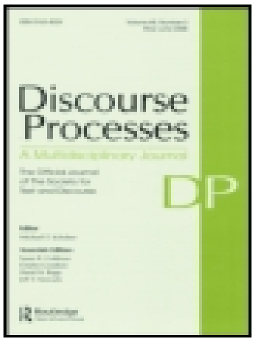
Events shape long-term memory for story information

Article *in* Discourse Processes · March 2023

DOI: 10.1080/0163853X.2023.2185408



Discourse Processes

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hdsp20

Events shape long-term memory for story information

Maverick E. Smith, Christopher A. Kurby & Heather R. Bailey

To cite this article: Maverick E. Smith, Christopher A. Kurby & Heather R. Bailey (2023): Events shape long-term memory for story information, Discourse Processes, DOI: 10.1080/0163853X.2023.2185408

To link to this article: https://doi.org/10.1080/0163853X.2023.2185408

DISCOURSE PROCESSES
https://doi.org/10.1080/0163853X.2023.2185408

Events shape long-term memory for story information

Maverick E. Smith[a,b], Christopher A. Kurby[c], and Heather R. Bailey[b]

[a]Department of Psychological & Brain Sciences, Washington University in St. Louis; [b]Department of Psychological Sciences, Kansas State University; [c]Department of Psychology, Grand Valley State University

ABSTRACT

We segment what we read into meaningful events, each separated by a discrete boundary. How does event segmentation during encoding relate to the structure of story information in long-term memory? To evaluate this question, participants read stories of fictional historical events and then engaged in a postreading verb arrangement task. In this task, participants saw verbs from each of the events placed randomly on a computer screen and then arranged the verbs into groups onscreen based on their understanding of the story. Participants who successfully comprehended the story placed verbs from the same event closer to each other than verbs from different events, even after controlling for orthographic, text-based, semantic, and situational overlap between verbs. Thus, how people structure story information into separate events during online comprehension is associated with how that information is stored in memory. Specifically, story information within an event is bound together in memory more so than information between events.

Introduction

When recalling the details of a novel, readers may organize their description into separate episodes using the story's events (Davis & Campbell, 2022; Davis et al., 2020; Ezzyat & Davachi, 2011; Kurby & Zacks, 2022), shared situational features between events (Zwaan, Langston et al., 1995), their knowledge and scripts for the actions that make up typical events (Bower et al., 1979; Trabasso & Van Den Broek, 1985; Van den Broek & Gustafson, 1999), and proximity of the information in the text (Graesser et al., 1994; Kintsch, 1988; Kintsch & van Dijk, 1978). For instance, if remembering the details from *Charlie and the Chocolate Factory* (Dahl, 1964), readers may recall the event of Charlie Bucket finding a golden ticket in a chocolate bar, granting him the opportunity to tour Willy Wonka's factory. They may recall each independent episode of the four other children on the tour giving into their impulses, misbehaving, and ultimately getting ejected from the factory in a darkly comedic fashion. The reader may also recall the event of Willy Wonka making Charlie his heir.

Although stories describe a large host of situational information about individual objects, actions, spatial settings, emotions, and so on, readers tend to remember stories as events that integrate across these individual pieces of information. For example, multiple events in *Charlie and the Chocolate Factory* contain Veruca Salt. Readers may recall the event of her winning a golden ticket by describing how workers in her father's factory were forced to unwrap chocolate bars until they found a ticket. Readers may also recall how Veruca got ejected from Willy Wonka's factory (e.g., she got flushed down the garbage chute). Despite being in separate events (e.g., winning a ticket, getting ejected from the factory), these are likely connected in memory through shared situational features, such as involving

CONTACT Maverick E. Smith mavericks@wustl.edu Department of Psychological & Brain Sciences, Washington University in St. Louis, 1 Brookings Drive, Campus Box 1125, St. Louis, MO 63130
Results of the current study were previously reported in a poster presented at the 60th Annual Meeting of the Psychonomic Society, Montreal, Canada.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/0163853X.2023.2185408

the same character. In addition, episodic memories for the individual actions within each event are likely bound together by virtue of sharing the same overarching event, even if those actions do not share all the situational features. For example, consider the "Veruca got ejected from the factory" event. The event begins with Veruca going under the gate that guards the squirrel's workplace (i.e., a change in location). The squirrels attack her (i.e., a change in goal from stealing to fleeing) and flush her down the garbage chute. Although these actions occur through situational changes in the story, they are likely associated in memory because they are from the same overall event.

Using a unique verb arrangement task, we explored how two important factors––shared situational properties (e.g., characters, causation etc.; Zwaan, Langston et al., 1995) and event structure (Davis & Campbell, 2022; Davis et al., 2020; Ezzyat & Davachi, 2011; Kurby & Zacks, 2022)—organize long-term memory for story information. Both have previously been shown to carry unique variance in performance on a variety of different tasks (Bailey & Zacks, 2015; Kurby & Zacks, 2012; Magliano et al., 2012; Speer et al., 2007). We expected that that both would contribute to the organization of long-term memory for story information. However, we suggest that the process of event segmentation contributes uniquely to the organization of episodic memories.

To successfully comprehend a story, readers construct a mental model, called a *situation model*, of the situation described in the text. Situation models serve as a mental simulation that includes information from the text itself, the events described in the text, the reader's background knowledge, and the inferences that readers generate while reading (McNamara & Magliano, 2009). Successful comprehension depends on the construction of a coherent situation model that connects information within events, even if the information within an event is described in sentences that are far apart in terms of the text-based structure of the text.

Event indexing model

Influential theories of situation models (Gernsbacher, 1990, 1997; Kintsch, 1998), such as the Event Indexing Model (Zwaan, Langston et al., 1995; Zwaan & Radvansky, 1998), claim that readers actively monitor 5 indices of the situation as they read stories and that these situational indices organize long-term memory for story information. These situational indices include the characters, their spatiotemporal location, their goals, and the causal relationships between events (Rinck & Weber, 2003; Zwaan, Langston et al., 1995; Zwaan, Magliano et al., 1995; Zwaan & Radvansky, 1998), although other indices, such as emotion, have also been proposed (Gernsbacher et al., 1992; Komeda & Kusumi, 2006; Therriault & Rinck, 2007). Readers update their situation model when there are changes along these indices, for example, when there is a shift in time, space, or a causal break.

Updating of the situation model during online comprehension is reflected in the availability of information immediately after a change in one or more of the situational indices. For instance, responses to recognition memory probes shown while reading take more time to execute and they are less accurate after a change in the situation (Bailey & Zacks, 2015; Pettijohn & Radvansky, 2016a; Radvansky & Copeland, 2010; Rinck & Bower, 2000; Zwaan, 1996). Presumably, information in the current situation model is more highly activated than information represented in prior situation models. This has been interpreted to suggest that readers update their situation models when they experience a change while they are reading.

Changes in the situation also shape how situation models are structured in long-term memory. Specifically, story events that share situational indices are more highly associated in long-term memory than story events that do not share indices. Zwaan, Langston et al. (1995) developed a method to measure the organization of memory for story information as a function of situational overlap (see also, Zwaan & Brown, 1996). After reading a text, participants saw a set of verbs, representing the story events, taken from the text. Participants sorted the verbs into groups that they thought "belonged together" based on their memory of the text. Zwaan, Langston et al. (1995) found that participants were more likely to sort verbs into the same group when they came from clauses that shared the same characters, time frame, spatial region, goal hierarchy, or causation.

Interestingly, participants did so, even after statistically controlling for the effects of the text-based distance and the base-rate relatedness of the verbs in semantic memory, which were also significant predictors of verb clustering. Thus, events are connected in a reader's long-term memory, and the strength of the connection depends on the proximity of the information in the story and whether the events share the same situation. In addition, research has also shown that less skilled comprehenders construct weaker situation models because they are less likely to cluster verbs along the situational indices than those who have high comprehension ability (Zwaan & Brown, 1996). Instead, less skilled comprehenders may rely more on the lexical similarity between verbs when sorting them from stories (Zwaan, Langston et al., 1995).

Event segmentation during reading

The Event Indexing Model may provide an adequate description of how situation models change during online comprehension and how situation models are stored in long-term memory offline. However, the Event Indexing Model does not address another important phenomenon in event processing: the segmentation of experience into separate mental models (see, Zacks, 2020 for a review) and how such a process may result in distinct memory representations.

The Event Horizon Model, which adopts Event Segmentation Theory (Zacks, 2020; Zacks et al., 2007), proposes that people spontaneously parse experience into meaningful discrete events and that events are stored as separate units in long-term memory (Radvansky, 2012; Radvansky & Zacks, 2014). According to this theory, the situation model[1] generates predictions about what will happen next. As features of an event changes, predictions become less accurate, and readers update their situation models to reflect these changes. The theory argues that readers engage in a global update, whereby the situation model resets, and a new situation model is created to reflect the new unfolding situation when there are transient spikes in prediction error (Franklin et al., 2020; Zacks, 2020; Zacks et al., 2007).

A cascade of cognitive processing occurs when people perceive a new event. Information from the previous event becomes less accessible (Copeland et al., 2006; Pettijohn & Radvansky, 2016b; Speer & Zacks, 2005), processing load increases (Delogu et al., 2018; Hard et al., 2011), changes occur in the allocation of spatial and visual attention (Eisenberg & Zacks, 2016; Huff et al., 2012; Ringer, 2018; Swets & Kurby, 2016), and information becomes less predictable (Franklin et al., 2020; Huff et al., 2014; Zacks et al., 2011). Research has also found that people read segments of text more slowly (Zacks et al., 2009), make more regressive eye movements to previously read text (Swets & Kurby, 2016), are slower to access information from previous events (Bailey & Zacks, 2015; Pettijohn & Radvansky, 2016a; Radvansky & Copeland, 2010; Swallow et al., 2009), and show transient increases in brain activity when people experience an event boundary (Kurby & Zacks, 2018; Speer et al., 2007; Zacks, Braver et al., 2001).

Unique influences of event segmentation beyond the processing of individual changes

The perception of event boundaries often coincides with the perception of shifts on story indices; however, not all changes in situational indices are considered event boundaries (Huff et al., 2014; Kurby & Zacks, 2012; Magliano et al., 2012; Speer et al., 2007; Zacks et al., 2009). Event segmentation is subjective, and it depends on how the reader directs attention to situational information and how readers construct their situation model (Bailey et al., 2017). Interestingly, research shows that the updating process that occurs when perceiving an event boundary is different from that which occurs when one experiences a change in a situational index (i.e., change in spatial location). That is, readers not only update their mental models at situation changes by integrating the new information, as argued by the Event Indexing Model, but they also engage in event updating to construct new situation models when they perceive an event boundary, a phenomena referred to as incremental and global updating, respectively (Bailey & Zacks, 2015; Kurby & Zacks, 2012; Speer et al., 2007; Zacks et al., 2009). For example, Bailey and

Zacks (2015) found that response times to recognition memory probes during reading were slow when people were probed for an aspect of the situation that just changed, consistent with the Event Indexing Model; however, they also found that participants were also slower in responding to unchanged indices of the situation compared to a no-change control condition. Thus, readers appear to update the entire model when they perceive a boundary, regardless of whether some of the information has changed in the story. Similarly, Kurby and Zacks (2012) found that people mention situational indices, during a think-aloud task, when those indices change individually, but they also found that readers mention both changed and unchanged situational information at event boundaries. Further, in a functional magnetic resonance imaging experiment, Speer et al. (2007) found a larger phasic brain response at event boundaries than nonboundary locations, but this effect was only partially accounted for by the presence of situation changes. Taken together, these findings indicate that readers may construct their situation models by tracking and updating individual situational indices while they are reading, but they segment their model when they perceive a new event. This implies that event structure may organize memories for story information above and beyond situational overlap (e.g., Zwaan & Brown, 1996; Zwaan, Langston et al., 1995).

Event boundaries and memory

Some behavioral evidence provides support for this possibility (Davis & Campbell, 2022; Ezzyat & Davachi, 2011). For instance, Ezzyat and Davachi (2011) had participants read stories containing multiple events denoted by changes in time ("a while later" vs. "a moment later") (see also, Zwaan, 1996). After reading the stories, participants recalled the next sentence shown in the story after they were given a single sentence as a cue. They found that sentences cued memory retrieval of subsequent sentences better when the two belonged to the same event than different events. Analogous effects were observed in a perceptual priming procedure in which viewers watched videos and then, afterward, engaged in a sequential priming task (Kurby & Zacks, 2022).

Together, these results suggest that event segmentation shapes how events are encoded and organized in episodic memory. The Event Horizon Model argues that situation models are stored in long-term memory as episodic units and that event memories are connected to each other via shared situational indices (Radvansky, 2012; Radvansky & Zacks, 2014). However, it remains unknown if events uniquely organize episodic memory above and beyond changes in situational features, such as temporal changes (Ezzyat & Davachi, 2011) or goal changes (Kurby & Zacks, 2019) which often, but do not always, align with the perception of event boundaries (Huff et al., 2014; Kurby & Zacks, 2012; Magliano et al., 2012; Speer et al., 2007; Zacks et al., 2009).

Hypothesis and overview

We hypothesized that event structure organizes memory for story information above and beyond links in situational event indices. To test this, participants performed a novel verb arrangement task, inspired by the verb clustering task used in Zwaan, Langston et al. (1995). After reading a story, we tested participants' memory by asking them to spatially arrange verbs on a computer screen, using the computer mouse. We instructed participants to place verbs closer together on the screen if they were considered related in the context of the story, based on their understanding of the story. We recorded the x and y coordinates of each verb on the screen, and we used the Euclidean distance between them as a measure of their association in episodic long-term memory. Analogous tasks have been used in the perception and categorization literatures (see, Berman et al., 2014; Goldstone, 1994; Kriegeskorte & Mur, 2012) to reveal how participants mentally represent similarities between different items.

We preferred this novel approach over the verb sorting task used in previous literature (i.e., Zwaan & Brown, 1996; Zwaan, Langston et al., 1995) for several reasons. First, our task is easy for participants to understand as it relies on the spatial nature in which people conceptualize similarity when making judgments based on conceptual features (Casasanto, 2008). Second, it is more efficient because participants make several similarity judgments simultaneously through the act of placing verbs spatially.[2]

We asked participants to arrange verbs from the story because verbs denote actions, they are situationally and semantically rich, and they often signal changes in the state of characters (Zwaan, Langston et al., 1995). We calculated the Euclidean distance between every possible verb pair as a measure of their association in episodic long-term memory. If events organize memory for story information, as predicted by the Event Horizon Model, then verbs from the same event will be placed closer together than verbs from different events, after controlling for shared situational indices, as well as other factors known to affect memory for story information and lexical items: orthographic similarity, knowledge, and text-based similarity. Alternatively, the strength of the association between events in memory may depend solely on the shared situational indices between them, as suggested by the Event Indexing Model (Zwaan, Langston et al., 1995). According to this possibility, verbs from the same event will not be placed closer together, but verbs that share situational indices will be, after controlling for the other predictors of similarity. A third possibility is that the strength of the memory representation may depend solely on the text-based structure of the story (e.g., whether two verbs shared the same clause or sentence) (see however Zwaan & Brown, 1996; Zwaan, Langston et al., 1995) because the closer two words appear in a text, the more likely they share arguments and make references to the same ideas (Graesser et al., 1994; Kintsch, 1988).

We evaluated if events uniquely structure memory by using segmentation data from one sample to predict verb arrangements in an independent sample. The approach we used is analogous to the three-pronged method (Magliano & Graesser, 1991), which links theory to converging sources of empirical evidence to make claims about the processes under investigation. The first prong outlines the theory, such as the Event Horizon Model., to guide the generation of hypotheses (i.e., events uniquely structure memory). The second and third prongs outline the use of different behavioral measures, such as event segmentation and verb arrangements, to find converging evidence to support a phenomenon (Graesser et al., 1987; Kendeou et al., 2019; Magliano & Graesser, 1991).

Methods

Participants

We collected data from 209 undergraduate students (145 women and 64 men, M_{age} = 19.01, SD_{age} = 2.10) enrolled in general psychology. We determined the sample size from the experimental design and from recommendations provided by Brysbaert and Stevens (2018) because the task was novel.[3] Brysbaert and Stevens (2018) suggested that one should collect approximately 160 observations per condition, per subject for a total of 28,000 observations to detect an effect with a power of .80. Each event in our experiment contained multiple verbs. Across all four texts, there were 145 pairs of verbs that shared an event. With 209 participants, we collected a total of (209 × 144) 30,096 observations per the condition of interest. Thus, we should be adequately powered to detect the effects of interest on the verb arrangement task. We collected data from 10 additional participants, but we removed their data because their comprehension questions were not recorded due to experimenter error (n = 1) or because they did not perform the verb arrangement task as instructed (n = 9). These participants placed the verbs in some identifiable spatial arrangement (e.g., in a line or in columns).

Materials

Participants read a total of five passages on a computer screen with a resolution of 1024 × 768. Participants began by reading a 440-word excerpt from the book *One Boy's Day* (Barker & Wright, 1951) for practice to familiarize them with the verb arrangement task. Participants read four experimental texts after reading the practice text. The stories contained 58 to 85 sentences and described the following: *The Beanie Baby Craze, A Farmer's Rebellion, New York in the Future*, and *Spy Identification*

Equipment. The texts were written as historical narratives, and they have been used in previous studies of discourse comprehension (Fisher & Radvansky, 2018; Radvansky et al., 2001; Therriault et al., 2006). Text length varied between stories (*The Beanie Baby Craze*: 516 words; *A Farmer's Rebellion*: 652 words; *New York in the Future*: 703 words; *Spy Identification Equipment*: 613 words). Participants always read the excerpt from *One Boy's Day* first. We counterbalanced the order of the experimental texts across participants using a 4 × 4 Williams Latin square (Williams, 1949). The texts were shown in black 14-point font against a white background, and they took up an entire page of the screen. Participants answered four true or false comprehension questions per text (e.g., "This story takes place in the future?," "Bob Collins was a well-to-do bean farmer?") after performing the verb arrangement task. Questions assessed understanding at the situation model level as opposed to surface form or text-based details. We took questions from Therriault et al. (2006) for the identical texts used in our experiment, and the questions assessed comprehension of protagonist, temporal, and spatial information. For each text, the correct answer was true for two questions and false for two questions.

Selection of event boundaries and verbs

We conducted a norming study at a large university in the Midwestern part of the United States, using an independent sample of 30 undergraduate students, to determine the normative boundaries between events in each story. Participants read each text from paper with no paragraph breaks so as not to influence their segmentation behavior. Participants segmented the text into the largest units that seemed natural and meaningful to them by placing a line between two words to indicate the end of one unit of activity and the start of the next. We gave instructions used to elicit coarse event segmentation because we were interested in testing participants' memory for events from a longer timescale (Newtson, 1973; Zacks, Tversky et al., 2001). We computed the proportion of participants that placed an event boundary within each clause, per text (see Figure S1 in the Supplemental Materials). To find moments that most people perceived as an event boundary and to ensure that the number of events were similar across stories, we identified normative boundaries as clauses that were segmented by more than or equal to 1.9 SDs above the mean segmentation probability for the story. Increasing the criteria to 2 SDs removed six events from the stories, which resulted in very few observations. The number of events varied between stories (*The Beanie Baby Craze*: 5 events; *A Farmer's Rebellion*: 10 events; *New York in the Future*: 7 events; *Spy Identification Equipment*: 5 events).

We then selected verbs from each text to use in the verb arrangement task. To do this, we identified the verb of each clause. Then, we eliminated any non-unique verbs from the pool, leaving 38 unique verbs from the *Beanie Baby Craze*, 58 unique verbs from *A Farmer's Rebellion*, 54 unique verbs from *New York in the Future*, and 35 unique verbs from *Spy Identification Equipment*. Last, we randomly selected 24 of the remaining verbs from each text, with the constraint that each event contained at least two of the selected verbs, based on the normative event segmentation data described above. The number of verbs per event for each story ranged from two to six verbs with a mode of three verbs per event. The list of verbs and the stories are provided in the supplemental materials and on the Open Science Framework (https://osf.io/h6g75/?view_only=b9c047f270954071ad85f5aaa1417012).

We chose to use normative event boundaries identified from one sample to predict verb arrangements in another sample for two reasons. Logistically, we could not ensure that there would be multiple unique verbs from the same event if we were to rely on event boundaries identified from the same individuals who performed the verb arrangement task. Second, groups of participants tend to agree on the moments they perceive as coarse event boundaries (Sasmita & Swallow, 2022); therefore, we assumed participants between samples would perceive similar event boundaries. Finally, it is common practice when using the three-pronged method to examine relations between behavioral measures using different samples of participants (Graesser et al., 1987; Magliano & Graesser, 1991).

Coding of text and verb features

We coded for several predictors to describe the similarity between the verbs in the stories.

Orthographic similarity. We computed the orthographic similarity between any two verbs using Levenshtein Distances. The Levenshtein Distance between any two verbs indicates the number of character edits (insertions, deletions, substitutions) needed to change one word into the other (Levenshtein, 1966; Yarkoni et al., 2008). The larger the Levenshtein Distance, the more orthographically dissimilar two words are.

Semantic similarity. Zwaan, Langston et al. (1995) had a group of participants sort verbs from their stories without having them read the texts. They assumed that participants would arrange the verbs based on their general knowledge and semantic similarity when they did not read the stories. We followed a more theoretical approach. We used latent semantic analysis (LSA) cosines to represent the semantic similarity of two verbs based on general knowledge (Landauer & Dumais, 1997). Briefly, LSA computes semantic similarity via a corpus analysis, which assess the co-occurrence of words and phrases across large sets of texts, paragraphs, and sentences. Through a statistical procedure, word meaning is modeled in a conceptual structure of connected latent variables. When two text inputs are given to the system, the system finds their "location" in this conceptual structure, and computes a similarity measure to reflect the result, called a cosine. The larger the LSA cosine, the more semantically similar two verbs are.

Surface distance. We created two variables to indicate the distance between any two verbs in terms of where they appeared in the text: whether any two verbs came from the same sentence and the number of intervening clauses between verbs in the story.

Situational overlap. We used the situational coding from Radvansky et al. (2001) to determine the situational overlap between every pair of verbs. Verbs were considered to overlap situationally if the verbs came from clauses that shared the same situation (space, time, entity, goal, causality). For each clause, Radvansky et al. (2001) indicated whether the clause contained a change in spatial location, temporal location, entities (objects or characters), goals, and causation from the previous clause. Text from the *The Beanie Baby Craze* story is shown in Table 1 to illustrate the situational coding of one story.

Table 1. Clauses Taken From *The Beanie Baby Craze* story. The situational coding performed by Radvansky et al. (2001) is shown. Clauses that overlap situationally share a value. A number in a cell indicates the identity of that dimension. For example, clauses 1–7 are all from the same spatial region in the story (region 1). Clauses 8 and 9 are both from the next spatial region in the story (region 2). This coding was used to determine whether any two verbs shared a situational dimension and event. For example, the verb *come* shares the same event as the verb *purchased*, but does not share the same event as the verb *taken*. (See Figure S1 in the Supplemental Materials for event segmentation norms per clause used to determine event membership.)

Clause Number	Clause	Space	Time	Entity	Cause	Goal	Event	Verb
1	Beanie babies became very popular in Mary's town of Lakewood	1	1	1	1	1	1	
2	because they are cute little toys	1	1	1	1	1	1	
3	that come in a variety of animals.	1	1	1	1	1	1	come
4	She purchased them as gifts at Christmas,	1	2	1	1	2	1	purchased
5	to give to her children.	1	2	2	1	2	1	give
6	While most collectors enjoy beanie babies,	1	2	3	1	2	1	enjoy
7	it is unlikely that they would be valued as highly today	1	3	3	1	2	1	valued
8	as they were in Lakewood last year.	2	4	3	1	3	1	
9	For some reason, beanie babies became very popular at this time.	2	4	3	2	3	1	
10	Beanie babies were first brought into Lakewood from Los Angeles.	3	4	3	3	3	1	brought
11	In 1996, Ellen Smith told Mary of seeing these toys.	3	5	4	4	3	2	told
12	She had been quite taken with the cuteness of these stuffed animals.	3	5	4	4	3	2	taken
13	In a few weeks, some toys were shipped out.	3	6	4	4	3	2	shipped
14	They were to be sold at toy shops in Lakewood.	4	6	4	4	4	2	

Space. The space dimension indicated whether two consecutive clauses took place in the same spatial region. Pairs of verbs were coded as a 0 if they came from clauses that shared the same spatial region in the depicted story situation. Otherwise, they received a 1. For instance, reading about how beanie babies were brought from Lakewood to Los Angeles reflects a change in space. Pairs of verbs that occurred before the change (i.e., come-purchased, come-give, purchased-give) were coded as a 0 because they shared the same spatial region. Pairs of verbs that came after the change (i.e., come-brought, come-told, purchased-brought) were coded as a 1.

Time. The time dimension indicated whether two clauses took place in the same time frame. Pairs of verbs were coded as a 0 if they came from consecutive clauses that shared the same time frame. Otherwise, they received a 1. For instance, reading about how beanie babies were Christmas gifts and then reading about how they would not be valued as much today reflects a change in time.

Entity. The entity dimension indicated whether two clauses involved the same object or character. Pairs of verbs were coded as a 0 if they came from consecutive clauses that shared the same entity. Otherwise, they received a 1.

Cause. The causality dimension indicated whether two clauses were directly related causally. Pairs of verbs were coded as a 0 if the prior context provided a necessary and sufficient cause for the incoming information. For instance, reading that beanie babies were popular *because* they are cute reflects continuity in causality, but then reading that beanie babies were brought to Lakewood from Los Angeles and then reading that Ellen Smith told Mary of the toys reflects a change in causality.

Goal. The goal dimension indicated whether two clauses were part of the same goal plan. Pairs of verbs were coded as a 0 if they came from consecutive clauses that shared the same goal structure. Otherwise, they received a 1.

Event overlap. Finally, using the normative segmentation data and stimuli construction procedures described above, we recorded whether each pair of verbs came from the same event, or a different event. A subset of verbs that shared the same events from the beanie babies craze are shown in Table 1. For instance, *come, purchased, enjoy, valued*, and *brought* were from the same event. *Told* and *collected* were from different events. This is because participants in the norming study judged an event boundary between the 10th and 11th clause (See Figure S1 in the Supplemental Materials). Correlations between the segmentation behavior and situation changes are provided for each text in the Supplemental Materials. We found that event boundaries tended to correspond with changes in characters, time, causality, and goals.

Verb arrangement task

After reading each story and performing a filler task (described below), we told participants they would see 24 verbs taken from the story, arranged in random non-overlapping locations on the computer screen. We instructed participants to place the 24 verbs from the story around the screen in such a way that reflected their understanding of the story. We instructed them to use the distance between each verb as a measure of how close or far they were from each other, based on their understanding of the story. We instructed participants to place verbs closer together on the screen when they felt, based on their understanding of the story, the words should go together, and to place verbs far from one another on the screen when they felt the words did not go together. Note here that the verb arrangements were not necessarily based on how close the words were located to one another in the text but based on the participants' comprehension of the story. Participants made responses by clicking and dragging each of the verbs with the mouse. We told them to use the entire screen when making their spatial arrangements. Verbs changed to a blue font when they were selected. Participants

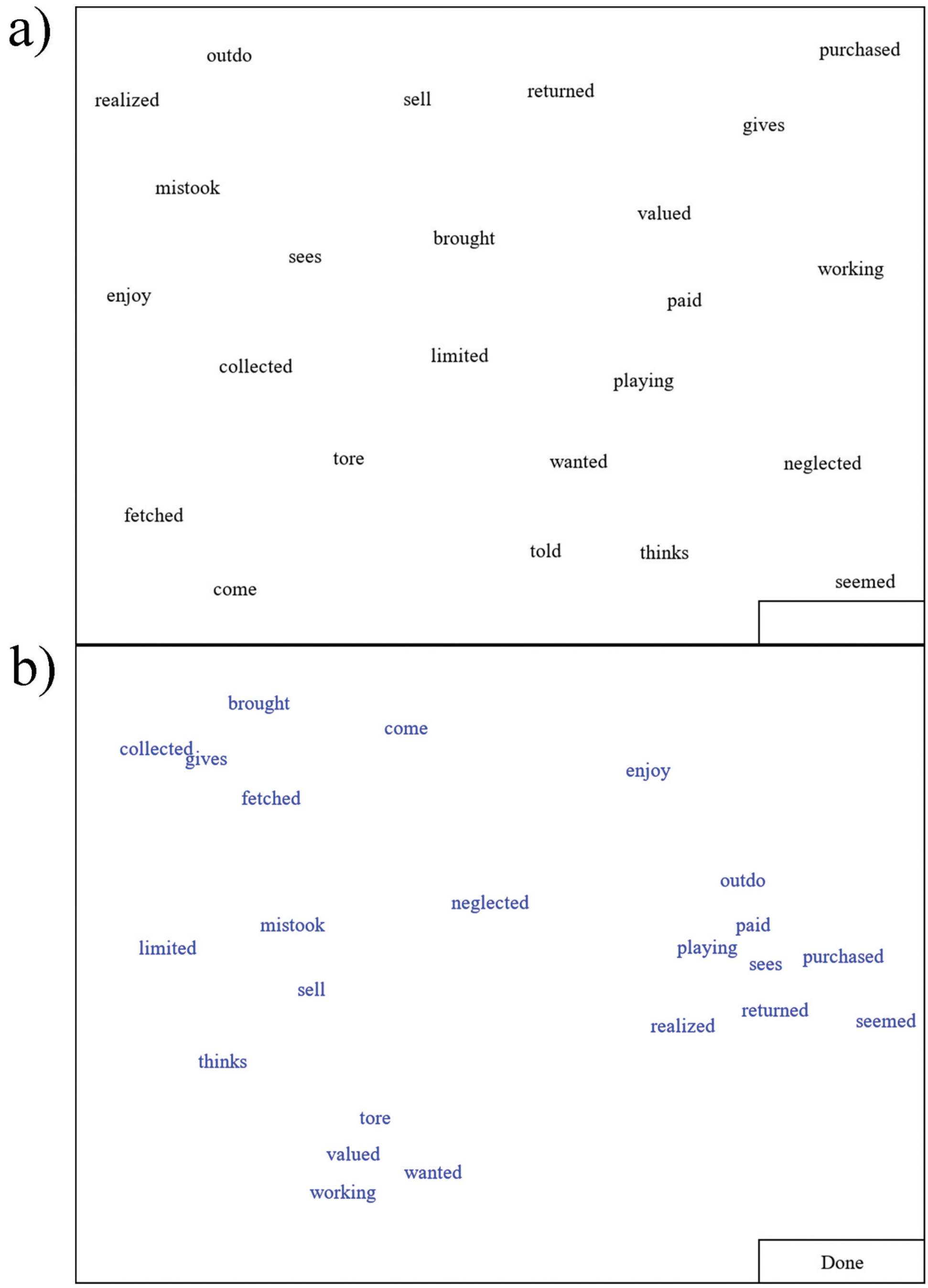


Figure 1. (A) Initial random arrangement of the verbs for one participant. (B) Postarrangement clustering of the verbs from the participant.

were not allowed to continue through the experiment until they selected each of the 24 verbs. Examples of the task are shown in Figure 1.

We did not have participants arrange the verbs without reading the stories. Zwaan, Langston et al. (1995) previously found that participants sort verbs using the lexicon when doing a verb sorting task.

We assumed that the orthographic and semantic features would be the only predictor of distance between the verbs if we included the control condition.

Filler tasks

Participants completed a filler task after reading each story. They performed a script sequencing task after the first and third texts and a script generation task after the second and fourth texts. We describe each task below.

Script sequencing

We assessed general event knowledge with a script sequencing task. Participants saw a description of an everyday activity (e.g., making a campfire) at the top of the screen along with its corresponding 12 steps. We showed participants these steps in a random order in a 3 × 4 grid. We instructed participants to arrange them into the order in which the activity would normally be executed (e.g., find a location, clear the surroundings, get the tinder, get some logs, etc.). Participants arranged the steps with the computer mouse by clicking and dragging each step into the correct position 1 through 12. Correct responses came from the normative sequences of the steps outlined for the activity (Galambos, 1983). We scored performance by calculating an error score. We calculated this score as the absolute distance of each step's location from the correct ordinal location. Larger values correspond to more errors.

Script generation

We also assessed general event knowledge with the script generation task used in (Rosen et al., 2003). Participants were instructed to write down, in order, all the steps involved in two everyday activities (going out to eat, shopping for groceries; Galambos, 1983; Rosen et al., 2003). We gave participants 5 minutes to write as many steps as possible for each activity. We scored performance as the number of words in the descriptions divided by the number of actions in the activity.

Analyses revealed no relation between these event knowledge measures and the dependent variable of interest in the verb arrangement task, and thus they will not be discussed further. However, those results are provided in the Supplemental Materials.

Procedure

We ran participants in groups of 1 to 6. They were greeted by an experimenter and seated in front of a computer. The experimenter told the participant that they would read stories about fictional events and were instructed to read each story carefully because they would complete memory tests after each story. Participants then read the practice story and performed the practice verb arrangement task.

The experimenter viewed how participants arranged the verbs after they completed the verb arrangement task for the practice story. If the verbs were not spread out around the screen (i.e., if participants put the words in some identifiable arrangement), the experimenter told the participant to try their best to arrange all verbs around the screen in such a way that reflected their understanding of the story and to use the distance between each verb as a measure of how close or far they were based on their understanding of the story. The experimenter then pressed a combination of keys on the keyboard, which permitted the participant to continue with the experiment. The experimental trials progressed as follows: Participants read a story, completed a filler task (see above), completed the verb arrangement task for the story, and then completed a four-question true or false test of their memory of the story. Comprehension questions came from Therriault et al. (2006), and they assessed participants' comprehension of protagonist, space, and temporal features. This sequence repeated for each of the four stories.

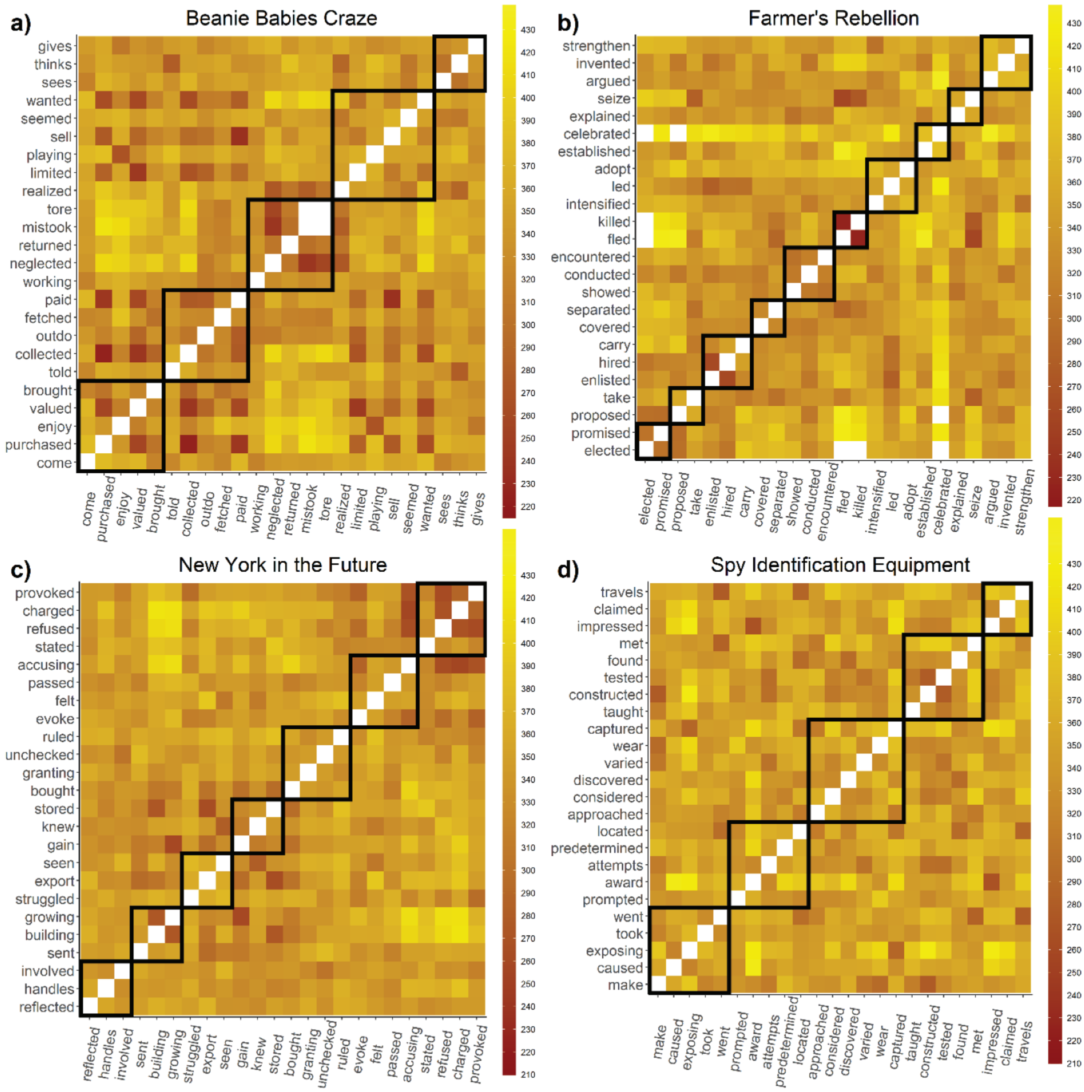


Figure 2. Mean Euclidean distance (in pixels) between pairs of verbs from the same or different events.

Results

We first examined how verbs were clustered descriptively. The mean Euclidean distance between each pair of verbs from each of the four stories is represented as a heat map in Figure 2. Verbs are organized by the order they were shown in each of the stories. Event membership is indicated by black lines drawn around clusters of cells in the matrices. Darker values in each of the cells indicate closer verb placement. As evident from the heat maps, there was substantial variability in the distance between verb pairs. Visual inspection hints that verbs that shared the same event tended to be placed closer together than verbs that came from different events; however, it is difficult from the descriptive plots alone to identify what additional factors may have influenced verb arrangements.

To assess how participants arranged the verbs from each story, we first fit a linear mixed-effects model to the Euclidean distance between each pair of verbs, per text, per subject, to test whether sharing an event was related to closer verb placement, controlling for orthographic similarity, semantic similarity, text-based similarity, and the extent to which verbs shared situational indices in

Table 2. Results of the linear mixed-effects model predicting distance between verbs.

Predictor	Estimate	SE	*t* value	*P*
Intercept	348.10	6.16	56.55	<.001***
Levenshtein Distance	2.68	0.25	10.53	<.001***
LSA similarity	−46.64	3.74	−12.48	<.001***
Clause distance	0.26	0.03	8.86	<.001***
Shared sentence	9.68	4.12	2.35	.02*
Shared space	6.97	1.45	4.80	<.001***
Shared time	−4.96	1.96	−2.53	.01*
Shared entity	8.21	2.71	3.03	.002**
Shared cause	1.67	2.32	0.72	.47
Shared goal	−8.90	1.46	−6.11	<.001***
Shared event	−8.16	1.78	−4.58	<.001***

We dummy coded the categorical predictors and centered the continuous predictors at their means before entering them into the analyses. The estimates are unstandardized weights, showing a change in Euclidean distances based on a single unit of change in the predictor. *$p < .05$; **$p < .01$; ***$p < .001$

the story. In the model, the fixed effects were orthographic similarity between each verb (Levenshtein Distances), semantic similarity (LSA similarity), text-based similarity (shared sentence, clause distance), shared situational indices (cause, goal, time, space, and entity), and whether the verb pair shared an event.

We conducted all analyses in R (version 4.0.5) using the lme4 (Bates et al., 2014) and stats packages (R Core Team, 2021). We obtained simple slopes and least square means using the emmeans library (Lenth et al., 2020). We dummy coded the categorical fixed effects and centered the continuous predictors at their means before entering them into the analyses. As shown in Table 2, all the fixed effects significantly predicted verb distance, except for having a shared cause. Importantly, the analysis shows that controlling for the orthographic and text-based distance, semantic similarity, and situational overlap, verbs were placed closer together when they shared an event than when they did not.

Exploratory analysis of comprehension accuracy

The perception of event structure and tracking of situational dimensions may depend on successful comprehension of the stories. In our data, we detected meaningful individual differences in story comprehension, as evident from participants' poor performance on the 4 comprehension questions ($M = 0.67$, $SD = 0.12$). It is possible that the shared event effect may be weaker for readers who did not successfully comprehend the stories and stronger for those who did (see also, Zwaan & Brown, 1996). To examine this possibility, we reran the above analysis and included the interaction between participant accuracy on the comprehension questions and the shared event predictor. Again, we treated the participant and story intercepts as random effects. We found that the second model was a better fit than the model that only included the main effects, χ^2 (2) = 11.78, $p = .003$, indicating that the interaction explained unique variance in the verb distances (interaction term model results: $\beta = -18.25$, $SE = 5.41$, $t = 3.38$, $p < .001$). We probed this interaction by comparing the distance between verbs when the verbs did and did not share an event at all five levels of comprehension accuracy, illustrated in Figure 3. We corrected the p values using a Bonferroni adjustment. Participants who scored 0% ($\beta = -3.88$, $SE = 3.95$, $z = -0.98$, $p = .33$) or 25% ($\beta = 0.73$, $SE = 2.81$, $z = 0.26$, $p = .79$) on the comprehension test did not place verbs closer together when they shared the same event than when the verbs came from different events. However, those who scored 50% ($\beta = 5.34$, $SE = 1.96$, $z = 2.72$, $p = .007$), 75% ($\beta = 9.96$, $SE = 1.86$, $z = 5.36$, $p < .0001$), or 100% ($\beta = 14.57$, $SE = 2.59$, $z = 5.63$, $p < .0001$) on the comprehension test did show event-based clustering, even after controlling for low-level similarity between the verbs and changes in the situational indices. Thus, those who successfully comprehended the stories showed more pronounced event-based memory organization.

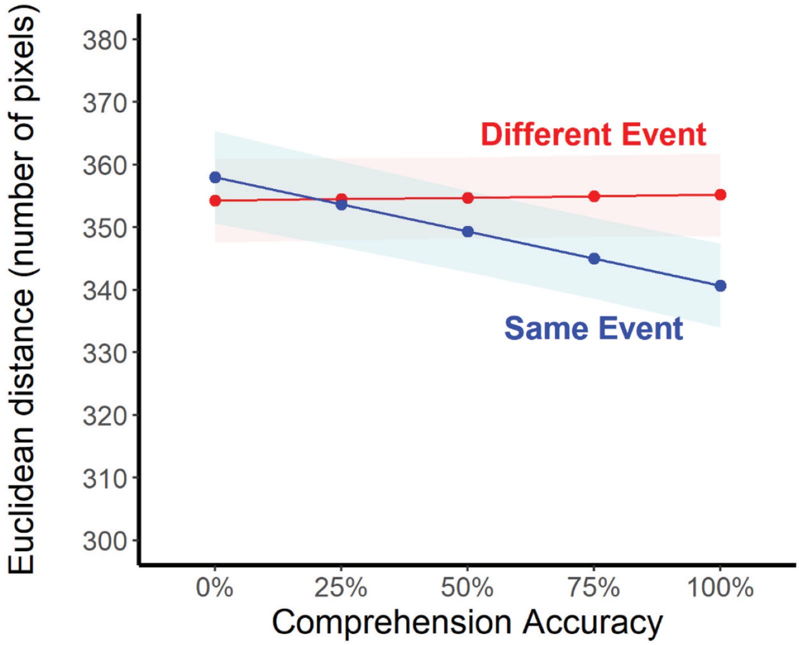


Figure 3. Least-square Euclidean distance between pairs of words as a function of comprehension accuracy and event similarity. Error ribbons correspond to ±1 SE around the estimated least-square means. We computed least-square means by setting the clause distance, the LSA difference, and the Levenshtein difference to their mean. We set all of the situational predictors to 0.

Underlying psychological representation of the verbs

The interaction between comprehension scores and the shared event factor suggests that successful comprehension of a story is related to the event-based structuring of situation models. Thus, to further characterize the theoretical semantic space of readers who comprehended the stories, we carried out a multidimensional scaling (MDS) analysis, which is commonly used to spatially represent the interrelations among a set of items using a similarity matrix. MDS is set of statistical techniques that takes a matrix containing item-to-item similarity judgments as input (such as data shown in Figure 2). It then uses data-reduction procedures to minimize the complexity of the similarity matrix, enabling one to plot a visual representation of the underlying relationships that govern the similarity judgments. In our analysis, we used the mean Euclidean distance between each verb as input to the MDS to further reveal the underlying psychological representation of the verbs. Each resulting dimension is empirically determined by the MDS analysis and is assumed to correspond to a dimension of the psychological space used to represent the verbs in memory. Before running this analysis, we removed individuals who scored less than 50% on the comprehension questions. This resulted in the removal of 102 combinations of participant and story (35.29% of total observations). We then recalculated the mean distance between pairs of verbs for each story and submitted the mean distances to a classical metric MDS analysis using the cmdscale() function in the stats library in R.

We used a scree plot to determine the appropriate number of dimensions because we did not have a priori assumptions about the number of dimensions. To assess the goodness of fit of the models, we calculated the correlation between the raw mean distances and the new MDS distances for each story. We found that a three-dimensional solution yielded a good fit to the data (*Beanie Baby Craze* $r = .72$; *A Farmer's Rebellion* $r = .75$; *New York in the Future* $r = .71$; *Spy Identification Equipment* $r = .77$). Each of the words plotted in the three-dimensional space, color coded by the events, is shown in Figure 4. Animated GIFs of the MDS analysis results can be viewed at https://osf.io/h6g75/?view_only=b9c047f270954071ad85f5aaa1417012. Consistent with the Event Horizon Model, we hypothesized

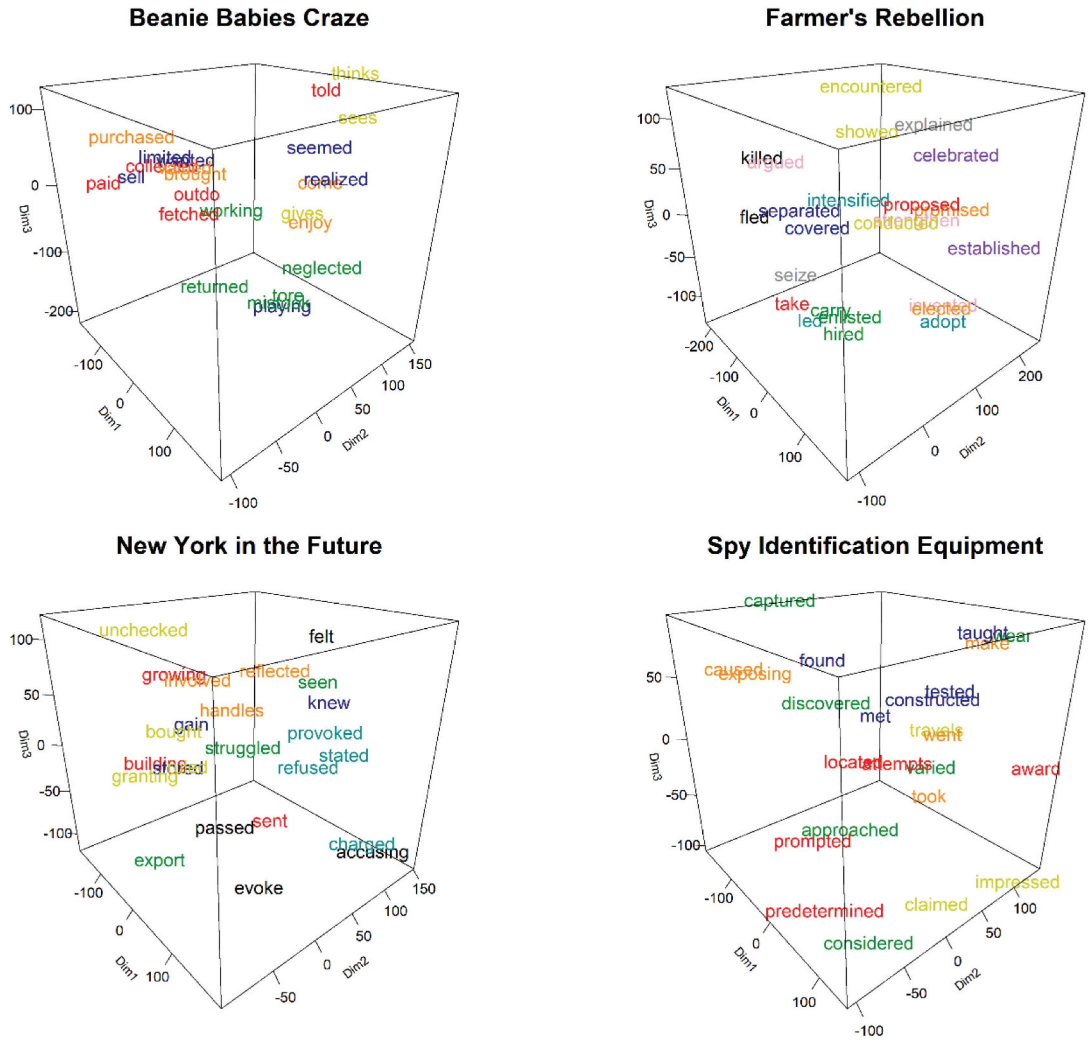


Figure 4. Results of the MDS analysis were used to plot the verbs according to their similarity matrix in a three-dimensional space, color coded by event. Downloadable animated gifs are available at https://osf.io/h6g75/?view_only=b9c047f270954071ad85f5aaa1417012.

that verbs from the same event would be closer together in this multidimensional space. Thus, the data driven MDS analysis should separate verbs based on their event membership. This expectation appears to hold true for many of the events shown in Figure 4.

To assess whether these empirically derived representations of the psychological space reflect event-based structuring, we recomputed the Euclidean distance between each of the verbs using the three-dimensional location of each verb in the multidimensional space and refit the linear mixed-effects model using the event, situational, and low-level similarity as fixed effects. We treated the story intercept as a random effect. As shown in the Table 3, participants placed verbs closer together when they were more orthographically (Levenshtein Distance) and semantically similar (LSA similarity), when they came from nearby clauses, and when they shared a goal. Critically, consistent with the Event Horizon Model, we found that participants placed verbs closer together when they shared the same event than when they did not, controlling for all low-level and situational-level similarity variables. This further demonstrates that events organize long-term memory for story information.

Table 3. Results of linear mixed-effects model predicting the recomputed distance between verbs after the MDS analysis.

Predictor	Estimate	SE	*t* value	*p*
Intercept	175.10	12.42	14.10	<.001***
Levenshtein Distance	4.60	1.39	3.30	<.001***
LSA similarity	−108.37	20.44	−5.30	<.001***
Clause distance	0.52	0.16	3.21	.001**
Shared sentence	5.92	22.61	0.26	.79
Shared space	14.10	7.97	1.77	.08
Shared time	−5.34	10.75	−0.50	.62
Shared entity	15.00	14.87	1.01	.31
Shared cause	17.34	12.74	1.36	.17
Shared goal	−23.48	8.00	−2.93	.003**
Shared event	−25.74	9.79	−2.63	.009**

Categorical predictors were dummy coded and continuous predictors were centered at their means prior to entry into the analyses. $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Discussion

When asked to recall details of a story, readers will often recall details as discrete episodes. One could argue that readers recall details in this way simply because they follow the maxims of conversation in a pursuit to communicate effectively (e.g., the maxim of manner; Grice, 1975). Our results suggest that the organization of individuals' reports have implicit structure that is systematic across people. Thus, event-organized recall is not an epiphenomenon or solely a maxim of communication. Describing what we comprehend in this way may reflect how events are structured as we comprehend stories.

In the current study, participants read stories taken from prior research (Fisher & Radvansky, 2018; Radvansky et al., 2001; Therriault et al., 2006) and then completed a novel verb arrangement task. We found evidence that the participants relied on at least four sources of information when performing their verb relatedness judgments. Participants placed verbs, according to their understanding of the story, closer together when (1) the verbs were orthographically and semantically similar, (2) the verbs were close together in the text, (3) the verbs shared a situational index, and, most importantly, (4) when the verbs belonged to the same events. This was true even after statistically controlling for, orthographic distance, semantic similarity, text-based distance, and situational overlap. Further, those who performed well on tests of their comprehension showed more event-based clustering. Together, these results suggest that event segmentation during online comprehension uniquely contributes to the organization of episodic memory for story information and contributes to how memories for actions are associated in memory.

This study was not designed to pit the Event Indexing Model and the Event Horizon Model against one another; however, we can interpret the results from those who performed well on the comprehension questions using both as frameworks. For instance, the organization of memory according to event structure is likely the result of segmenting that structure into events when readers comprehend the story (Davis & Campbell, 2022; DuBrow & Davachi, 2016; Ezzyat & Davachi, 2011; Kurby & Zacks, 2022). According to the Event Horizon Model (Radvansky, 2012; Radvansky & Zacks, 2014), readers actively maintain a mental model describing the current situation when reading, and they update the situation model whenever they perceive boundaries between events. Our results suggest that the same situation models that are constructed and maintained during online comprehension are stored in long-term memory as independent cohesive episodes when readers update their situation models at event boundaries. As proposed by the Event Horizon Model, the storage of situation models as independent episodic units may serve an adaptive function. Namely, the segmentation of events in memory may reduce the amount of information that competes for retrieval of any single piece of information (Pettijohn et al., 2016; Radvansky, 2012; Radvansky & Zacks, 2014).

It is important for us to note that we assume that event segmentation during reading (an online process) affects how memories are structured (an offline process). There was much debate early in the study of discourse comprehension on the merit of making inferences on how online processes relate to offline processes (McKoon & Ratcliff, 1986; Potts et al., 1988). However, there are also reasons, in the context of the method used here, to make this assumption (Magliano & Graesser, 1991). Both the Event Horizon Model and the Event Indexing Model propose that long-term memories are structured by how situation models are updated while reading. Our results provide converging support for this claim. Namely, event segmentation behavior from one sample of participants was able to account for how an independent sample, who successfully comprehended the stories, arranged the verbs.

Our results also raise an interesting question: What information did participants who failed to comprehend the stories use when performing the verb arrangement task? Our results do not provide a definitive answer. One possibility is that participants fall back to using simpler characteristics, such as orthographic and semantic similarity, to sort the verbs when they fail to comprehend the stories. An alternative possibility is that the poor comprehenders were inattentive when performing the experiment. Future research could evaluate these possibilities.

We also found that participants placed verbs that shared the same goal closer together than verbs that did not belong to the same goal structure, consistent with hypotheses generated from the Event Indexing Model. These results are compatible with a large body of research showing that readers actively track the intentions and goal plans of characters (Graesser & Clark, 1985; Long et al., 1992; Lutz & Radvansky, 1997; Trabasso & Nickels, 1992; Trabasso & Suh, 1993; Zwaan, Magliano et al., 1995), and that goal hierarchies help organize long-term memory for goal-related actions (Brewer & Dupree, 1983; Lichtenstein & Brewer, 1980; Zwaan & Brown, 1996; Zwaan, Langston et al., 1995). Our results extend prior work by showing that goal plans organize memory; however, our results also show that events do too. The Event Horizon Model (Radvansky, 2012; Radvansky & Zacks, 2014) also argues that separate situation models are connected in memory through their causal similarities in the story. Readers can access other situation models in memory via their causal connections to other situation models formed of the story. Although causality and goals are different, they are highly related in that they both indicate how and why story characters engage in actions, feel the way they do, and so forth (Trabasso & Wiley, 2005).

Except for changes in goals, we did not find that all the situational dimensions reliably contributed to explaining variance in verb relatedness judgments in the direction we hypothesized, based on the Event Indexing Model. For instance, we found that shared causality did not predict verb arrangements in any of our analyses and that shared time and entity predicted verb arrangements in the reverse direction (i.e., regression coefficients were positive rather than negative for those factors). Some work has failed to find effects of causality on situation model construction, in the context of when participants are asked to think-aloud as they read (Kurby & Zacks, 2012). However, breaks in causality reliably predict reading times (McNerney et al., 2011; Radvansky et al., 2014; Singer et al., 1992; Zwaan, Langston et al., 1995, 1998) even in the texts we used here (Radvansky et al., 2001), so it is unclear why causality did not affect the distance between pairs of verbs. It is also unclear why individuals placed verbs farther apart if they shared a common time frame or an entity. Further, we found that the spatial location only marginally impacted verb placement. Some work has shown that readers may not track space (Zwaan et al., 1998) or only do so the second time they read a story (Zwaan, Langston et al., 1995). Radvansky and Copeland (2010) found that observers rapidly update on the spatial dimension. Different effects of situational dimensions between our study and prior work are likely due to differences in experimental design. For instance, the verbal arrangement procedure may have changed how these situational dimensions were processed. The present study was not designed to assess this possibility, because our primary research question was on the role that event segmentation plays in organizing long-term memory. Therefore, more work will need to be conducted to understand the reversed spatial, time, and entity effects.

Last, we found that readers organized the verbs according to their orthographic and semantic similarity and by the distance of the verbs in the story. These results conceptually replicate Zwaan, Langston et al. (1995), who also found that participants sort verbs using the lexicon and text-based features. It is unclear whether the sorting of verbs using orthographic and semantic features likely reflected how people comprehended the stories, or how they engaged in the verb arrangement task, or both. Studies of semantic memory have demonstrated that the lexicon is partially organized by orthographic and semantic similarity (Landauer & Dumais, 1997; Spieler & Balota, 2000). Without further investigation, it is likely that the structure of the lexicon had an influential effect on the verb arrangement behavior observed here. Our observed effects of verb proximity in the text base on verb arrangement is consistent with the established interpretation that readers construct a text-based representation during reading (Kintsch, 1998), and our obtained results were thusly expected.

In conclusion, this study demonstrated that readers comprehend stories by constructing a situation model of the events represented in the text. Using a novel verb-arrangement procedure, our results suggest that situation models are updated and stored in episodic memory as cohesive, independent units when readers perceive a boundary between two events and that event segmentation uniquely contributes to the organization of long-term memory for story information.

Notes

1. Radvansky and Zacks (2014) use the term *event model* to capture the online representation for *what is happening now*. According to Radvansky and Zacks (2014), situation models are a subtype of event model that captures representations tied to discourse. We use the term situation model throughout the article but acknowledge that the claims we make can be applied to the more abstract concept of event models.
2. For instance, making isolated similarity judgments of 24 verbs would require (24 choose 2) 276 trials.
3. We also ran a power analysis using data from eight participants that served as pilot participants after we collected data from 209 participants. Data from these eight pilot participants were not included in the final data set. We ran the power analysis using the mixedpower library in Kumle et al. (2021). Estimating power in (generalized) linear mixed models: an open introduction and tutorial in R. *Behavior Research Methods*, 1–16. We were unaware how to use the approach prior to data collection. This approach estimates power from a simulation. The power analysis begins by first fitting a statistical model to the data. Power is calculated by repeating three steps. First, new values for the response variable are simulated using the pilot data. Second, the model refits to the simulated responses. We repeated the second step for 1,000 simulations. Third, a statistical test is applied to the simulated data with an alpha set equal to .05. We calculated power as the proportion of significant tests relative to the number of simulations (n = 1,000). We repeated each of these three steps for sample sizes from 20 to 200 participants, incrementing by 20 participants for each simulation. We found that a total sample size of 180 would be needed to observe a difference between verbs that do and do not share an event at a power of .90. Therefore, we should be adequately powered to detect a difference between verbs that share and do not share an event.

Acknowledgments

We thank J. Mac Stewart, Madaline Merle, Tori Evans, Taylor Capko, and Jennica Rogers for their help with data collection and scoring participant responses. We also extend our gratitude to Dr. G. A. Radvansky for providing us the stimuli and the situational coding of the stories we used. We also thank Dr. Lester Loschky and the members of the Event Cognition Reading Group at Kansas State University and Dr. Jeffrey Zacks and the Dynamic Cognition Lab at Washington University in St Louis for engaging in thoughtful discussion of this project with us.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Maverick E. Smith http://orcid.org/0000-0003-0345-2371

References

Bailey, H., Kurby, C., Sargent, J. Q., & Zacks, J. M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. *Memory and Cognition*, *45*(6), 940–955. https://doi.org/10.3758/s13421-017-0707-2

Bailey, H., & Zacks, J. M. (2015). Situation model updating in young and older adults: Global versus incremental mechanisms. *Psychology of Aging*, *30*(2), 1–25. https://doi.org/10.1037/a0039081

Barker, R. G., & Wright, H. F. (1951). *One boy's day: A specimen record of behavior*. Harper & Brothers.
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *Lme4: Linear mixed-effects models using Eigen and S4 (Version 1.1-7)*. http://CRAN.R-project.org/package=lme4
Berman, M. G., Hout, M. C., Kardan, O., Hunter, M. R., Yourganov, G., Henderson, J. M., Hanayik, T., Karimi, H., & Jonides, J. (2014). The perception of naturalness correlates with low-level visual features of environmental scenes. *PLoS ONE*, *9*(12), e114572. https://doi.org/10.1371/journal.pone.0114572
Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*(2), 177–220. https://doi.org/10.1016/0010-0285(79)90009-4
Brewer, W. F., & Dupree, D. A. (1983). Use of plan schemata in the recall and recognition of goal-directed actions. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *9*(1), 117–129. https://doi.org/10.1037/0278-7393.9.1.117
Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 1. https://doi.org/10.5334/joc.10
Casasanto, D. (2008). Similarity and proximity: When does close in space mean close in mind? *Memory & Cognition*, *36* (6), 1047–1056. https://doi.org/10.3758/MC.36.6.1047
Copeland, D., Magliano, J. P., & Radvansky, G. A. (2006). Situation models in comprehension, memory, and augmented cognition. In C. Forsythe, M. L. Bernard, & T. E. Goldsmith (Eds.), *Cognitive systems: Human cognitive models in systems design* (pp. 37–66). Erlbaum.
Dahl, R. (1964). *Charlie and the chocolate factory*. Alfred A. Knopf.
Davis, E., & Campbell, K. (2022). Event boundaries structure the contents of long-term memory in younger and older adults. *Memory*, *31*(1), 47–60. https://doi.org/10.1080/09658211.2022.2122998
Davis, E., Chemnitz, E., Collins, T. K., Geerligs, L., & Campbell, K. L. (2020). Looking the same, but remembering differently: Preserved eye-movement synchrony with age during movie-watching. *Psychology and Aging*, *36*(5), 604–615. https://doi.org/10.1037/pag0000615
Delogu, F., Drenhaus, H., & Crocker, M. W. (2018). On the predictability of event boundaries in discourse: An ERP investigation. *Memory & Cognition*, *46*(2), 315–325. https://doi.org/10.3758/s13421-017-0766-4
DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of Learning and Memory*, *134*(Part A), 107–114. https://doi.org/10.1016/j.nlm.2016.07.011
Eisenberg, M. L., & Zacks, J. M. 2016. Ambient and focal visual processing of naturalistic activity. *Journal of Vision*, *16*(2), 5: 1–12. Article 5. https://doi.org/10.1167/16.2.5
Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychological Science*, *22*(2), 243–252. https://doi.org/10.1177/0956797610393742
Fisher, J. S., & Radvansky, G. A. (2018). Patterns of forgetting. *Journal of Memory and Language*, *102*, 130–141. https://doi.org/10.1016/j.jml.2018.05.008
Franklin, N. T., Norman, K. A., Ranganath, C., Zacks, J. M., & Gershman, S. J. (2020). Structured event memory: A neuro-symbolic model of event cognition. *Psychological Review*, *127*(3), 327. https://doi.org/10.1037/rev0000177
Galambos, J. A. (1983). Normative studies of six characteristics of our knowledge of common activities. *Behavior Research Methods & Instrumentation*, *15*(3), 327–340. https://doi.org/10.3758/BF03203654
Gernsbacher, M. A. (1990). *Language comprehension as structure building* (Vol. xi). Erlbaum.
Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes*, *23*(3), 265–304. https://doi.org/10.1080/01638539709544994
Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. (1992). Do readers mentally represent characters' emotional states? *Cognition & Emotion*, *6*(2), 89–111. https://doi.org/10.1080/02699939208411061
Goldstone, R. (1994). An efficient method for obtaining similarity data. *Behavior Research Methods, Instruments, & Computers*, *26*(4), 381–386. https://doi.org/10.3758/BF03204653
Graesser, A. C., & Clark, L. F. (1985). *Structures and procedures of implicit knowledge: Advances in discourse processes*. Praeger Publishers Inc.
Graesser, A. C., Haberlandt, K., & Koizumi, D. (1987). How is reading time influenced by knowledge based inferences and world knowledge. In B. K. Britton & S. M. Glynn (Eds.), *Executive control in reading* (pp. 217–251). Erlbaum.
Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. https://doi.org/10.1037/0033-295X.101.3.371
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Studies in syntax and semantics III: Speech acts* (pp. 41–58). Elsevier.
Hard, B., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of Experimental Psychology: General*, *140*(4), 586. https://doi.org/10.1037/a0024310
Huff, M., Meitz, T. G., & Papenmeier, F. (2014). Changes in situation models modulate processes of event perception in audiovisual narratives. *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*(5), 1377–1388. https://doi.org/10.1037/a0036780
Huff, M., Papenmeier, F., & Zacks, J. M. (2012). Visual target detection is impaired at event boundaries. *Visual Cognition*, *20*(7), 848–864. https://doi.org/10.1080/13506285.2012.705359

Kendeou, P., Butterfuss, R., Kim, J., & Van Boekel, M. (2019). Knowledge revision through the lenses of the three-pronged approach. *Memory & Cognition*, *47*(1), 33–46. https://doi.org/10.3758/s13421-018-0848-y

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kintsch, W., & van Dijk, T. A. (1978, September). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 1363–1394. https://doi.org/10.1037/0033-295X.85.5.363

Komeda, H., & Kusumi, T. (2006). The effect of a protagonist's emotional shift on situation model construction. *Memory & Cognition*, *34*(7), 1548–1556. https://doi.org/10.3758/BF03195918

Kriegeskorte, N., & Mur, M. (2012). Inverse MDS: Inferring dissimilarity structure from multiple item arrangements. *Frontiers in Psychology*, *3*, 245. https://doi.org/10.3389/fpsyg.2012.00245

Kumle, L., Vo, M. L.-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research MethodsBehavior research methods*, *53*(6), 2528–2543. https://doi.org/10.3758/s13428-021-01546-0

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition*, *40*(5), 812–826. https://doi.org/10.3758/s13421-011-0179-8

Kurby, C. A., & Zacks, J. M. (2018). Preserved neural event segmentation in healthy older adults. *Psychology and Aging*, *33*(2), 232–245. https://doi.org/10.1037/pag0000226

Kurby, C. A., & Zacks, J. M. (2019). Age differences in the perception of goal structure in everyday activity. *Psychology and Aging*, *34*(2), 187. https://doi.org/10.1037/pag0000321

Kurby, C. A., & Zacks, J. M. (2022). Priming of movie content is modulated by event boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition 48*(11), 1559–1570. https://doi.org/10.1037/xlm0001085

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211. https://doi.org/10.1037/0033-295X.104.2.211

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). Package 'emmeans': Estimated marginal means, aka least-squares means. *R Package Version*, *1*(1), 3. https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, *10* (8), 707–710.

Lichtenstein, E. H., & Brewer, W. F. (1980). Memory for goal-directed events. *Cognitive Psychology*, *12*(3), 412–445. https://doi.org/10.1016/0010-0285(80)90015-8

Long, D. L., Golding, J. M., & Graesser, A. C. (1992). A test of the on-line status of goal-related inferences. *Journal of Memory and Language*, *31*(5), 634–647. https://doi.org/10.1016/0749-596X(92)90032-S

Lutz, M. F., & Radvansky, G. A. (1997). The fate of completed goal information in narrative comprehension. *Journal of Memory and Language*, *36*(2), 293–310. https://doi.org/10.1006/jmla.1996.2491

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics*, *20*(3), 193–232. https://doi.org/10.1016/0304-422X(91)90007-C

Magliano, J. P., Kopp, K., McNerney, M. W., Radvansky, G. A., & Zacks, J. M. (2012). Aging and perceived event structure as a function of modality. *Aging, Neuropsychology, and Cognition*, *19*(1–2), 264–282. https://doi.org/10.1080/13825585.2011.633159

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 82. https://doi.org/10.1037//0278-7393.12.1.82

McNamara, D. S., Magliano, J. P. (2009). Toward a comprehensive model of comprehension, and B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 51, pp. 297–384). Academic Press. https://doi.org/10.1016/s0079-7421(09)51009-2

McNerney, M. W., Goodwin, K. A., & Radvansky, G. A. (2011). A novel study: A situation model analysis of reading times. *Discourse Processes*, *48*(7), 453–474. https://doi.org/10.1080/0163853X.2011.582348

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal Of Personality And Social Psychology*, *28*(1), 28–38. https://doi.org/10.1037/h0035584

Pettijohn, K. A., & Radvansky, G. A. (2016a). Narrative event boundaries, reading times, and expectation. *Memory & Cognition*, *44*(7), 1064–1075. https://doi.org/10.3758/s13421-016-0619-6

Pettijohn, K. A., & Radvansky, G. A. (2016b). Walking through doorways causes forgetting: Environmental effects. *Journal of Cognitive Psychology*, *28*(3), 329–340. https://doi.org/10.1080/20445911.2015.1123712

Pettijohn, K. A., Thompson, A. N., Tamplin, A. K., Krawietz, S. A., & Radvansky, G. A. (2016). Event boundaries and memory improvement. *Cognition*, *148*, 136–144. https://doi.org/10.1016/j.cognition.2015.12.013

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language*, *27*(4), 399–415. https://doi.org/10.1016/0749-596X(88)90064-2

Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science*, *21*(4), 269–272. https://doi.org/10.1177/0963721412451274

Radvansky, G. A., & Copeland, D. E. (2010). Reading times and the detection of event shift processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 210. https://doi.org/10.1037/a0017258

Radvansky, G. A., Tamplin, A. K., Armendarez, J., & Thompson, A. N. (2014). Different kinds of causality in event cognition. *Discourse Processes*, *51*(7), 601–618. https://doi.org/10.1080/0163853X.2014.903366

Radvansky, G. A., & Zacks, J. M. (2014). *Event Cognition*. Oxford University Press.

Radvansky, G. A., Zwaan, R. A., Curiel, J. M., & Copeland, D. E. (2001). Situation models and aging. *Psychology and Aging*, *16*(1), 145. https://doi.org/10.1037/0882-7974.16.1.145

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rinck, M., & Bower, G. H. (2000). Temporal and spatial distance in situation models. *Memory & Cognition*, *28*(8), 1310–1320. https://doi.org/10.3758/BF03211832

Rinck, M., & Weber, U. (2003). Who when where: An experimental test of the event-indexing model [journal article]. *Memory & Cognition*, *31*(8), 1284–1292. https://doi.org/10.3758/bf03195811

Ringer, R. V. (2018). *The spatiotemporal dynamics of visual attention during real-world event perception* [dissertation, Kansas State University].

Rosen, V., Caplan, L., Sheesley, L., Rodriguez, R., & Grafman, J. (2003). An examination of daily activities and their scripts across the adult lifespan. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 32–48. https://doi.org/10.3758/BF03195495

Sasmita, K., & Swallow, K. M. (2022). Measuring event segmentation: An investigation into the stability of event boundary agreement across groups. *Behavior Research Methods*, *55*(1), 428–447. https://doi.org/10.3758/s13428-022-01832-5

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language*, *31*(4), 507–524. https://doi.org/10.1016/0749-596X(92)90026-T

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language*, *53*(1), 125–140. https://doi.org/10.1016/j.jml.2005.02.009

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, *18*(5), 449–455. https://doi.org/10.1111/j.1467-9280.2007.01920.x

Spieler, D. H., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychology and Aging*, *15*(2), 225. https://doi.org/10.1037/0882-7974.15.2.225

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236–257. https://doi.org/10.1037/a0015631

Swets, B., & Kurby, C. A. (2016). Eye movements reveal the influence of event structure on reading behavior. *Cognitive Science*, *40*(2), 466–480. https://doi.org/10.1111/cogs.12240

Therriault, D. J., & Rinck, M. (2007). Multidimensional situation models. In F. Schmalhofer & C. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 311–327). Erlbaum.

Therriault, D. J., Rinck, M., & Zwaan, R. A. (2006). Assessing the influence of dimensional focus during situation model construction. *Memory & Cognition*, *34*(1), 78–89. https://doi.org/10.3758/BF03193388

Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes*, *15*(3), 249–275. https://doi.org/10.1080/01638539209544812

Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. Special Issue: Inference generation during text comprehension. *Discourse Processes*, *16*(1–2), 3–34. https://doi.org/10.1080/01638539309544827

Trabasso, T., & Van Den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*(5), 612–630. https://doi.org/10.1016/0749-596X(85)90049-X

Trabasso, T., & Wiley, J. (2005). Goal plans of action and inferences during comprehension of narratives. *Discourse Processes*, *39*(2–3), 129–164. https://doi.org/10.1080/01638539209544812

Van den Broek, P., & Gustafson, M. (1999). Comprehension and memory for texts: Three generations of reading research. In S. R. Goldman, A. Graesser & P. van de Broek (Eds.), *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso* (pp. 15–34). Erlbaum.

Williams, E. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, *2*(2), 149–168. https://doi.org/10.1071/CH9490149

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979. https://doi.org/10.3758/PBR.15.5.971

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, *71*(1), 165–191. https://doi.org/10.1146/annurev-psych-010419-051101

Zacks, J. M., Braver, T., Sheridan, M., Donaldson, D., Snyder, A., Ollinger, J., Buckner, R., & Raichle, M. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, *4*(6), 651–655. https://doi.org/10.1038/88486

Zacks, J. M., Kurby, C., Eisenberg, M., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *Journal of Cognitive Neuroscience*, *23*(12), 4057–4066. https://doi.org/10.1162/jocn_a_00078

Zacks, J. M., Speer, N., & Reynolds, J. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*(2), 307–327. https://doi.org/10.1037/a0015305

Zacks, J. M., Speer, N., Swallow, K., Braver, T., & Reynolds, J. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273
Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology-General*, *130*(1), 29–58. https://doi.org/10.1037/0096-3445.130.1.29
Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1196. https://doi.org/10.1037/0278-7393.22.5.1196
Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation−model construction. *Discourse Processes*, *21*(3), 289–327. https://doi.org/10.1080/01638539609544960
Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x
Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 386–397. https://doi.org/10.1037/0278-7393.21.2.386
Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162
Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading*, *2*(3), 199–220. https://doi.org/10.1207/s1532799xssr0203_2