcl.wustl.edu/DCL/stimuli.html.

In addition to situation model changes, each clause in the visual presentation conditions was coded for four structural variables: the *number of syllables* in the clause, whether or not the clause contained any *terminal punctuation* (e.g., periods and question marks) and/or *non-terminal punctuation* (e.g., commas and semi-colons), and the *narrative* from which the clause originated. Each clause in the auditory presentation condition was coded for three structural variables: the *length of the clause*, the *length of the pause following each clause*, and the originating *narrative*. These structural variables, as well as the six situation model dimensions, were used to predict the locations of the coarse and unit boundaries identified by each participant.

Results

The initial analyses examined the degree of hierarchical alignment between the locations of the coarse and fine unit boundaries. Following these, logistic regressions were conducted for each participant to determine if and how the situation change variables predicted patterns of segmentation.

Alignment Effects.—The number of clauses in a narrative was divided by the number of boundaries identified in that narrative to give a measure of the average length (in clauses) of the units identified in the coarse and fine segmentation tasks. Descriptive statistics for these unit lengths are given in Table 1. The number of clauses for each participant identified as both coarse and fine unit boundaries (number of overlapping clauses) was used to assess the degree to which the coarse and fine unit boundaries aligned with one another. Expected overlap values were computed for each participant (see Zacks et al., 2001b), and subtracted from the observed overlap values to generate a measure of overlap that controls for the amount of overlap expected by chance alone. Observed overlap significantly

exceeded that expected by chance for all narratives for all groups [See Figure 3; auditory: smallest $t(31) = 5.02$, $p < .001$; visual-sequential: smallest $t(26) = 2.97$), $p = .006$; visual-continuous: smallest $t(31) = 7.60$, $p < .001$] . That is, coarse and fine event boundaries aligned significantly. (A statistically significant degree of alignment also was observed when the first tasks were compared across participants, indicating that the order of the tasks was not solely responsible for the observed alignment of coarse and fine boundaries.)

Predictors of Event Structure – Number of Changes.—To assess the overall relation between situation changes and event segmentation we carried out logistic regressions predicting the pattern of segmentation for each individual from the number of changes in a given clause, as well as the structural variables. Each clause received a value according to the number of situation changes present in that clause. For example, if a clause had a spatial change and a goal change, that clause would receive a score of two. Because there were only 25 clauses with more than three changes (accounting for 3.30% of the trials), any clause with three or more changes was given a score of three. (There were no changes in 280 of the clauses; 210 clauses had one change; 150 clauses had two changes; and 109 clauses had three or more changes.)

As the number of changes in a clause increased, the probability of segmentation during that clause increased. Figure 2 depicts this effect, and Table 2 gives the mean coefficients for each group and segmentation grain. A 2×3 ANOVA with grain and presentation group as the two independent variables determined that the coefficients for the number of changes variable were larger for participants in the visual-continuous group than those in the visual-sequential group, and both visual groups had larger coefficients than the auditory group, $F(2, 88) = 42.43$, $p < .001$. Coefficients for the number of changes variable were larger in the coarse grain segmentation task than in the fine grain segmentation task, $F(1, 88) = 5.52$, $p = .021$, but only for the visual-continuous group [auditory: $t(31) = .73$, $p = .47$; visual-sequential: $t(26) = -.42$, $p = .68$; visual-continuous: $t(31) = 3.92$, $p < .001$; grain-by-group interaction: $F(2,88) = 4.38$, $p = .015$].

Hierarchical regressions were conducted to assess the unique contribution of the number of situation changes to event segmentation. For each participant, separate hierarchical regressions predicted the pattern of segmentation in the coarse segmentation task and in the fine segmentation task. Models first were fit using only the structural variables as predictors; the number of situation changes was added in a second stage. The mean Nagelkerke's R^2 (Nagelkerke, 1991) values for the models including the number of changes in a clause as a predictor is shown in the middle of Table 4 for each of the three groups, along with 95% confidence intervals for each R^2 under the null hypothesis. The null hypothesis distributions were constructed for each group and each segmentation grain by Monte Carlo simulation. Specifically, we 1) sampled the segmentation data from a single participant, 2) randomly distributed segment boundaries within each narrative for that participant, 3) ran the hierarchical regression described above, 4) computed the Nagelkerke's R^2 values for the structural and situation change variables, 5) repeated steps one through four 32 times (or 27 times for the visual-sequential group), 6) calculated the mean Nagelkerke's R^2 values for the structural and situation change variables across all "participants", and 7) repeated steps one through six 1,000 times. The mean Nagelkerke's R^2 values were well above the 95%

confidence interval for the null hypothesis, indicating that the number of changes variable accounted for a statistically significant amount of variability above and beyond the structural variables. As can be seen in the top of Table 4, the structural variables also accounted for a statistically significant amount of variability in coarse and fine segmentation.

Predictors of Event Structure – Type of Change.—Logistic regressions were conducted for each participant in order to determine which of the structural and situation change variables predicted patterns of segmentation. For each segmentation condition for each participant, a logistic regression predicted whether or not a clause was identified as an event boundary as a function of the structural variables and zero-one variables coding for changes on each situation dimension. Descriptive statistics for and correlations between the predictor variables (with the exception of the dummy variables coding for narrative) are shown in Table 3. Because there were moderate correlations between terminal and non-terminal punctuation ($r = -.33$), and between goal changes and character and causal changes ($r = .38$ and $r = .34$, respectively), the condition index was computed to determine whether substantial multicollinearity was present (Lattin, Carroll, & Green, 2003). The condition index was 4.07 in the model for the auditory group, and 15.32 in the model for the visual groups, both of which were well within the acceptable range. Each structural variable was centered around its mean (by subtracting the mean value from each individual value) in order to facilitate interpretation of the situation change coefficients in the regression analyses. The graphs in Figure 4 show odds ratios for each coefficient in the regressions. These odds ratios represent the multiplicative change in the odds of identifying a boundary given the presence of a situation change. For example, an odds ratio of 1.5 for a situation change variable would indicate that the odds of identifying a boundary on a given clause were 1.5 times greater if that clause contained a change on that dimension. An odds ratio of 1.0 would indicate that changes on that dimension were unrelated to segmentation.

The odds ratios shown in Figure 4 demonstrate that individual situation change variables reliably predicted patterns of segmentation in both the coarse and fine grain tasks. In the coarse grain segmentation task, temporal references, as well as spatial, character, causal, and goal changes, were each associated with increases in the odds that participants would identify a clause as an event boundary. These results were mirrored in the fine grain segmentation task. In addition, changes in characters' interactions with objects were associated with increases in fine grain segmentation. To assess the statistical significance of differences between these associations, the coefficients for each change variable were entered into a $3 \times 2 \times 6$ ANOVA, with group (auditory, visual-continuous, and visual-sequential) as the single between-subjects variable, and grain (coarse and fine) and situation change (temporal, spatial, object, character, cause, and goal) as the two within-subject variables. Segmentation grain influenced the coefficients for each of the situation change variables differentially: Coefficients for causal changes, goal changes, and spatial changes were larger for coarse grain segmentation than fine grain segmentation, whereas coefficients for object changes were larger for fine segmentation than for coarse segmentation (smallest t for coarse vs. fine difference = 1.97, $p = .05$). This pattern held both for participants who segmented first at a coarse grain and for those who segmented first at a fine grain. Coefficients for character changes and temporal references did not differ reliably across

segmentation grains, largest $t = 1.71$, $p = .09$. However, this interaction between situation change and segmentation grain varied across the three groups, $F(10, 440) = 2.85$, $p = .002$. The effects of grain on regression coefficients for character, object, spatial, and goal changes did not differ across the three groups, largest $F = 1.91$, $p = .15$. For causal changes and temporal references, coefficients were greater for coarse segmentation than for fine segmentation in only the visual-continuous group, smallest $t = 4.61$, $p < .001$.

Overall, the influence of the situation change variables on coarse and fine segmentation was relatively consistent across presentation groups. However, participants in the visual-continuous group had larger coefficients than participants in the visual-sequential group, and both visual groups had larger coefficients than participants in the auditory group, $F(2, 88) = 32.61$, $p < .001$. The size of the coefficients varied significantly across the different situation changes, $F(5, 440) = 19.04$, $p < .001$, but only in the visual presentation conditions, $F(10, 440) = 11.99$, $p < .001$. There was not an overall effect of segmentation grain on the size of the coefficients, $F(1,88) = .51$, $p = .48$, but the effect of segmentation grain differed across the three groups, $F(2, 88) = 3.97$, $p = .02$.

Hierarchical regressions were conducted to assess the amount of variance in segmentation accounted for by the situation changes. The hierarchical regressions had the same form as those for the number of situation change, described above, except that the situation change variables themselves were entered in the second step. The bottom of Table 4 shows the Nagelkerke's R^2 values that can be attributed to each set of variables for each group, along with the means and 95% confidence intervals for the Monte Carlo-simulated null hypothesis distributions. The bottom of Table 4 shows that the situation change variables accounted for larger amounts of incremental variability in coarse and fine segmentation than would be expected by chance across all three groups.

Discussion

Patterns of segmentation for the narratives were similar to patterns of segmentation for movies of everyday events. Readers and listeners identified smaller units of activity in the fine segmentation task than in the coarse segmentation task, indicating that they were able to distinguish between small and large meaningful units of activity. Readers and listeners also tended to segment the narratives hierarchically, grouping smaller units of activity into larger wholes. Viewers show these same patterns of segmentation when segmenting movies of everyday events (Zacks et al., 2001b), suggesting that similar mechanisms may be involved in segmenting activities in narratives and in real-world events.

Readers and listeners perceived changes in the six situation model dimensions as event boundaries. In all three groups, situation changes predicted the locations of individuals' event boundaries: temporal references, changes in causal relationships, characters, their goals, locations, and interactions with objects all were significant predictors of segmentation. These results suggest that a number of different conceptual cues are important in the comprehension of narrated activities, across visual and auditory modalities. Regardless of whether participants read or heard the stories, situation changes predicted patterns of event segmentation. The consistency across modalities suggests that common mechanisms govern the segmentation of events independent of whether the pacing of

information presentation is controlled by the comprehender or not, and independent of whether previously presented information is reviewable.

Across all three groups, changes in characters' interactions with objects predicted the locations of fine event boundaries. These interactions typically described characters picking up objects or putting them down, such as putting a coat into a bicycle basket, or picking up a pencil. One possibility is that descriptions of object interactions encourage comprehenders to simulate the movements involved in those interactions. Movement features previously have been found to relate to event segmentation (Hard et al., 2006; Newtson et al., 1977), particularly for fine segmentation (Zacks, 2004). However, in one other study in which the current materials were presented one word at time (Speer, Reynolds, & Zacks, 2007), object changes were associated with both coarse and fine event boundaries, so it does not appear to be the case that changes in object interactions *only* affect segmentation at fine temporal grains.

Although situation changes influenced the pattern of event segmentation on their own, they also acted together to influence event segmentation. As the number of changes in a clause increased, the likelihood of perceiving an event boundary also increased. This result suggests that event boundaries are perceived when there is an unusual degree of mismatch between incoming information and the current event model, and mirrors claims that readers update their situation models when there is less overlap between consecutive pieces of information (Gernsbacher, 1990; Magliano et al., 1999; Zwaan et al., 1998). The fact that this relationship held for both the coarse and fine segmentation tasks indicates that the effect of the number of changes on event segmentation is not restricted to a particular perceptual timescale.

Experiment 2

Although Experiment 1 used descriptions of real-world events rather than actual real-world events, the influence of these situational variables on event segmentation may not be limited to descriptions. Temporal and spatial changes also influence event segmentation of film (Magliano et al., 2001), indicating that these conceptual aspects of narrated situations are abstract enough to modulate perception of live activity as well as activity presented through language. EST (Zacks et al., 2007) proposes that segmentation of narrative and live-action events reflects the operation of common memory updating processes. If so, the relations between situation changes and event segmentation observed in Experiment 1 should hold with live-action film as well as with text narratives. To test this hypothesis, we asked participants in Experiment 2 to segment an extended narrative film.

Experiment 2 also afforded the opportunity to examine the role of an important structural cue that may play a role in the segmentation of film. *Cuts* are those points at which two sequences of continuous film (*shots*) are edited together. Cuts produce full-field visual discontinuities. One might expect that such discontinuities would be highly salient and would produce the subjective experience of an event boundary. Previous studies using extended narrative films have shown that larger interruptions, namely commercial breaks, do affect comprehension and memory for film (Boltz, 1995; Boltz, 1992), and cuts have

been shown to affect memory for brief films (Carroll & Bever, 1976; Schwan, Garsoffky, & Hesse, 2000). On the other hand, if event segmentation depends on discontinuities in the situational features monitored by viewers, cuts by themselves may not be strongly related to event boundaries. Magliano and colleagues (2001) found that when a spatial or temporal change occurred, the shot that contained the change was more likely to be perceived as an event boundary. Many of these changes happened at cuts, which is consistent with the notion that cuts themselves tend to produce event boundaries. However, Magliano and colleagues did not compare intervals with cuts to those without cuts. Thus, it remains an open question whether cuts in and of themselves are associated with event segmentation.

Method

Participants.—Forty-one undergraduates (ages 18–22, 10 men) participated in Experiment 2 for course credit or a cash stipend. Five additional participants were excluded due to failure to complete most of the testing session (n = 2), problems understanding or following the instructions (n = 3), or having previously seen the stimulus movie (n = 1). An additional two participants were able to complete most but not all of the testing session; their partial datasets were analyzed.

Materials, Design, and Procedure.—Participants viewed the movie *The Red Balloon* (Lamorisse, 1956), a French film for children from the 1950s. *The Red Balloon* was selected because it has very little spoken language, contains many situation changes, and has virtually no discontinuous jumps in time. Narrative films often use cuts to present temporal sequences out of order (flashbacks and flash-forwards). By selecting a film that did not make use of these devices we hoped to maximize generalization to the perception of real activities. Other psychologists have found these features of *The Red Balloon* attractive; it was previously used in a comprehensive comparison of memory across narrative text and film (Baggett, 1979).

Participants were asked to identify event boundaries while watching the film by pressing a button box button whenever they judged that one meaningful unit of activity had ended and another had begun (Newtson, 1973). The film was digitized from VHS tape for the experiment. It was presented on a Macintosh computer (Apple, Cupertino CA) with a 19" monitor using PsyScope experimental presentation software (Cohen, MacWhinney, Flatt, & Provost, 1993). The film was presented in four clips, with lengths of 463.3, 468.4, 446.2, and 600.6 s.

Design and Procedure.—Each participant watched the movie twice, once to identify coarse boundaries and once to identify fine boundaries. Order of segmentation grain was counterbalanced across participants. After providing informed consent, participants were given instructions for the segmentation task. Those who did fine segmentation first were asked to identify the smallest units that were meaningful to them; those who did coarse segmentation first were asked to identify the largest units they found meaningful. They were asked to practice by segmenting a brief (155.4 s) movie of a man assembling a boat using Duplos construction blocks (Lego Group, www.lego.com). To reduce idiosyncratic variability in participants' interpretation of how long coarse and fine units should generally

be, we used a shaping procedure (Zacks, Speer, Vettel, & Jacoby, 2006). If a participant identified fewer than three coarse units or six fine units during the practice session, the experimenter repeated the instructions and asked the participant to repeat the practice segmentation. The practice also was repeated if the participant appeared unsure of the instructions after completing the practice segmentation task. Following the practice, the participant segmented *The Red Balloon*, with breaks between each clip. They then were given instructions for the second segmentation grain and the practice procedure was repeated for that grain, after which they segmented *The Red Balloon* again using the new grain.

Scoring.—Each clip in the film was divided into 5 s intervals (398 intervals total). (This interval length was chosen based on pilot data, such that the probability of segmentation in each unit at each grain was similar to those observed in Experiment 1, and to maximize the variability in the predictor variables of interest.) An interval was coded as an event boundary if the participant had identified at least one boundary during that interval. Two research assistants coded the film frame by frame for situation changes. *Spatial changes* were coded when a character changed direction of motion within the scene or when the point of view of the camera changed location. Changes in direction of motion included starting, stopping, and sudden turning. New camera points of view generally occurred when the camera position before a cut was different from the position after a cut, but also could occur when the camera tracked around a corner or through a door. *Temporal changes* were coded whenever the frame after a cut was temporally discontinuous with the frame preceding the cut. Whereas the texts used in Experiment 1 contained no temporal changes but only temporal references, this film did contain a small number of temporal changes (27, or 6.8% of intervals). When a temporal change occurred there was always also a spatial change. Therefore, temporal changes could be used in the analysis of the number of changes, but not in the analysis of the effects of individual changes. *Object changes* occurred whenever the nature of a character's interaction with an object changed. Object changes included picking up an object, putting down an object, and beginning to use an object already in hand in a new way. *Character changes* occurred whenever the focus of action was an animate character or characters and this focus was different than in the preceding frame. For example, a cut from one boy running to a pack of boys chasing him was coded as a character change. (It is worth noting that this film includes non-human animate characters, including dogs and animate balloons.) In addition to changes in characters themselves, we also coded changes in *character interactions*. These were defined as changes in the physical or abstract interactions between characters, such as touching, talking or gesturing, or joining together while walking or running. Changes in character interactions are often not mentioned explicitly in a text (and so were not coded in Experiment 1) but are a salient feature of narrative film. *Causal changes* were coded whenever the activity in a frame could not be described as having been caused by something viewed in the previous frame. Finally, *goal changes* were coded whenever a character performed an action associated with a goal different than the goal in the previous frame. In sequences in which a series of shots cut back and forth between characters with different goals, this was coded as a series of goal changes. In addition to the situation changes, we coded for *cuts*, which are those points in time when two continuous film shots are edited together. The situation changes and the presence of cuts were used as predictors in models of segmentation.

Results

As in Experiment 1, the initial analyses compared the size of the units identified in the coarse and fine segmentation tasks, as well as the degree of overlap between the locations of the coarse and fine unit boundaries. Logistic regressions for each participant then determined if and how the situation change variables predicted patterns of segmentation.

Alignment Effects.—The total number of seconds in a movie was divided by the number of event boundaries identified in that movie to give a measure of the mean length (in seconds) of the units identified in the coarse and fine segmentation tasks for each participant. Participants identified longer units during the coarse segmentation task ($M = 51.53$ s, $SEM = 5.26$) than during the fine segmentation task ($M = 13.76$ s, $SEM = 1.11$), as instructed. The number of 5-s intervals each participant identified as coarse boundaries, fine boundaries, or both was used to assess the degree to which the coarse and fine boundaries aligned with one another. As in Experiment 1, expected overlap values were computed for each participant, and compared to the observed overlap values to control for the amount of overlap expected by chance alone. A paired sample t-test indicated that the mean number of bins on which coarse and fine boundaries aligned ($M = 34.95$, $SEM = 4.26$) was significantly larger than would be expected on the basis of chance ($M = 24.672$, $SEM = 3.74$), $t(40) = 9.3$, $p < .001$.

Predictors of Event Structure – Number of Changes.—As for Experiment 1, we carried out logistic regressions predicting the pattern of segmentation for each individual as a function of the number of situation changes that occurred in each 5-s interval. There were 121 intervals with no changes, 76 with one change, 85 with two changes, 56 with three changes, 42 with four changes, and 18 with five changes. As can be seen in Figure 5, the probability of segmentation increased with increasing numbers of changes for both coarse and fine segmentation. For coarse boundaries the mean logistic regression coefficient was .30 (SD = .15), $t(40) = 12.7$, $p < .001$. For fine boundaries the mean coefficient was .39 (SD = .15), $t(40) = 16.4$, $p < .001$. The mean for fine boundaries was significantly greater than that for coarse boundaries, $t(40) = -3.26$, $p = .002$.

Predictors of Event Structure – Type of Change.—Logistic regressions were conducted for each participant to assess the relations between the situation change variables and event segmentation. Cuts also were included in the models. (As noted previously, temporal changes were not included because they were rare and were a strict subset of spatial changes.) Descriptive statistics for and correlations between the predictor variables are shown in Table 5. As for Experiment 1, alpha was set at .001 due to the large number of correlations considered. Temporal, spatial, and character changes tended to occur at cuts (though this was not necessarily the case), leading to significant correlations for these three variables. As for Experiment 1, we computed the condition number for the matrix of predictor variables to check for multicollinearity. The condition index was 4.86, well within the acceptable range. As for Experiment 1, the predictor variables were centered around their means. Odds ratios giving the multiplicative change in the odds of identifying a boundary given the presence of a situation change are shown in Figure 6.

The odds ratios indicate that, as in Experiment 1, individual situation change variables reliably predicted patterns of segmentation in both the coarse and fine tasks. This was true for all of the change variables during fine segmentation and for all but goal changes during coarse segmentation. The most predictive changes were those in characters and objects; the least predictive were changes in goals. Overall, situation changes were associated with slightly larger effects on fine segmentation than coarse segmentation. These differences among the situation change variables, and between coarse and fine segmentation, were assessed using a 2×6 repeated measures ANOVA, with segmentation grain and situation change as the independent variables. The main effects of both grain and feature were significant. $F(1, 40) = 7.59$, $p = .008$; $F(5, 200) = 34.9$, $p < .001$. The two-way interaction was not significant, $F(5, 200) = 1.49$, $p = .19$. That is, situation changes produced larger increases in fine segmentation than in coarse segmentation, but this did not vary significantly across the different changes.

Hierarchical regressions were conducted to assess the amount of variability in segmentation accounted for by the situation changes. The analyses paralleled those of Experiment 1: Each participant's coarse and fine segmentation were modeled separately. Cuts were entered first, followed by either the number of situation changes or the situation change variables. These were compared to the null hypothesis using Monte Carlo simulation, as was done for Experiment 1. Table 6 shows that both the segmentation change variables and the number of changes accounted for significant increases in variability accounted for above that accounted for by cuts, for both coarse and fine segmentation. In sum, larger numbers of situation changes were associated with greater likelihood of identifying an event boundary, and when all the situation changes were entered individually they accounted for 9 to 11% of the variance in segmentation.

Cuts accounted for a significant amount of variability in fine but not coarse segmentation in the hierarchical regression analyses. Inspection of the regression coefficients revealed that in the first stage models (with only cuts as predictors), cuts were associated with increases in the probability of segmentation; these increases were significant only for fine segmentation [Coarse: odds ratio from mean regression coefficient = 1.09, $t(40) = .12$; Fine: odds ratio = 1.40, $t(40) = 4.98$, $p < .001$]. However, once situation changes were included in the model the direction of the effect reversed: Cuts were now associated with decreases in the probability of segmentation; these were significant for both coarse and fine segmentation [Coarse: odds ratio = .70, $t(40) = -5.75$, $p < .001$; Fine: odds ratio = .85, $t(40) = -2.1$, $p = .04$]. This suggests that cuts, by themselves, do not cause people to perceive event boundaries. They are weakly associated with event boundaries because they are the points in time at which changes in the depicted situation tend to occur. If anything, cuts in and of themselves appear to *decrease* the probability that a viewer will perceive an event boundary.

Discussion

The main findings of Experiment 2 converged with those of Experiment 1: Changes in the situation depicted by the films were associated with increases in event segmentation. This was true for both coarse and fine segmentation. Increases cumulated: The more situational features changed, the larger the probability that a viewer would identify an

event boundary. Further, there was some consistency in which situation changes were associated with increases in segmentation. Changes in objects and characters both were strongly associated with increases in segmentation, similar to Experiment 1. Changes in spatial location also were associated with segmentation increases, which replicates the pattern observed previously by Magliano and colleagues (2001).

In Experiment 2, situation changes were stronger predictors of fine segmentation than of coarse segmentation. This result differs from Experiment 1. One possibility is that this difference reflects the different sources of information available to language comprehenders and film comprehenders. Film provides a number of physical visual and auditory cues that are not available from narrative language, including quantitative information about the movement of actors and objects, changes in object contact relations, facial expressions, and environmental sounds. In particular, movement features have been found to be powerful cues to event understanding (Bassili, 1976; Heider & Simmel, 1944) and to segmentation —particularly to segmentation at finer temporal grains (Hard et al., 2006; Zacks, 2004; Zacks, Swallow, Vettel, & McAvoy, 2006). If film viewers' fine segmentation depends preferentially on physical features such as object and actor motion, which were not available for analysis in the present study, this dependence may reduce the strength of relation between situation changes and fine segmentation, relative to coarse segmentation. However, another possibility is that the difference between Experiments 1 and 2 reflects incidental differences between the stimuli or between how the segmentation grain instructions were interpreted by participants. In future work it would be valuable to compare segmentation of film and text stimuli matched for information content. Such matching has been instructive in analyses of memory for narrative texts, picture stories, and movies (Baggett, 1979; Gernsbacher, 1985).

The fact that cuts were not by themselves associated with increases in event segmentation may seem surprising at first blush. However, this result actually is consistent with film-theoretic accounts (e.g., Anderson, 1996). Simple cuts are often intended by filmmakers to be unobtrusive, and film editing traditions are quite successful at making this work. When viewers are explicitly instructed to detect cuts, miss rates as high as 32% have been reported for some cut types (Smith & Henderson, in press). *The Red Balloon* is edited in a naturalistic style consistent with these traditions, with minimal use of more obtrusive editing techniques. Edits such as fades and dissolves are much more noticeable to viewers, and are used by filmmakers to indicate things such as the passage of time or a change in location. Also, film theory distinguishes between different types of cuts (Zacks & Magliano, in press). Across cuts, the action may switch to a new scene or maintain scene continuity . Within a scene it may maintain continuity of action or not. These distinctions are related to the situation change coding used here but do not correspond to it exactly. Two important questions for future research are whether more obtrusive edits lead to the perception of event boundaries, and whether the formal categories used by film theorists to classify cuts are predictive of event segmentation.

In sum, Experiments 1 and 2 both suggest that viewers tend to segment ongoing activity into events at points when features in a narrated situation are changing. This is consistent with the Event Indexing Model (Zwaan, 1999), which specifically predicts that set of

situational dimensions studied here are monitored by comprehenders. These results extend the model by showing it applies to narrative film, not just on the scale of the shot (Magliano et al., 2001) but within shots. These results also are consistent with EST (Zacks et al., 2007), which proposes that situation changes render an activity less predictable, and this leads to the perception of an event boundary. They extend EST by applying its predictions to the comprehension of printed and spoken discourse.

Discourse theories and EST make predictions not just about the *causes* of event segmentation, but also about its *consequences*. Experiment 3 investigated consequences for one variable that is important for text comprehension: reading rate.

Experiment 3

The results of Experiments 1 and 2 indicated that situation changes contribute to comprehenders' explicit perception of narrative structure. Some theories of narrative comprehension predict that the mechanisms that produced these explicit event boundary judgments also should lead to slowing in reading. Both the Event Indexing Model (Zwaan, 1999) and the Structure Building Framework (Gernsbacher, 1990), propose that readers are more likely to update working memory representations when situational features in a narrative change; this memory updating takes time, leading to a transient slowing in reading. EST (Zacks et al., 2007) makes a similar prediction for similar reasons. EST proposes that when situational changes occur, failures in prediction tend to result. This transient prediction failure triggers the updating of event models, which would be expected to slow reading. Therefore, both theories of reading and theories of event segmentation argue that reading times should increase when situation changes occur.

A pair of previous studies have measured reading times for excerpts from short fiction stories that contained changes in a number of dimensions, such as time, space, and causality (Zwaan et al., 1998; Zwaan, Magliano, & Graesser, 1995c). Spatial changes did influence reading times, but only when participants had memorized a map of the narrative location. Such results suggest that comprehenders *can* track spatial location when it is made salient, but may not do so under normal reading circumstances. Therefore, it was important to replicate the results of these previous studies using the narratives from *One Boy's Day* (Barker & Wright, 1951), whose narrated situations were substantially different from the fiction stories used in previous studies.

Narrative comprehension theories predict that the probability of updating one's situation model increases as the amount of situational discontinuity increases. Similarly, EST suggests that larger numbers of situational changes should produce larger prediction failures, leading to a greater likelihood of triggering an event boundary. Therefore, these theories predict that on average, reading times should be positively correlated with the number of situation changes in clause. Previous studies have found some support for this claim (Magliano et al., 1999; Zwaan et al., 1998). Experiment 3 provided a further test of this hypothesis.

Thus, discourse comprehension theories predict that reading times should be correlated with the event boundary judgments of Experiment 1, because they both reflect the same

processing of situational changes and consequent memory updating. EST makes two strong predictions. First, reading times for event boundaries should be slower than for non-boundaries. Second, this relationship should be mediated by the presence of situation changes. Experiment 3 tested these predictions.

Method

Participants.—Thirty-two undergraduates (ages 18–25, 13 men) participated in Experiment 3 for course credit or a cash stipend. Seven additional participants were excluded due to failure to follow the task instructions (n = 1), overall reading times more than two standard deviations above the group mean (n = 3), or less than 50% accuracy on the comprehension test (n = 3).

Materials, Design, and Procedure.—The materials were identical to those used for the visual-sequential group in Experiment 1. The narratives were presented one clause at a time on the computer monitor, and participants were asked simply to read each clause as it appeared on the screen, and press a button on the keyboard to advance through the clauses. The time between clause onset time and the button press served as the dependent reading time measure. The order of the narratives was counterbalanced across participants. Participants were told to read the stories for comprehension, as they would receive a comprehension test at the end of each narrative. Each comprehension test consisted of five cued recall questions, for a total of 20 questions throughout the experiment. (Example: "On whose bike did Raymond hang his jacket.") Participants made relatively few incorrect responses on the comprehension test ($M = 3.34$ of 5, $SEM = 0.33$), indicating that they were adequately comprehending the narratives. The comprehension questions for each narrative can be found online at http://dcl.wustl.edu/DCL/stimuli.html.

Results

The analysis procedure for Experiment 3 closely followed the analysis procedure for Experiments 1 and 2. However, rather than using logistic regression to predict the probability of identifying a given clause as coarse or fine event boundary, Experiment 3 used linear regression to predict reading times for each clause. These linear regressions predicted reading time based on terminal and non-terminal punctuation, originating narrative, and the situation change variables[i]. To determine the amount of variability in reading times that perceived event boundaries accounted for above and beyond the influence of the structural and situation change variables, the number of individuals that identified each clause as a coarse event boundary in the visual-continuous group in Experiment 1 was also included in

---

[i]Number of syllables, although highly related to reading times, was not included in these models. There were moderate positive zero-order relationship between number of syllables and the situation changes (largest $r = 0.17$), and in conjunction with the strong relationship between number of syllables and reading times (largest $r = .75$), these positive relationships suppressed the relationship between individual situation changes and reading times. Suppression leads to artificially inflated coefficients for one or more predictor variables, and can lead to negative coefficients when all of the zero-order relationships between independent and dependent variables are positive (for an expanded discussion of this issue, see (Cohen & Cohen, 1983). Indeed, when number of syllables was included as a predictor in these models, the coefficients for many of the situation changes, which had positive zero-order correlations with reading time, became negative. To avoid these artificial relationships between the situation changes and reading time, number of syllables was excluded from the regression models. However, the analyses were recomputed using only clauses with six syllables (the modal number of syllables) as a way to control for differences in reading times. The situation change variables were still predictive of reading times, even in this matched-syllable analysis, indicating that they are related to reading times, above and beyond the influence of the number of syllables.

a hierarchical regression[ii]. Segmentation data from the visual-continuous group was used as the predictor rather than segmentation data from the visual-sequential group because these data were the more reliable of the two data sets due to participants' ability to precisely indicate perceived event boundaries.

Clauses with reading times less than 500 ms were excluded from the analyses. The remaining reading times were trimmed on an individual basis, such that for each participant, trials with overall reading times more than two standard deviations above the individual participant's mean were excluded from the analyses. On average, this procedure resulted in the exclusion of less than 5% of each participant's data ($M = 4.73\%$, $SEM = .10$).

Predictors of Reading Time – Number of Changes.—A first series of regressions predicted reading times for each individual using the number of changes in a given clause and the structural variables. An increase in the number of changes in a given clause was associated with slower reading times while controlling for the effects of the structural variables; the mean estimated reading time increase for each additional situation change was 92.8 ms (SEM 4.12 ms), $t(31) = 22.53$, $p < .001$. The mean increase in R^2 due to the number of changes in a clause was .02 (*95% CI* for null distribution = .001 to .002), and the mean additional increase in R^2 due to the coarse boundaries was .01 (*95% CI* for null distribution = .001 to .002). (We note that these R^2 values refer to variance accounted for in reading times for *individual clauses*, which is not typically reported. Many variables contribute to reading time, so the overall proportion of variance is small though the effects are systematic and reliable.)

Predictors of Reading Time – Type of Change.—A second series of regressions predicted reading times using the individual situation changes, coded using a set of zero-one variables, rather than the number of changes. Figure 7 shows the mean coefficients for each type of change. T-tests were conducted to determine whether the coefficients for each situation change variable differed significantly from zero. Changes in causal relationships, characters, their goals, and interactions with objects all were associated with increased reading times, smallest $t(31) = 5.15$, $p < .001$. Changes in characters' spatial locations were associated with faster reading times, $t(31) = -4.64$, $p < .001$, and temporal references failed to have a statistically significant association with reading times, $t(31) = -1.66$, $p = .106$.

For each participant, we computed the correlation between reading times and the probability of a coarse event boundary from the visual-sequential group in Experiment 1. Coarse event boundaries were coded as the percentage of participants in Experiment 1 who had identified a coarse boundary during each clause. Across participants the average correlation coefficient was .19 (*95% CI* for null distribution = −.013 to .014). To determine whether event boundaries predicted reading times above and beyond the situation changes, we fit a set of hierarchical regressions predicting reading time for each participant based on the structural variables (first step) and the coarse event boundaries (second step). The mean R^2

[ii]Including fine unit boundaries in the models led to an unacceptable degree of multicollinearity in the models (condition index = 54.13), even when coarse unit boundaries were not included (condition index = 52.21). Including only coarse unit boundaries resulted in an acceptable condition index (13.68).

for the structural variables across participants was .08 (*95% CI* for null distribution = .012 to .026). Adding coarse segmentation to the structural variables increased R^2 by .023 (95% CI for the null distribution = .001 to .002). The mean effect of the event boundary variable was 10.8 ms (SEM = .45 ms). In other words, for each 10% increase in segmentation probability, reading times increased by 108 ms.

To assess whether the relationship between event boundaries and reading time was mediated by situation changes, we fit a set of hierarchical regressions in which the structural variables were entered first, followed by the situation change variables. The mean increase for the situation change variables was .04 (*95% CI* for null distribution = .007 to .010). We then added the coarse event boundaries in a third regression step; the mean increase in R^2 for the coarse event boundaries after having added the situation change variables was .01, which though small was statistically significant (95% CI for null distribution = .001 to .002). The mean effect of the event boundary variable in this model was 7.0 ms (SEM = .50 ms), which was significantly smaller than the mean effect of 10.8 ms from the model that did not include the situation change variables, *t*(31) = 15.5, *p* < .001. In short, event boundaries were strongly related to reading time, but a substantial component of this effect could be accounted for by situation changes.

Discussion

Overall, the results of Experiment 3 support the hypothesis that situation changes lead to transient increases in cognitive processing. Many of the situation changes were associated with slower reading times, and situation changes had a cumulative effect such that reading times became increasingly slower as the number of changes in a clause increased. The number of situation changes in a clause had a strong linear relationship with reading time. This result replicates previous results (Magliano et al., 1999; Zwaan et al., 1998), and is consistent with the proposal that readers are increasingly likely to update their situation models as incoming information becomes less consistent with the information in the readers' current situation model (Gernsbacher, 1990; Zwaan, 1999).

An important novel finding was that reading tended to slow down for those clauses that corresponded to coarse event boundaries. Moreover, this relationship was substantially reduced once the situation changes were taken into account. This suggests that the relationship between the event boundaries and reading time was mediated by the situation changes. Such a pattern provides strong support for EST's claim that event models are updated when features of the situation change, and that this updating is perceived as an event boundary.

The particular situation changes that were associated with slower reading times are meaningful: Changes in the causal relations in the narrative, changes in characters, their interactions with objects, and goals were all associated with slower reading times. These results extend the results of previous studies by including shifts in characters and changes in characters' interactions with objects as relevant dimensions to be coded in narratives. Previous studies have defined character changes as points when a new character is introduced into the story world, rather than points where the story shifts focus from one character to another. The results of Experiment 3 suggest that changing the subject of a

sentence is sufficient to increase reading times, and does not require the introduction of an entirely new character. This finding is consistent with the idea that breaks in *referential coherence* lead readers to update their situation models during narrative comprehension (Gernsbacher, 1990).

In contrast to causal, object, character, and goal changes, temporal references failed to have a significant relationship with reading times, and changes in characters' spatial locations were associated with *faster* reading times. The lack of influence of temporal references is not too surprising, as the narratives used here did not have discontinuous changes in narrative time. *Changes* in narrative time (e.g., "The next day…"), as opposed to references to the passage of small time intervals, have previously been associated with increased reading times (Anderson, Garrod, & Sanford, 1983; Zwaan, 1996).

The negative relation between change in characters' spatial locations and reading times is somewhat counterintuitive. However, it should not be too surprising, as previous studies of the effects of spatial changes on reading time have produced mixed but generally negative results (Zwaan et al., 1998; Rinck & Weber, 2003; Scott Rich & Taylor, 2000). It may be that readers determined that tracking location did not facilitate comprehension, and therefore skimmed clauses containing spatial references. Spatial changes were frequently repetitive within the narratives. For example, one narrative involved an extended chase between Raymond and his friend, Susan, and described them repeatedly running up and down a hill in the schoolyard. Readers may have learned quickly to ignore these repetitive changes. This conscious inattention to spatial changes could have led to the observed facilitation in reading times for clauses containing spatial changes. At this point it remains an open question what determines whether a reader will track spatial changes during reading (Scott Rich & Taylor, 2000; Zwaan et al., 1998).

Experiment 4

Experiment 1 demonstrated that situation changes increase the probability of identifying a clause as an event boundary during comprehension, and Experiment 3 demonstrated that situation changes influence reading times. Theories of text comprehension and event perception suggest that situation changes lead readers to update their situation models, and that this updating process leads to slower reading times and to the explicit perception of event boundaries. EST proposes that this updating occurs as a response to transient increases in prediction error (Zacks et al., 2007). When a new event begins, features of the situation may change unpredictably, producing transient increases in error in one's predictions about what will happen next in the narrative. It would be adaptive to update situation models when this happens, and prediction error may provide a signal that can be used without awareness to control memory updating. Computational simulations support this hypothesis for simple biological motion events (Reynolds et al., 2007). In short, EST proposes that the moment-to-moment predictability of an activity is coupled to the perception of event boundaries and the control of cognition during comprehension

In Experiment 4, we set out to test three hypotheses: First, that clauses corresponding to event boundaries would tend to be judged less predictable than other clauses; second,

that clauses corresponding to event boundaries would tend to be associated with slowing in reading; and third, that these two effects would be mediated by increases in situation changes. To test these hypotheses, we had readers rate each of the clauses used in Experiments 1 and 3 for perceived predictability. We then compared these predictability ratings to segmentation data from Experiment 1, and to reading rate data from Experiment 3.

Method

Participants.—Thirty-two undergraduates (ages 18–21, 11 men) participated in Experiment 4 for course credit or a cash stipend. Seven additional participants were excluded due to failure to follow the task instructions (n = 2), computer error (n = 1), less than 50% accuracy on the comprehension test (n = 2), or incomplete data sets (n = 2).

Materials, Design, and Procedure.—The narratives and the presentation method were similar to those used for the visual-sequential group in Experiment 1 and for Experiment 3. A 7-point rating scale with values ranging from 1 ("Completely unpredictable") to 7 ("Completely Predictable") was presented throughout each narrative at the bottom of the computer monitor. The narratives were presented one clause at a time on the computer monitor, and participants were simply asked to read each clause as it appeared on the screen, and to press the number on the keyboard to rate the predictability of the information in the clause, given the previous information they had encountered in the story. The order of the narratives was counterbalanced across participants. Participants were told to read the stories for comprehension, as they would receive a comprehension test at the end of each narrative (identical to the test used in Experiment 3). Participants made relatively few incorrect responses on the comprehension test (M = 4.84 of 5, SEM = 0.41), indicating that they were adequately comprehending the narratives.

Results

The initial analyses for Experiment 4 closely followed the analysis procedure for Experiment 3. Linear regressions predicted ratings based on terminal and non-terminal punctuation, originating narrative, number of syllables, and the situation change variables. An additional set of analyses determined whether the situation change variables accounted for variability in Experiments 1 and 3 above and beyond the mean predictability rating for each clause.

Clauses with response times less than 500 ms were excluded from the analyses. The remaining response times were trimmed on an individual basis, such that for each participant, trials with overall response times more than two standard deviations above the individual participant's mean were excluded from the analyses. On average, this procedure resulted in the exclusion of less than 5% of each participant's data (M = 4.09%, SEM = .21).

Predictors of Predictability – Number of Changes.—As the number of changes increased, the perceived predictability decreased. The mean estimated decrease in predictability for each additional situation change was .07 (SEM .014) $t(32) = -4.88$, $p < .001$. The overall increase in R^2 due to the number of changes in a clause was small but statistically significant (M = .005, *95% CI* for null distribution = .001 to .002).

Predictors of Predictability – Type of Change.—Mean coefficients for each of the situation change variables are shown in Figure 8. Changes in the causal relationships between clauses were associated with reduced predictability ratings, $t(32) = -10.95$, $p < .001$, whereas changes in characters were associated with increased predictability ratings, $t(32) = 7.24$, $p < .001$. The remaining situation changes were not significantly associated with predictability ratings, largest $t(32) = -1.12$, $p = .273$. The mean R^2 was .06 for the structural variables (*95% CI* for null distribution = .019 to .032), and .04 for the situation change variables (*95% CI* for null distribution = .007 to .010).

Relations of Predictability Ratings to Event Boundaries and Reading Times.—
To ask whether predictability was related to event segmentation and to reading time, we used the mean predictability for each clause to predict the segmentation data from Experiment 1 and the reading time data from Experiment 3. For each participant in Experiments 1 and 3, mean predictability ratings from the current experiment were correlated with coarse and fine event boundaries and reading times. These correlation coefficients were then averaged across individuals in each experimental condition, and comparisons to Monte Carlo-simulated null hypothesis distributions determined whether more participants showed statistically significant correlations than would be expected by chance. The results are shown in Table 7. As predictability decreased, participants were more likely to identify event boundaries and reading time increased.

Situation Changes as Mediators of Relations Between Predictability, Segmentation, and Reading Time.—We conducted hierarchical regressions to assess whether situation changes mediated the relation between predictability and event boundaries (using data from Experiment 1), and between predictability and reading time (using data from Experiment 3). The event boundary effects were estimated using logistic regression; the reading time effects were estimated using linear regression. We first estimated the relation between predictability and event segmentation or reading time using models that included only the predictability ratings (first column of Table 8). We then estimated the same relation using models that included the structural variables (second column of Table 8). Finally, we added the situation change variables (third column of Table 8). As can be seen in the table, the first-level coefficients relating predictability to event boundaries were all negative. They became less negative or crossed zero and became positive when the structural variables were included, and did so further when the situation change variables were included. In all cases but one the increases were significant. Negative coefficients mean that as predictability increased segmentation decreased (see Table 7), so an increase in the coefficients indicates that this tendency was reduced. Thus, to the extent that predictability was related to event segmentation, this relation was partially mediated by situation changes and the structural variables. As can be seen in the last row of Table 8, the strong relation between predictability and reading time was reduced once the structural variables were included in the model, but was not further affected by adding the situation change variables in the next step. Thus, there was no evidence that the relation between predictability and reading time was mediated by unique variance in situation changes.

Discussion

Participants' ratings of the predictability of a clause were reliably related to event segmentation. As predictability decreased, the probability of identifying an event boundary increased. This relation was reduced once the effects of structural variables were taken into account, and further reduced when the unique effects of the situation changes were taken into account. According to EST (Zacks et al., 2007), this effect occurs because when information from the environment is no longer consistent with the contents of the current event model, increased prediction error leads to the updating of memory and the subjective experience that a new event has begun. Thus, this result provides important support for EST.

This finding also supports narrative comprehension theories that propose readers update memory representations when a text becomes less coherent (Zwaan, 1999; Gernsbacher, 1990; van Dijk & Kintsch, 1983). Although neither coherence nor memory were directly measured here, it is likely that self-judged predictability is correlated with coherence (Graesser & Wiemer-Hastings, 1999; van den Broek, Linzie, Fletcher, & Marsolek, 2000), and there is evidence that self-reported event boundaries are associated with memory updating (Speer & Zacks, 2005).

Predictability ratings also were reliably related to reading time. Clauses that were less predictable were read more slowly, replicating previous results (Langston & Trabasso, 1999; McKoon & Ratcliff, 1992; Rayner, Ashby, Pollatsek, & Reichle, 2004). This finding is consistent with EST's proposal that readers engage in specialized processing at those points in time when prediction error transiently increases.

The relation between predictability and reading time was partially mediated by the structural features of the text, similar to the relation between predictability and segmentation. However, we did not find evidence that the relation between predictability and reading time was mediated by situation changes, above and beyond any mediation by structural variables. This result is difficult to reconcile with theoretical accounts that propose the effect of situation changes on reading rate is mediated by prediction error, including EST. However, one alternative possibility is that predictability does in general mediate relations between situation changes and reading rate, but that in naturalistic materials the relevant situation changes covary with structural variables. As noted previously, in these stimuli situation changes are correlated with the structural variables, particularly clause length. In future research it will be valuable to experimentally manipulate the presence of situation changes independent of length and examine the relations between those changes, predictability, and reading rate.

Some aspects of the data raise questions about the validity of subjective ratings of predictability in this context. As the number of situation changes in a clause increased, perceived predictability decreased; this effect was as predicted. However, this effect was small. Only one type of situation change, changes in causes, was uniquely associated with reductions in predictability. Moreover, changes in characters were associated with small but significant *increases* in predictability ratings, which was unexpected. We know of no theory that would predict this result. One possibility is that this reflects a limitation in participants' ability to retrospect about how predictable a clause was, having just read it.

Regarding character changes in particular, many of these changes occur when one character finishes doing something and another responds or starts doing something else. For example, consider this sequence: "Raymond walked off slowly in the direction of the terrace. Susan deliberately grabbed the jacket from the handle bars of Thelma's bike." After reading the second sentence (which contains a character change), participants may have had the sense that they had predicted that *something* new was going to be done, probably by a different character. This may have provided the subjective impression that the sentence was predictable from the preceding. But it is unlikely that readers actually could have predicted that it was Susan who would act next, or that she would grab Raymond's jacket. In future research it will be important to compare subjective ratings of predictability with objective tests of prediction accuracy.

In sum, the results of Experiment 4 show that event boundaries tend to occur when activity is rated less predictable, and replicate previous findings that reading slows at these points. In future research it will be important to extend these results for self-rated predictability to objective measures of prediction performance.

General Discussion

This series of studies set out to investigate how situational cues affect event segmentation. The results of these four studies suggest that a number of conceptual cues influence event perception, and this influence extends to implicit measures of reading comprehension. In addition, changes in these conceptual situational features influenced the perceived predictability of the activities described in the narratives. Taken together, the results of these four studies support those theories of event perception and reading comprehension which suggest that comprehenders build models of the current event or situation, and update those models when incoming information no longer matches the contents of the current model (Gernsbacher, 1990; Zacks et al., 2007; Zwaan & Radvansky, 1998). These data further refine these theories by identifying some of the variables that lead readers and listeners to perceive these changes, and update their event models during comprehension.

For both text and film narratives, comprehenders were able to identify meaningful units of activity in the narratives, and these patterns mirrored those found when individuals are asked to identify meaningful units of activity in films of everyday events (Newtson, 1973; Zacks et al., 2001b). Comprehenders in Experiments 1 and 2 were able to modulate their grain of segmentation, identifying both large and small units of activity in the narratives, and the locations of these fine and coarse boundaries aligned to a greater degree than would be expected by chance.

Comprehenders Segment Activities at Points of Change

In both reading and movie viewing, event boundaries became more likely as the number of dimensions that changed in the situation increased. This relationship was true for both coarse and fine segmentation. This finding gives support to the idea that comprehenders constantly monitor a number of different dimensions for changes. The fact that readers processed clauses with more situation changes more slowly in Experiment 3 further supports this conclusion. These effects of situation changes on reading times have been observed

in other studies that have used more standard fiction stories and different change variables (Zwaan et al., 1995a; Zwaan et al., 1998). These results suggest that this finding is robust, and provides strong support for models that take into account the dissimilarity between incoming information and information in the current situation/event model (Gernsbacher, 1990; Zacks et al., 2007; Zwaan & Radvansky, 1998).

The variables that predicted patterns of segmentation in Experiment 1 and 2 were quite consistent, whether narratives were presented as visual text, spoken text, or film. Changes in the causal relationship of narrated activities, changes in characters, their goals, and locations were all associated with an increased likelihood of identifying a clause as an event boundary, regardless of the medium of presentation and mostly regardless of the grain of segmentation. These situation change variables predicted patterns of segmentation above and beyond the influence of structural variables such as punctuation (in the case of the visually presented narratives), pause duration (in the case of the auditory narratives), and the presence of a cut (in the case of the narrative film).

The fact that event segmentation was strongly related to situation changes in both narrative texts and films suggests an important conclusion about the role of physical cues in event segmentation. Previous research has provided evidence that physical cues such as actor and object movement are significant predictors of event segmentation (Newtson et al., 1977; Hard et al., 2006; Zacks, 2004). It is likely that in viewing real activity or movies, physical movements are correlated with conceptual features such as causes, characters, goals, and locations. The results of Experiment 1 demonstrate that physical changes are not necessary to produce robust effects on event segmentation: In reading and listening to the stories, there were no physical cues in the stimuli to influence the perception of event boundaries, yet purely conceptual changes were robustly related to segmentation. Thus, the data from Experiment 1 provide strong support for the claim that conceptual cues are sufficient to guide event perception.

However, Experiment 1 demonstrated that not all of the conceptual cues exert equal influences on event perception, and the exact cues that influence segmentation vary depending on task context. Characters' interactions with objects predicted only fine grain segmentation, indicating that there is a difference in the cues readers and listeners attend to when looking for large and small units of activity. Comprehenders may implicitly track a number of different cues in the process of understanding events, but task demands and context may focus their attention on a subset of these cues. The fact that reading times, an online measure of the effect of these changes on comprehension, were influenced by changes in characters' interactions with objects suggests that comprehenders may constantly process these changes, but only attend to them when the task demands highlight small units of activity.

The features proposed by the Event Indexing Model (Zwaan et al., 1995a; Zwaan et al., 1998) were correlated with event segmentation; this result provides support for the model's proposal that readers monitor this set of features. However, it is worth keeping in mind that what counts as a change depends on which features a comprehender is monitoring. For example, consider the motion of objects. The same trajectory might be described as

a change in position or as motion with a constant velocity; another trajectory might be described as a change in velocity or as a constant acceleration. In the case of motion, previous research suggests that people monitor several moments of the position variable simultaneously (Bassili, 1976; Zacks, 2004; Mann & Jepson, 2002; Rubin & Richards, 1985).

EST claims that event boundaries occur when predictions fail. Is prediction failure the only thing that produces an event boundary? An attractive possibility is that some event boundaries are detected in a purely top-down fashion based on one's currently activated event schemata. For example, let us suppose that most readers possess a schema corresponding to the morning routine (Rosen, Caplan, Sheesley, Rodriguez, & Grafman, 2003). When reading "Raymond went quickly, almost eagerly, into the kitchen," one might recognize based on one's schema that the kitchen is the typical site of eating breakfast, and therefore infer that a new event had begun. A reader might also have represented that one typically dresses just before eating breakfast (if this is how the reader usually does it), in which case the fact that the new event would be breakfast would be quite predictable. Such intuitive examples abound, and they suggest that one may identify event boundaries in a purely top-down fashion based on one's event schemata. However, there is another more parsimonious alternative. It may be that at moments such as this, though it is predictable that Raymond's next actions will comprise eating breakfast, there is much about the activity that is unpredictable at this point: What will he eat? Who is in the kitchen? Where will he sit? Thus, it may be that even in situations in which comprehenders have well-formed schemata that specify information about the temporal order of events, prediction becomes less reliable at event boundaries. The question of whether event schemata can produce psychological event boundaries in the absence of prediction failures is an important one for future research.

Reading, Situation Model Updating, and Prediction

The results of Experiment 3 indicate that event boundaries are related to reading time, but that this relationship is most likely mediated by the situation changes. EST (Zacks et al., 2007) predicts this relation, because event boundaries are thought to arise from prediction errors that occur when the current event model no longer accurately predicts incoming information. These errors lead to updating the comprehender's current event model. The fact that readers showed effects of the situation changes on reading times suggests that readers are updating their event models to reflect those changes.

In particular, the results of Experiment 4 suggest that readers update their event models because they detect errors in predictions they have made about what they will encounter in the text. Clauses that were identified as event boundaries in Experiment 1 were rated as less predictable than non-boundary clauses. This was partially mediated by the fact that as the number of situation changes in a clause increased, the judged predictability decreased. This pattern is consistent with EST's proposal that changes in a situation lead to errors in prediction, which lead to the perception that a new event has begun (Zacks et al., 2007). However, at this point it is not clear whether this was due to situation changes *per se*, to structural changes that co-occur with situation changes, or to both.

Individual-level Analyses

Although previous studies have tested the effects of variables on group segmentation and group-averaged reading times (Magliano et al., 2001; Speer & Zacks, 2005; Zacks, 2004; Zwaan et al., 1995c; Zwaan et al., 1998), the current studies captured the relationships between the situation change variables and individual patterns of segmentation and reading times. In this way, we were able to determine the effects of the situation change variables across individuals rather than across groups. Across individuals there is good agreement as to where event boundaries are, but this agreement is greater within individuals (Speer et al., 2003). We believe that individual-level analyses will be extremely useful in future research, particularly when testing hypotheses about individual differences in event perception and reading comprehension. For instance, the strength of the relationship between perception, reading time, and the situation change variables may relate to the comprehending ability of the reader. Because comprehension ability is thought to rely on general cognitive functions, the influence of the change variables on event perception and reading times should relate to more general measures of cognitive function, such as working memory, or general comprehension skill (e.g., Gernsbacher, 1990).

One important point to keep in mind in interpreting these data is that the target of the analyses was a single comprehender's segmentation judgment, reading time, or predictability rating for a single clause. This contrasts with the more typical procedure of averaging reading times or ratings across participants for linear regression analyses. Individual reading times and individual judgments can be affected by many sources of unsystematic variance, including moment-to-moment fluctuations in attention or mental set. Thus, although the effects of situation changes on individual segmentation judgments, reading times, and predictability judgments may appear relatively small, they were generally large relative to the total systematic variance in the data. It is instructive to compare these individual-level analyses to analyses of responses aggregated across participants. For example, in the visual-sequential condition of Experiment 1, the number of situation changes accounted for an additional 4.8% of variance in coarse segmentation and an additional 8.8% of variance in fine segmentation in the average individual-level hierarchical logistic regressions. In a set of hierarchical linear regressions predicting event boundary data summed across participants, the corresponding increases in variance were 12.8% and 25.3%.

The Control of Comprehension

The ability to segment ongoing activity—be it described in narrative, depicted in film, or directly observed—is an important component of comprehension. The current studies suggest that changes in conceptual features of a situation are monitored by perceivers and form a basis for segmenting activity into meaningful events, across presentation modality and medium. This is good methodological news for researchers who hope that studies of discourse processing can inform us about the comprehension of real life. It also is important news for researchers interested in how cognitive processes are controlled over time: In the domain of comprehension, at least, the processing of situational changes is an important predictor of the conscious experience of events and of the control of reading.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Abelson RP (1981). Psychological status of the script concept. American Psychologist, 36, 715–729.

Anderson A, Garrod SC, & Sanford AJ (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 35(3-A), 427–440.

Anderson J (1996). The reality of illusion: an ecological approach to cognitive film theory Carbondale: Southern Illinois University Press.

Baggett P (1979). Structurally equivalent stories in movie and text and the effect of the medium on recall. Journal of Verbal Learning & Verbal Behavior, 18(3), 333–356.

Baird JA & Baldwin DA (2001). Making sense of human behavior: Action parsing and intentional inference. In Malle BF, Moses LJ, & Baldwin DA (Eds.), Intentions and intentionality: Foundations of social cognition (pp. 193–206). Cambridge, MA: MIT Press.

Baldwin DA, Baird JA, Saylor MM, & Clark MA (2001). Infants parse dynamic action. Child Development, 72(3), 708–717. [PubMed: 11405577]

Barker RG & Wright HF (1951). One boy's day: A specimen record of behavior New York: Harper & Brothers.

Barker RG & Wright HF (1954). Midwest and its children: The psychological ecology of an American town Evanston, Illinois: Row, Peterson and Company.

Bassili JN (1976). Temporal and spatial contingencies in the perception of social events. Journal of Personality & Social Psychology, 33(6), 680–685.

Boltz MG (1995). Effects of event structure on retrospective duration judgments. Perception & Psychophysics, 57(7), 1080–1096. [PubMed: 8532497]

Boltz M (1992). Temporal accent structure and the remembering of filmed narratives. Journal of Experimental Psychology: Human Perception & Performance, 18(1), 90–105. [PubMed: 1532194]

Carroll JM & Bever TG (1976). Segmentation in cinema perception. Science, 191(4231), 1053–1055. [PubMed: 1251216]

Cohen J & Cohen P (1983). Applied multiple regression/correlation analysis for the behavioral sciences Hillsdale, N.J.: Nature Publishing Group.

Cohen JD, MacWhinney B, Flatt M, & Provost J (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. Behavior Research Methods, Instruments & Computers, 25(2), 257–271.

Copeland DE, Magliano JP, & Radvansky GA (2006). Situation models in comprehension, memory and augmented cognition. In Forsythe C, Bernard ML, & Goldsmith TE (Eds.), Cognitive systems: Human cognitive models in systems design (pp. 37–66). Mahwah, NJ: Erlbaum.

Dickman HR (1963). The perception of behavioral units. In Barker RG (Ed.), The stream of behavior (pp. 23–41). New York: Appleton-Century-Crofts.

Gernsbacher MA (1985). Surface information loss in comprehension. Cognitive Psychology, 17, 324–363. [PubMed: 25308975]

Gernsbacher MA (1990). Language comprehension as structure building Hillsdale: L. Erlbaum.

Glenberg AM, Meyer M, & Lindem K (1987). Mental models contribute to foregrounding during text comprehension. Journal of Memory & Language, 26(1), 69–83.

Graesser AC & Wiemer-Hastings K (1999). Situational models and concepts in story comprehension. In Goldman SR, Graesser AC, & et al. (Eds.), Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso (pp. 77–92). Mahwah, NJ, US: Lawrence Erlbaum Associates.

Grafman J (1995). Similarities and distinctions among current models of prefrontal cortical functions Structure and Functions of the Human Prefrontal Cortex. Annals of the New York Academy of Sciences, 769, 337–368. [PubMed: 8595036]

Hard BM, Tversky B, & Lang D (2006). Making sense of abstract events: Building event schemas. Memory & Cognition, 34(6), 1221–1235. [PubMed: 17225504]

Heider F & Simmel M (1944). An experimental study of apparent behavior. American Journal of Psychology, 57, 243–259.

Hommel B (2006). How we do what we want: a neuro-cognitive perspective on human action planning. In Jorna RJ, van Wezel W, & Meystel A (Eds.), Planning in intelligent systems: Aspects, motivations and methods (pp. 27–56). New York: John Wiley & Sons.

Lamorisse A (1956). The red balloon

Langston M & Trabasso T (1999). Modeling causal integration and availability of information during comprehension of narrative texts. In van Oostendorp H & Goldman SR (Eds.), The construction of mental representations during reading (pp. 29–69). Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc., Publishers.

Lattin JM, Carroll JD, & Green PE (2003). Analyzing Multivariate Data Pacific Grove, CA: Thomson Brooks/Cole.

Magliano JP, Miller J, & Zwaan RA (2001). Indexing space and time in film understanding. Applied Cognitive Psychology, 15(5), 533–545.

Magliano JP, Zwaan RA, & Graesser AC (1999). The role of situational continuity in narrative understanding. In van Oostendorp H & Goldman SR (Eds.), The construction of mental representations during reading (pp. 219–245). Mahwah, N. J.: Lawrence Erlbaum Associates.

Mann R & Jepson AD (2002). Detection and classification of motion boundaries. Paper presented at the Eighteenth National Conference on Artificial Intelligence, Edmonton, Canada.

McKoon G & Ratcliff R (1992). Inference during reading. Psychological Review, 99(3), 440–466. [PubMed: 1502273]

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

Nagelkerke NJD (1991). A note on a general definition of the coefficient of determination. Biometrika, 78(3), 691–692.

Newtson D (1973). Attribution and the unit of perception of ongoing behavior. Journal of Personality and Social Psychology, 28(1), 28–38.

Newtson D (1976). Foundations of attribution: The perception of ongoing behavior. In Harvey JH, Ickes WJ, & Kidd RF (Eds.), New directions in attribution research (pp. 223–248). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Newtson D, Engquist G, & Bois J (1977). The objective basis of behavior units. Journal of Personality and Social Psychology, 35(12), 847–862.

Pawley A & Lane J (1998). From event sequence to grammar: serial verb constructions in Kalam. In Siewierska A & Song JJ (Eds.), Case, Typology, and Grammar (pp. 201–228). Philadelphia: John Benjamins.

Pustejovsky J (1991). The syntax of event structure. Cognition, 41(1–3), 47–81. [PubMed: 1790655]

Radvansky GA & Copeland DE (2006). Walking through doorways causes forgetting: situation models and experienced space. Memory & Cognition, 34(5), 1150–1156. [PubMed: 17128613]

Radvansky GA, Zwaan RA, Federico T, & Franklin N (1998). Retrieval from temporally organized situation models. Journal of Experimental Psychology: Learning, Memory, & Cognition, 24(5), 1224–1237.

Rayner K, Ashby J, Pollatsek A, & Reichle ED (2004). The effects of frequency and predictability on eye fixations in reading: implications for the E-Z Reader model. Journal of Experimental Psychology: Human Perception & Performance, 30(4), 720–732. [PubMed: 15301620]

Reynolds JR, Zacks JM, & Braver TS (2007). A computational model of event segmentation from perceptual prediction. Cognitive Science, 31, 613–643. [PubMed: 21635310]

Rinck M & Bower G (2000). Temporal and spatial distance in situation models. Memory and Cognition, 28(8), 1310–1320. [PubMed: 11219959]

Rinck M & Weber U (2003). Who when where: An experimental test of the event-indexing model. Memory and Cognition, 31(8), 1284–1292. [PubMed: 15058689]

Rosen VM, Caplan L, Sheesley L, Rodriguez R, & Grafman J (2003). An examination of daily activities and their scripts across the adult lifespan. Behavioral Research Methods & Computers, 35(1), 32–48.

Rubin JM & Richards WA (1985). Boundaries of visual motion. AIM-835

Saylor MM, Baldwin DA, Baird JA, & LaBounty J, (2007). Infants' on-line segmentation of dynamic human action. Journal of Cognition and Development, 8(1), 113–128.

Schwan S, Garsoffky B, & Hesse FW (2000). Do film cuts facilitate the perceptual and cognitive organization of activity sequences? Memory & Cognition, 28(2), 214–223. [PubMed: 10790977]

Scott Rich S & Taylor HA (2000). Not all narrative shifts function equally. Memory & Cognition, 28(7), 1257–1266. [PubMed: 11126946]

Smith TJ & Henderson JM (in press). Edit blindness: The relationship between attention and global change in dynamic scenes. Journal of Eye Movement Research

Speer NK, Reynolds JR, & Zacks JM (2007). Human brain activity time-locked to narrative event boundaries. Psychological Science, 18(5), 449–455. [PubMed: 17576286]

Speer NK, Swallow KM, & Zacks JM (2003). Activation of human motion processing areas during event perception. Cognitive, Affective & Behavioral Neuroscience, 3(4), 335–345.

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

van den Broek P, Linzie B, Fletcher C, & Marsolek CJ (2000). The role of causal discourse structure in narrative writing. Memory & Cognition, 28(5), 711–721. [PubMed: 10983444]

van Dijk TA & Kintsch W (1983). Strategies of discourse comprehension New York: Academic Press.

Zacks JM (2004). Using movement and intentions to understand simple events. Cognitive Science, 28(6), 979–1008.

Zacks JM, Braver TS, Sheridan MA, Donaldson DI, Snyder AZ, Ollinger JM et al. (2001a). Human brain activity time-locked to perceptual event boundaries. Nature Neuroscience, 4(6), 651–655. [PubMed: 11369948]

Zacks JM & Magliano JP (in press). Film understanding and cognitive neuroscience. Art and the Senses New York: Oxford University Press.

Zacks JM, Speer NK, Swallow KM, Braver TS, & Reynolds JR (2007). Event perception: A mind/brain perspective. Psychological Bulletin, 133(2), 273–293. [PubMed: 17338600]

Zacks JM, Speer NK, Vettel JM, & Jacoby LL (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. Psychology & Aging, 21(3), 466–482. [PubMed: 16953710]

Zacks JM, Swallow KM, Vettel JM, & McAvoy MP (2006). Visual movement and the neural correlates of event perception. Brain Research, 1076(1), 150–162. [PubMed: 16473338]

Zacks JM & Tversky B (2001). Event structure in perception and conception. Psychological Bulletin, 127(1), 3–21. [PubMed: 11271755]

Zacks JM, Tversky B, & Iyer G (2001b). Perceiving, remembering, and communicating structure in events. Journal of Experimental Psychology: General, 130(1), 29–58. [PubMed: 11293458]

Zwaan RA (1996). Processing narrative time shifts. Journal of Experimental Psychology: Learning, Memory, & Cognition, 22(5), 1196–1207.

Zwaan RA (1999). Five dimensions of narrative comprehension: The event-indexing model. In Goldman SR, Graesser AC, & van den Broek P (Eds.), Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso (pp. 93–110). Mahwah, NJ, US: Lawrence Erlbaum Associates.

Zwaan RA, Langston MC, & Graesser AC (1995a). The construction of situation models in narrative comprehension: An event-indexing model. Psychological Science, 6(5), 292–297.

Zwaan RA, Magliano JP, & Graesser AC (1995c). Dimensions of situation model construction in narrative comprehension. Journal of Experimental Psychology: Learning, Memory, & Cognition, 21(2), 386–397.

Zwaan RA & Radvansky GA (1998). Situation models in language comprehension and memory. Psychological Bulletin, 123(2), 162–185. [PubMed: 9522683]

Zwaan RA, Radvansky GA, Hilliard AE, & Curiel JM (1998). Constructing multidimensional situation models during reading. Scientific Studies of Reading, 2(3), 199–220.

Clause	Cause	Character	Goal	Object	Space	Time
...[Mrs. Birch] went through the front door into the kitchen.	●				●	
Mr. Birch came in	●	●			●	
and, after a friendly greeting,	●					●
chatted with her for a minute or so.	●					●
Mrs. Birch needed to awaken Raymond.		●				
Mrs. Birch stepped into Raymond's bedroom,			●		●	
pulled a light cord hanging from the center of the room,				●		
and turned to the bed.						
Mrs. Birch said with pleasant casualness, "Raymond, wake up."						

Figure 1.
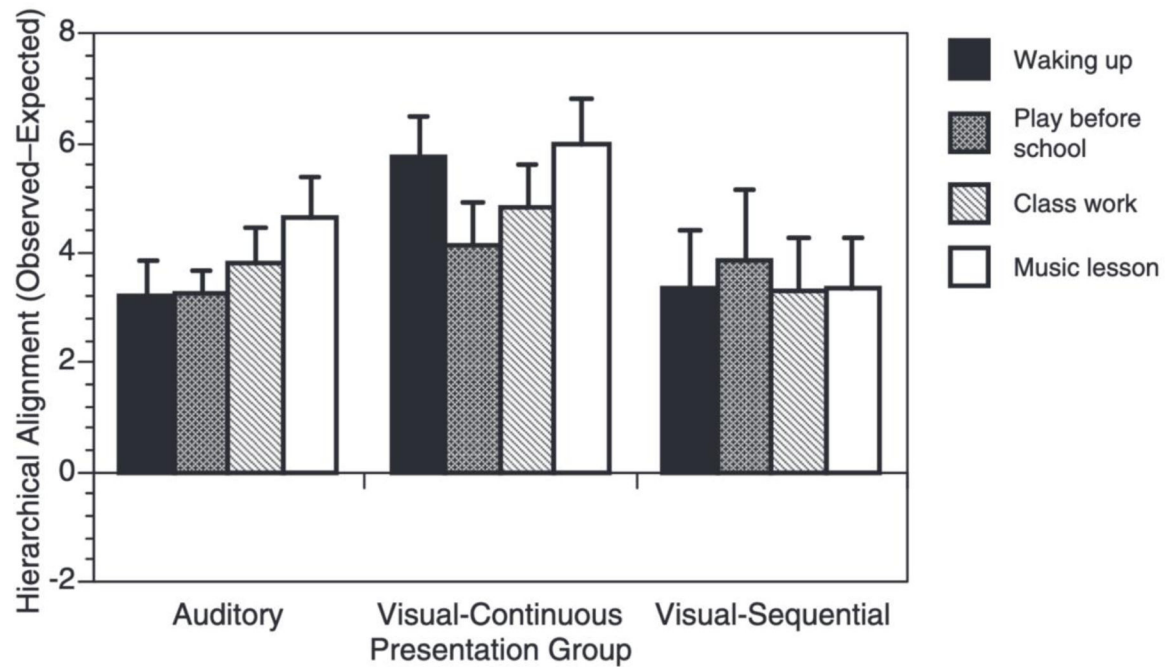Examples of changes in the narrative coding.

Figure 2.
As the number of situation changes during the stories increased, readers' probability of coarse and fine segmentation increased. Estimated probabilities were calculated by fitting logistic regression models for each participant, constructing a model using the mean regression coefficients across participants, and predicting each data point from that model. (In the models used for this figure, effects of story were not included because there are not natural "mean" story coefficients.)
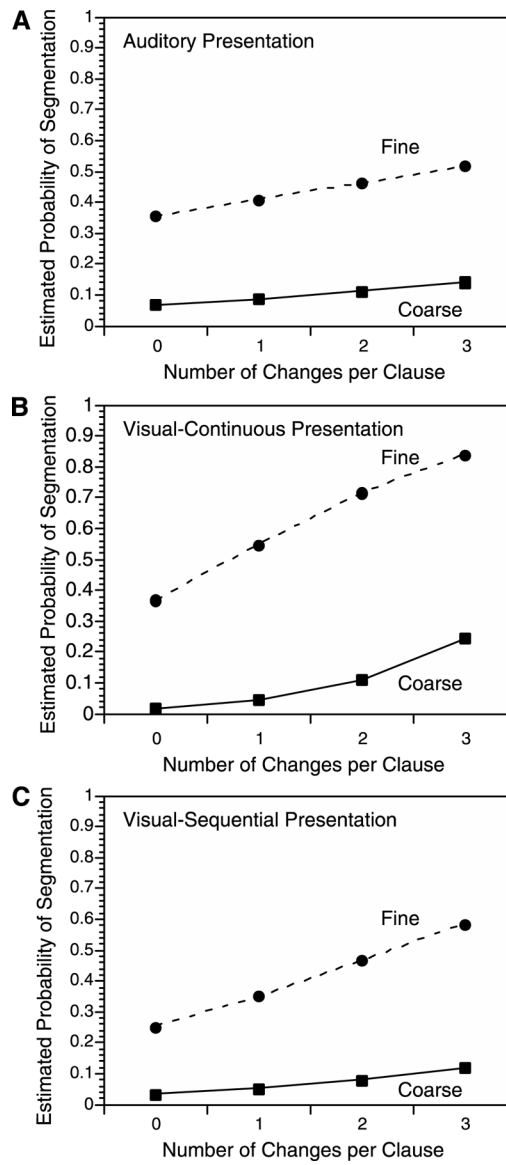
Figure 3.
Observers' coarse and fine unit boundaries for narrative texts showed a greater degree of alignment than would be expected by chance.
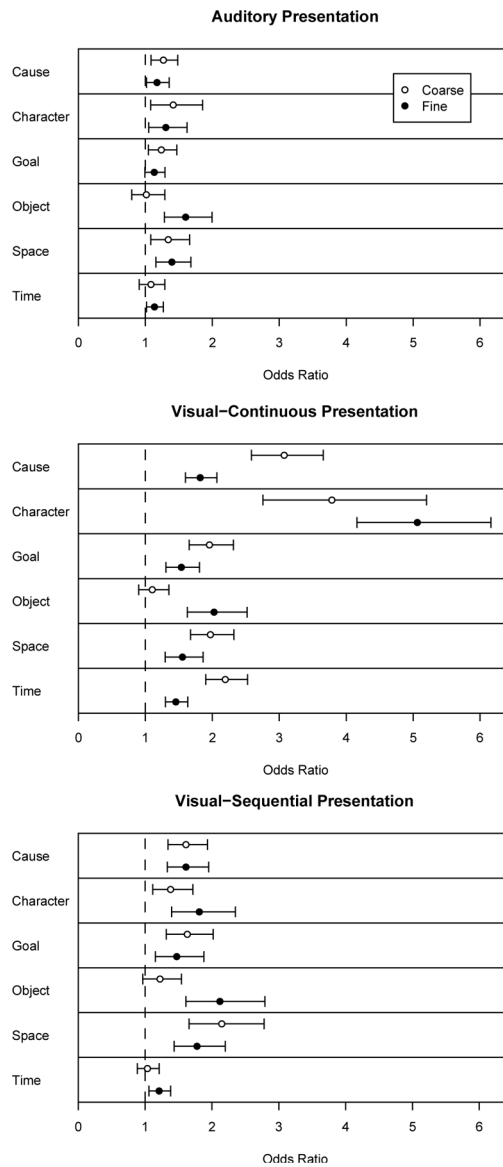
Figure 4.
The relations between situation changes and segmentation in narrative text. Circles represent the odds ratios for the mean coefficients in the logistic regressions, and error bars the 95% confidence intervals for these odds ratios. Values whose confidence intervals do not include 1.0 (dashed lines) indicate a statistically significant relationship between changes on that dimension and segmentation.

Figure 5.
As the number of situation changes during the movie increased, viewers' probability of coarse and fine segmentation increased. Estimated probabilities were calculated by fitting logistic regression models for each participant, constructing a model using the mean regression coefficients across participants, and predicting each data point from that model.
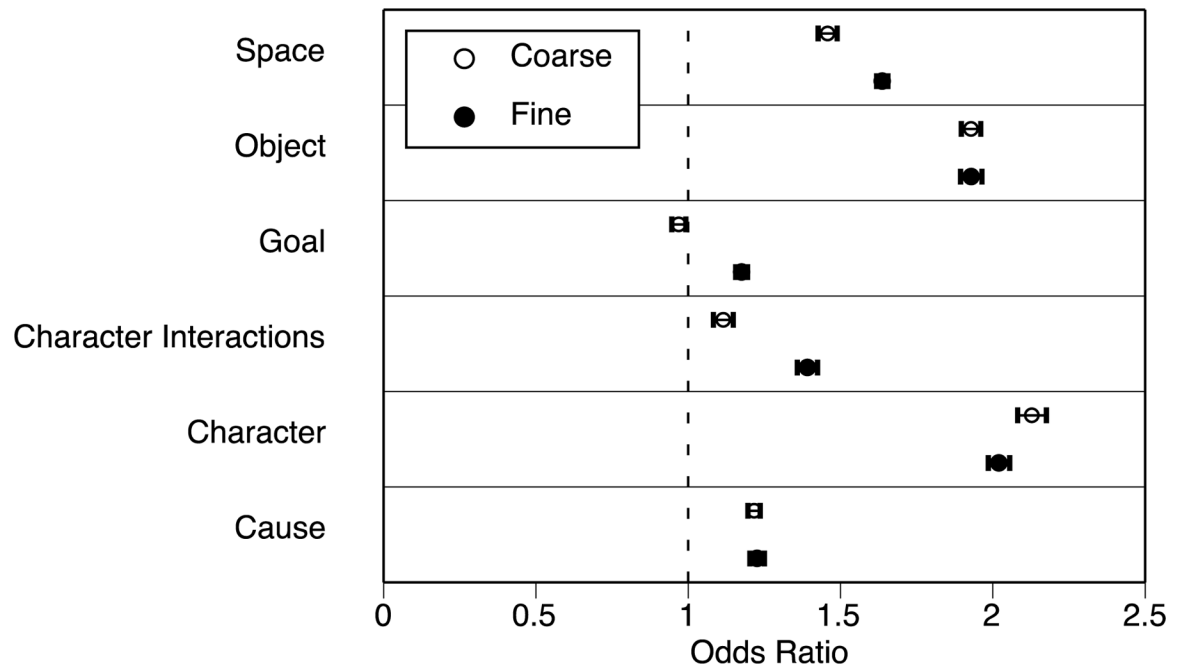
Figure 6.
The relations between situation changes and segmentation in narrative film. Circles represent the odds ratios for the mean coefficients in the logistic regressions, and error bars represent the 95% confidence intervals for these odds ratios. Values whose confidence intervals do not include 1.0 (dashed lines) indicate a statistically significant relationship between changes on that dimension and segmentation.
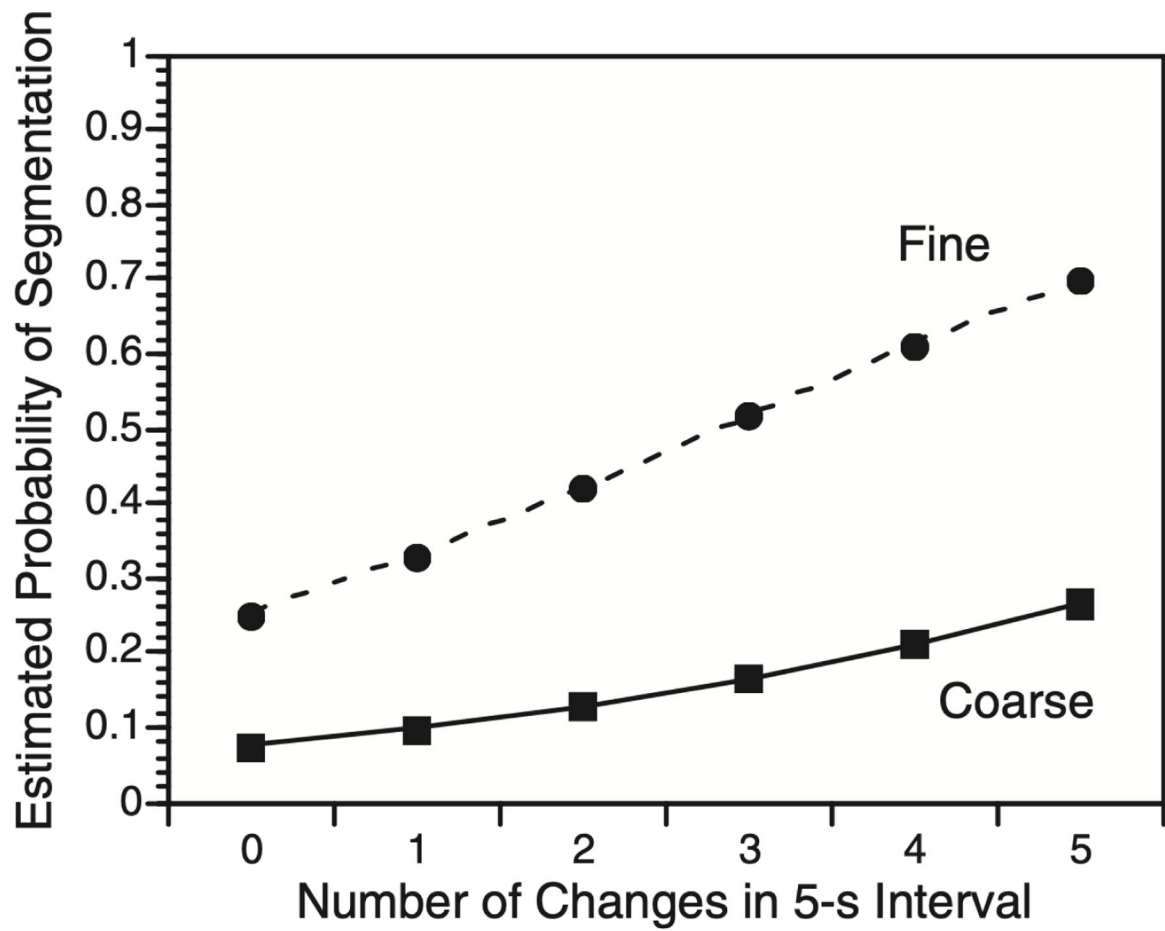
Figure 7.
The change in clause reading times associated with each of the situation changes (error bars represent 95% confidence intervals).
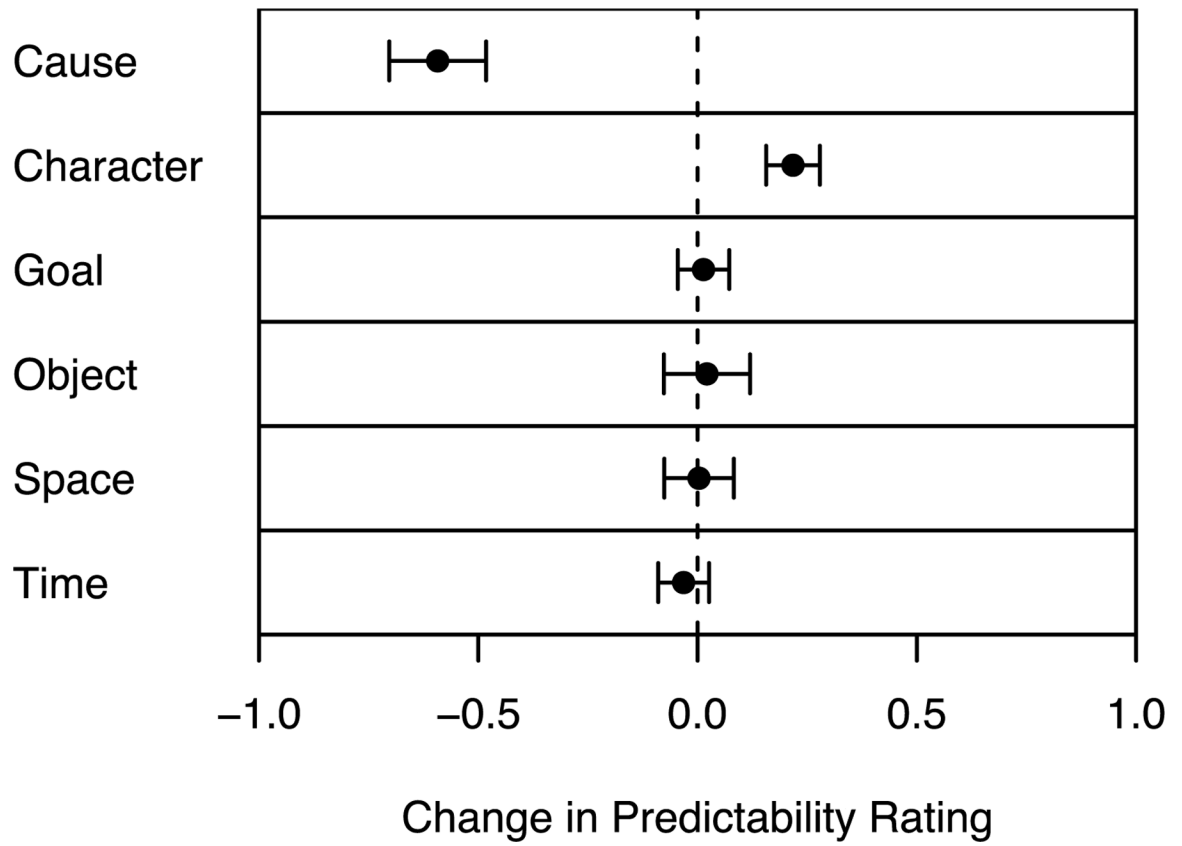
Figure 8.
The change in clause predictability ratings associated with each of the situation changes
(error bars represent 95% confidence intervals).

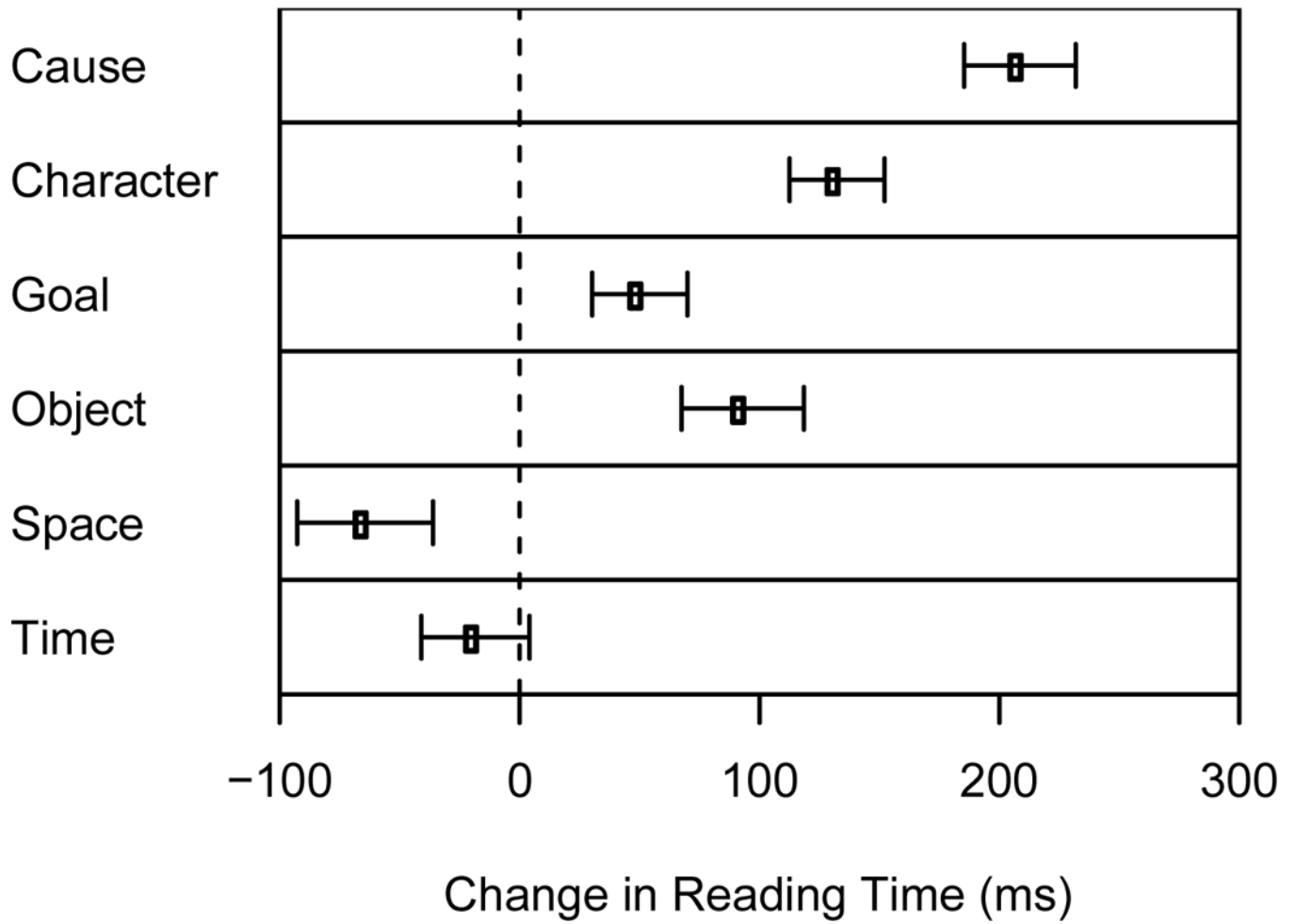**Table 1:**

Mean event lengths in clauses in Experiment 1. (Standard deviations in parentheses.)

		Auditory	Visual-Continuous	Visual-Sequential
Waking up	Fine	3.56 (2.42)	2.48 (1.54)	4.15 (3.12)
	Coarse	14.52 (10.19	15.11 (10.48)	20.24 (14.36)
Play before school	Fine	3.08 (1.82)	2.39 (1.32)	3.59 (2.70)
	Coarse	12.16 (9.29)	16.52 (17.02)	17.70 (11.22)
Class work	Fine	2.90 (1.60)	2.39 (1.39	3.44 (2.42)
	Coarse	11.75 (11.69)	12.14 (7.64)	18.28 (13.05)
Music lesson	Fine	3.18 (1.89)	2.50 (1.40)	3.75 (2.60)
	Coarse	12.24 (9.31)	14.58 (8.96)	24.12 (26.23)

Table 2.

Coefficients for number of changes in Experiment 1

Group	Coarse		Fine	
	M	*SEM*	*M*	*SEM*
Auditory	.27	.06	.23	.06
Visual-Continuous	.97	.03	.76	.05
Visual-Sequential	.46	.06	.48	.08

NOTE: Coefficients are means from subject-level logistic regressions predicting the presence of an event boundary from the number of situation changes in a clause (as well as from the structural variables). Positive coefficients indicate that as the number of changes in a clause increased, so did the probability of segmentation.

Table 3.

Descriptive Statistics for and Correlations Between Predictor Variables in Experiments 1, 3, and 4

Variable	1	2	3	4	5	6	7	8	9	10	11	*M*	*SEM*
1. Number of Syllables	---											9.46	4.28
2. Clause Duration	.92	---										2.18	.04
3. Pause Duration	.13	.11	---									.91	.03
4. Non-Terminal Punctuation	.19	.31	−.14	---								.31	.02
5. Terminal Punctuation	.22	.19	.50	−.33	---							.57	.02
6. Temporal References	.00	−.04	−.05	.09	−.06	---						.18	.01
7. Spatial Changes	−.07	−.04	−.04	−.11	−.03	.00	---					.12	.01
8. Object Changes	.02	−.02	.05	−.09	−.04	−.05	−.12	---				.11	.01
9. Character Changes	.17	.19	−.09	.10	−.13	.00	.15	−.04	---			.30	.02
10. Causal Changes	.16	.21	−.07	.04	−.10	.00	.15	.13	.17	---		.22	.01
11. Goal Changes	.17	.20	−.08	.12	−.09	.05	.12	.00	.38	.34	---	.22	.02

NOTE: With n = 757, $|t| > .12$ indicates a statistically significant correlation with $p < .001$. Mean clause and pause durations are in seconds.

The two rightmost columns give the mean and standard error for the structural variables and situation changes. For the situation changes, these correspond to proportions.

Correlations between the structural variables in the auditory group and the structural variables in the visual groups are shown for descriptive purposes only; the two sets of variables were never used in the same regressions, and the high correlations between the structural variables in these groups therefore have no effect on the following analyses.

Table 4.

Increment in Nagelkerke's R^2 for Predictor Variables Experiment 1

	Structural Variables			
	Coarse		Fine	
Group	*M*	*95% CI*	*M*	*95% CI*
Auditory	.093	.013 - .021	.175	.017-.035
Visual-Continuous	.097	.017 - .026	.107	.020 - .042
Visual-Sequential	.098	.016 - .029	.183	.019 - .033

	Number of Changes Variable			
	Coarse		Fine	
	M	*95% CI*	*M*	*95% CI*
Auditory	.034	.002 - .005	.041	.001 - .003
Visual-Continuous	.150	.002 - .005	.128	.001 - .003
Visual-Sequential	.048	.002 - .006	.088	.001 - .003

	Situation Change Variables			
	Coarse		Fine	
	M	*95% CI*	*M*	*95% CI*
Auditory	.064	.014 - .023	.060	.009 - .014
Visual-Continuous	.212	.015 - .025	.174	.010 - .017
Visual-Sequential	.081	.018 - .031	.113	.009 - .015

M = sample mean, 95% CI = 95% confidence interval for the null hypothesis, determined by Monte Carlo simulation.

Table 5.

Descriptive Statistics for and Correlations Between Predictor Variables in Experiment 2

Variable	1	2	3	4	5	6	7	*M*	*SEM*
1. Cuts	---							0.43	0.50
2. Temporal Changes	0.31	---						0.07	0.25
3. Spatial Changes	0.44	0.26	---					0.53	0.50
4. Object Changes	−0.11	−0.04	0.06	---				0.12	0.32
5. Character Changes	0.44	0.21	0.37	0.04	---			0.42	0.49
6. Character Interaction Changes	−0.05	0.01	0.21	0.18	0.15	---		0.21	0.41
7. Causal Changes	0.07	0.09	0.24	0.04	0.20	0.29	---	0.17	0.38
8. Goal Changes	0.06	0.01	0.31	0.07	0.20	0.51	0.31	0.24	0.43

Note that with n = 757, $|r| > .17$ indicates a statistically significant correlation with $p < .001$.

The two rightmost columns give the mean and standard error for the cuts variable and the situation changes. These correspond to proportions.

Table 6.

Increment in Nagelkerke's R^2 for Predictor Variables Experiment 2

Model	Coarse		Fine	
	M	*95% CI*	*M*	*95% CI*
Cuts	.006	.003 - .008	.017[*]	.002 - .006
Number of Changes Variable (increase of model with only cuts)	. 053[*]	.003 - .008	.090[*]	.002 - .005
Situation Change Variables (increase of model with only cuts)	.090[*]	.027-.041	.116[*]	.018 - .026

M = sample mean, 95% CI = 95% confidence interval for the null hypothesis, determined by Monte Carlo simulation.

[*] Differed significantly from the value expected by chance.

NOTE: Values reported to 3 decimal places to represent significant differences.

Table 7.

Correlations for Event Boundaries and Reading Time with Predictability

Coarse Boundaries	*M*	*r* 95% CI
Auditory	**−.043**	−.017 - .008
Visual-Continuous	**−.138**	−.014 - .011
Visual-Sequential	**−.067**	−.016 - .012
Fine Boundaries		
Auditory	**−.021**	−.016 - .010
Visual-Continuous	**−.062**	−.010 - .014
Visual-Sequential	**−.026**	−.017 - .010
Reading Time	**−.276**	−.018 - .010

*
95% CI = 95% confidence interval for the null hypothesis, determined by Monte Carlo simulation. Values in **bold** print represent statistically significant correlations.

Table 8.

Coefficients Relating Predictability to Event Boundaries and Reading Time

	Simple Effect	With Structural Variables	With Structural Variables and Situation Changes
Coarse Boundaries			
Auditory	−.16 (.04)	−.06 (.04)[*]	−.01 (.04)[*]
Visual-Continuous	−.49 (.03)	−.35 (,04)[*]	−.27 (.04)[*]
Visual-Sequential	−.29 (.04)	−.23 (.03)[*]	−.14 (.04)[*]
Fine Boundaries			
Auditory	−.05 (.02)	.03 (.02)[*]	,06 (,02)[*]
Visual-Continuous	−.15 (.03)	.00 (.02)[*]	.03 (.03)[*]
Visual-Sequential	−.07 (.04)	−.02 (.03)	.03 (−.03)[*]
Reading Time	−206.1 ms (8.4 ms)	−123.5 ms (7.0 ms)[*]	−124.9 ms (7.2 ms)

NOTE: Values for coarse and fine boundaries are logistic regression coefficients (log odds ratios); coefficients for reading time are linear regression coefficients. (Values in parentheses are standard errors across participants.) The models reflected in the first column included only predictability; the models reflected in the second column included the structural variables and predictability; the models reflected in the third column also included the situation change variables.

[*] The coefficient differed significantly from the one to its left, $p < .01$.