Event perception

Gabriel A. Radvansky[1] and Jeffrey M. Zacks*[2]

Events are central elements of human experience. Formally, they can be individuated in terms of the entities that compose them, the features of those entities, and the relations amongst entities. Psychologically, representations of events capture their spatiotemporal location, the people and objects involved, and the relations between these elements. Here, we present an account of the nature of psychological representations of events and how they are constructed and updated. Event representations are like images in that they are isomorphic to the situations they represent. However, they are like models or language in that they are constructed of components rather than being holistic. Also, they are partial representations that leave out some elements and abstract others. Representations of individual events are informed by schematic knowledge about general classes of events. Event representations are constructed in a process that segments continuous activity into discrete events. The construction of a series of event representations forms a basis for predicting the future, planning for that future, and imagining alternatives. *WIREs Cogn Sci* 2010 DOI: 10.1002/wcs.133

INTRODUCTION

What are the basic units of human experience? Objects are clearly central, and much research has been devoted to how they are represented by the mind and brain.[1] Collections of objects form scenes, and scenes as well have also figured prominently in theories of perception and conception (see COGSCI-287). People are clearly important to human experience, but people feel like a special case—one can easily imagine life on a desert island with no other people around, but not life without objects or scenes. In addition to these fundamental aspects of thought, we believe that *events* are fundamental to human experience. They seem to be the elements that constitute the stream of experience, the things that are remembered or forgotten in autobiographical memory, and the components of our plans for future action. In this article, we develop a framework for describing and studying how the mind and brain represent events, and explore accounts of how ongoing activity is segmented into events.

We start with a review of an event semantics framework for thinking about situations and events. Then, we introduce the concept of an *event model* as an actively maintained representation of a current event, and describe how it relates to similar concepts. We characterize how such representations are structured and how, in their creation, continuous activity is segmented into events. This includes such ideas as events taking place in a spatial–temporal framework, involving entities that have properties and various functional relations to one another. Along with this exposition, we illustrate points with recent developments in the field of event cognition that support these ideas. Following this we provide some basic principles of event models, including incompleteness, flexibility, and relation to more general semantic knowledge. We then discuss how people update their event models to take in and handle new information. Finally, we present a discussion of more recent work on embodied cognition and perceptual symbol theories and the importance that these hold for event cognition.

SEMANTICS OF SITUATION AND EVENTS

In their book *Situations and Attitudes*,[2] Barwise and Perry provided a formal description of how events are structured as part of their theory of situation

*Correspondence to: jzacks@artsci.wustl.edu

[1]Department of Psychology, University of Notre Dame, Notre Dame, IN, USA

[2]Departments of Psychology and Radiology, Washington University, Campus Box 1125, St. Louis, MO, USA

DOI: 10.1002/wcs.133

semantics. Here, we concentrate on what they had to say about event structure. Their ideas are grounded in the idea that the world can be organized in a variety a ways, but only some of these are recognized by people. As such, the important elements and structure of an event are often, in some way, imposed by a person. There is presumably a reasonably high degree of uniformity across people in the way that they conceive of events. Thus, the components and structure of an event are not deterministically derived from the world itself.

In Barwise and Perry's account, events have a number of important properties. The basic components of an event are individuals, the relations among individuals, their properties, event states, and spatiotemporal locations. These components are present in all events. Individuals can either be considered as whole entities (e.g., Joe), or they can be broken down into parts in which each part is considered as an individual (e.g., Joe's arm).

There are three basic types of relational information in Barwise and Perry's theory. *Binary relations* capture the interrelations among individuals and can capture actions among people, such as one person kicking another, or social and kinship relations, such as one person being another person's mother. *Properties* are a special class of relations that apply only to a single individual. Finally, *situational states* are relational information that captures general characteristics of the event, such as the fact that it is raining or that it is noon.

In addition to the different types of relational information noted above, events also are embedded in *spatiotemporal locations*. The time and place at which an event occurs serve as a framework for the event itself, constraining the individuals and relations and their configuration.

Barwise and Perry distinguished between two types of events: states of affairs and courses of events. A state of affairs is confined to a single spatiotemporal location, and thus is not dynamic. A state of affairs could be captured by taking a photograph. By contrast, a course of events is a collection of states of affairs, and thus is dynamic. A course of events unfolds over time and space. For a course of events to occur, there needs to be some concept(s) in common across the states of affairs that serve to unify them. Individuals can serve as such invariant uniformities.

Finally, Barwise and Perry distinguish between real and abstract events. Real events are parts of the world, whereas abstract events correspond to nonphysical domains, such as mathematics. From our perspective, under some circumstances, a common location may be generated not by an actual physical location, but by a *virtual location*. This virtual location is an abstract region that, when the two entities enter into it, allows them to interact with one another although their physical locations would be recognized as being quite distinct. For example, a virtual location may be created when two people are having a conversation over the telephone, or when a protagonist thinks about another location.[3] Thus, locations can be conceptual as well as physical.

In sum, situation semantics proposes that events consist of entities that have features. Some of these features relate entities to each other. A state of affairs is a static configuration of entities and features that is localized in time and space. A course of events is a sequence of states of affairs that unfolds over time, and is held together by some common attribute. Barwise and Perry refer to states of affairs and courses of events as two types of events. However, it is the second that corresponds to the psychological notion of event that we wish to characterize here. In the following sections we investigate how a sequence of states can be bound together into a coherent event.

MODELPALOOZA

We are now ready to turn from events themselves to their mental representation. The literature uses a variety of terms to describe the representations people construct of real or imagined experiences. These can be summarized with the hierarchy illustrated in Figure 1. At the broadest level is the *mental model*—a term introduced by Johnson-Laird[4] and widely adopted by cognitive psychologists. For us, a mental model is a representation of a set of circumstances. This set of circumstances may or may not be tied to a specific event. As such, the term mental model is quite broad and general. Mental models can be divided into two classes: *system models* and *event models*. System models are mental models representing a functional system, such as how a mechanical device works, how a computer programmer processes information, or how a theoretical construct operates. System models themselves can be divided into two general classes[4] These are *physical system models*, which capture our understanding of physical devices, such as a thermostat, a car engine, or a drawbridge, and *abstract system models*, which capture our understanding of systems that are either wholly abstract, such as theories of mathematics, or cannot be directly perceived, such as theories of subatomic physics. Other than noting their existence, we do not consider system models further.

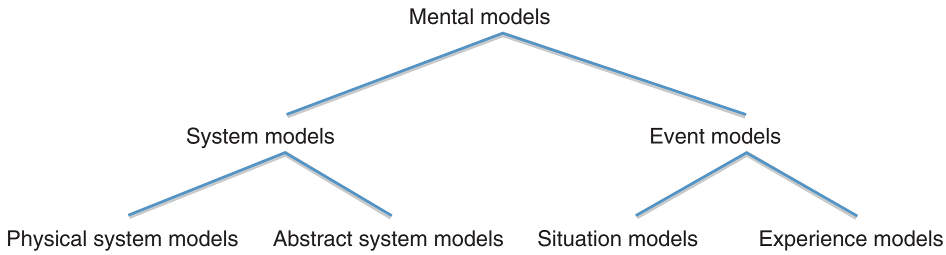

FIGURE 1 | Hierarchy of mental models.

Event models capture the entities and functional relations involved in understanding a specific state of affairs. This is in comparison to system models, which hold true across a range of circumstances. Event models can be derived from perceptual/motor experience, such as our own interaction with the world, television, film, and virtual reality technologies, or from linguistic descriptions. Because event models derived from language have received extensive attention, we believe that it is helpful to distinguish these two cases. The term *situation model* was introduced by van Dijk and Kintsch[5] to refer to event models derived from language. We use the term *experience model* to refer to event models derived from live experience. Some researchers working with narratives have taken as a working assumption that situation models derived from language share most of their properties with experience models.[6] Our focus here is on event models.

EVENT MODELS

Event models are multidimensional representations. Real-world events are composed of different types of information that are interrelated in complex ways. As such, the mental representations of events should reflect this complexity. In this section, we outline some of the major types of information that are involved in events. We draw heavily from previous work on the role of situation models in narrative comprehension and memory, in particular the event indexing model.[7,8]

Spatial–temporal Frameworks

First, each event occurs in the context of a spatial–temporal framework that serves as the basis of an event model when it is created. Spatiotemporal location is a major organizing factor in memory for events. All animals that have been studied show generally poorer memory for events that are more distant in time.[9] Although one might intuitively believe that the temporal distance effects are due to the decay of memory traces, the evidence suggests that they are mostly due to the differences between the spatiotemporal frameworks at encoding and retrieval. When the conditions of retrieval match those of encoding memory is good and when they mismatch memory is bad.

We distinguish between three types of spatial information: labels, locations, and relations. A spatial label is simply the name given to describe a location, such as 'Wrigley Field', 'The White House', or 'Slovakia'. Spatial labels are property information.

The spatial location itself defines where the event unfolds. This could be a physical location, such as a room, a park, a city, or an abstract location, such as a conference call, a web site, or a virtual environment. The centrality of spatial location to events is clear. Many studies of event cognition focus on either spatial knowledge itself, or the impact of spatial information on other cognitive processes.

Spatial relations involve the spatial interrelations among objects in an event. To illustrate this, consider the sentence 'The boy was next to the tree in the park'. The term 'park' is the label, the park itself is the location, and the boy's being next to the tree is a spatial relation. However, it is possible for people to treat spatial relations as a means to define subregions within a larger spatial context.[10]

It is also important to note that it is possible for event models to be 'viewpoint' dependent to some degree. That is, the model may be experienced or 'read' based on a particular perspective from within the model. This viewpoint may be based on the perspective in which the information was originally learned, such as whether a space was experienced from a survey (birds-eye view) or route perspective.[11,12] Recent work on embodied cognition has also demonstrated an influence of such perspectives on an event. For example, in a study by Borghi, Glenberg, and Kaschak,[13] people were asked to verify attributes of a named item. The speed with which they verified those attributes varied as a function of the perspective taken. For example, if asked to verify information about a car's steering wheel, this was done faster if one had the event perspective of being in the car as compared to being outside of it. Finally, the influence of event perspective can be seen in the difference between field and observer autobiographical memories.[14] Field memories are the memories that convey the field of view a person had when the event was actually experienced. By contrast, observer memories are those in which a person can see one's self in the event, which is, in some sense, inaccurate because one rarely sees one's self during an event unless it involves a mirror, television camera, or some such. The fact that people can have both field and observer memories suggests that there is some flexibility in the perspective that one takes on an event model.

For an event model to be constructed, a spatial location needs to be specified or inferred. When explicit information about a location is given, such as its label, or can be easily inferred from the context, then an event model will use that location for the framework. However, if little or no information is given, then a person may establish a general 'empty stage' which serves as the location information for the event model framework.[15] For several pieces of information to be integrated into an event model a person must be either explicitly told or must infer that the same location is involved.

The importance of spatial locations is reflected in the finding that certain brain regions seem to be specialized for processing location frameworks. As one example, the parahippocampal place area (PPA) has been reported to be more active when people are looking at pictures of locations as compared with pictures of faces or objects.[16,17]

The temporal framework is the period of time in which an event is unfolding. Like space, time is important for events. In many cases, a time frame is defined by the activities carried out by the agents involved in a spatial–temporal framework,[18] as well as how a person is parsing up the passage of time.[19] Activities that occur at different times are unlikely to be considered to be part of the same event, unless there is some unifying relationship, such as a causal one. For example, if someone sets a trip wire at one time, and then a person walks across it at some later point in time, this can be considered a larger, extended event with a larger temporal separation between the two sub-events.[20] However, in general, a discrepancy in time is likely to cause the information to be attributed to separate events. Two entities in a common location are less likely to be part of the same event if they were in that location at different times. For example, if a tiger were in the same room as a zookeeper, it would be inappropriate for us to draw the inference that the tiger is going to tear the poor zookeeper to shreds if we also know that they were in the room at different times. So, for an event model to be constructed, temporal location needs to be explicitly specified or inferred.

The duration enclosed by a time frame can vary widely depending on the actions or relations that are involved. For example, a telephone call from a telemarketer would comprise somewhere from a few seconds to a few minutes, but a vacation could be anywhere from several days to several months.[18] In many situations, one may not have detailed information about the duration of an event, such as Japanese tea ceremonies and a train trip from Amsterdam to Rotterdam[21,22] Moreover, the exact boundaries of a temporal frame may be more or less fuzzy. An inning in baseball or a courtroom trial session each has a precise beginning and ending, but an afternoon of reading or a dorm party may not.

Although time and space are primitives in the sense that they both are necessary for the creation of an event model framework, there are significant differences. First, spatial location is typically more restrictive in defining an event than is time frame. For example, knowing that a set of events took place in the emergency room of a hospital provides more of a constraint than knowing they took place on Thursday between 7:00 a.m. and 8:00 a.m. As such, we expect spatial framework information to be used more during comprehension and to serve as a better retrieval cue during memory retrieval, and research on autobiographical memory suggests that it does.[23]

Second, spatial location is less likely to be redundant with other information. For example, if you know a person is eating at a restaurant, you know general information about the location, but you do not know the exact restaurant they are at; information that can have a profound impact on understanding the event. There are no obligatory spatial markers that locate an event in space relative to the place of utterance. By contrast, temporal framework information is more likely to be redundant with knowledge of the activity, and is more likely to be linguistically marked, as with verb tense. Furthermore, time frames can be localized within a larger time line using temporal relational information. That is, an event can be defined relative to others or the current time.

Third, space is symmetric whereas time is not. There is nothing special about Joe's being east of the tree rather than west, or right rather than left. However, for Joe to leave the bank before the robber arrived is very different than for Joe to leave the bank after the robber arrived.

Entities and Properties

The entities in an event are often what are of most interest to people. A great deal of event processing is oriented around them. For example, a person may want to know what the moods of another person are, how do they relate to the other people in the situation, what their goals and desires are, etc. This property information is bound in the event model to the token standing for the entity.

Entity tokens are the points at which event models make contact with causal relations. To understand an event it is important to know how the entities involved both are the sources of causation and are affected by other sources of causation. As such, the degree to which an entity requires or is involved in linking relations influences its importance in an event model. Entities that are highly connected are more likely to be represented, whereas those that have little to no bearing are unlikely to be well integrated into the model.[24–27]

Associated with each entity is a collection of properties that identify it, such as its name, physical characteristics, internal characteristics (e.g., emotions and goals), and so forth. Information about properties can be directly associated with the entity. However, not all properties are stored in the event model. Those properties that are causally relevant are more likely to be stored. Property information that is important but not directly relevant may be stored in the event model, but may also be kept out and only be stored in a more generalized representation where it can be called upon as needed.

A basic property of an entity is its name or identity. While this is a seemingly central aspect of an entity, the name does not need to be specified in the model. For entities that recur across many different events, this information may be stored in a referent-specific way and be accessed as needed. For other entities, a *specific* name or identity is unnecessary. For those entities, a token is present in the model and is not identified, but serves as a placeholder. For example, if you know that Bill stood in line at the movie theater, the model may contain tokens for the people ahead and behind him in line. However, the identities of these people are left unspecified. For those entities that are not identified, if the context is sufficiently constraining, and the identity information is needed to understand a structural or linking relation, then identity can be inferred. For example, if a person is told that Bill is getting a root canal, one can infer with a high probability that a dentist is the one giving it to him.

Another type of property associated with an entity is its physical characteristics. These can include things such as size, color, texture, shape, and so forth. This is in line with perceptual symbol theories of cognition (see below) that assume that people use mental representations that incorporate and depend on perceptual qualities. Again, as with most other entity properties, these characteristics are less likely to be incorporated into a model unless they are important to understanding the functional structure of the event.

While some entity properties have a perceptual quality, others are internal to the entity, and would not be perceptual per se. This can include things such as physiological state or health. Probably one of the more interesting internal properties is a person's emotions. Emotions are important because they often provide a source of impetus for actions and provide a source of causal explanation for why people act the way they do. It is clear that people use emotions to aid comprehension.[28–31] Nonliving things also have internal properties. For example, a car may have a dead engine or a cabin may be warm or cold inside.

Goals are an important internal property of entities that serve as agents.[32,33] A goal is a state of affairs that an entity acts to bring about. They are related to desires, needs, and motivations. Goals may be caused by desires or needs. For example, a person who desires a martini or needs water to prevent dehydration may adopt the goal of ingesting a beverage. However, desires and needs can be distinguished from goals because it is possible to have the former without the latter. A castaway on a desert island may desire a television but not adopt the goal of obtaining one because it is futile. Because desires and needs cause goals, and goals cause behaviors, goals allow us to understand the actions of others.

Structural Relations and Linking Relations

Often what makes an event unique is not the space and time that it occurred, nor the entities involved, but the relations of the entities to one another, the location they are in for that event, or time at which that particular event is occurring relative to others. Relational information provides the unique structure for an event apart from the specific identity information assigned to the other elements.

There are two types of relational information. One is *structural relations* that specify the interrelations among entities. These can be either within a given time frame, such as the spatial arrangement of objects, or more stable characteristics, such as the layout of a building. Some examples of structural relations are spatial relations, ownership relations, kinship relations, social relations, etc.

Spatial relations are structural relations that convey the spatial configuration of entities within an event. However, while a spatial framework is needed for event model creation, spatial relation information is not. Instead, spatial relations are more likely to be ignored unless (1) the person has a goal of learning this information, (2) there are enough processing resources to devote to this information, or (3) the information is functionally or causally important to understanding the situation.[34]

Another important point about spatial relations is that they help define a framework from which other aspects of the event are interpreted. Franklin and Tversky[35] describe a set of spatial relations that locate objects relative to one's body. One critical feature of their account is that these reference directions are anisotropic. For example, whether something is up or down relative to the body is more accessible than whether it is left or right.[35] Moreover, these spatial frameworks can vary with respect to the perspective a person takes, such as whether a person reading a narrative takes a first-person or third-person perspective on the events being described.[36]

The other type relation is *linking relations* that serve to link different events into a sequence or collection. The most common types of linking relations are temporal and causal relations. Temporal relations convey when a given event occurs relative to others, rather than providing an absolute time period. Causal relations provide information about a causal chain of events.

Causal relations are central for cognition. For example, when placed in classical conditioning situations, people often try to make a causal attribution,[37] even if their behavior is actually guided by mechanisms that do not take causality into account. That said, people do not always draw causal connections between two pieces of information. For example, in a study by Fenker, Waldmann, and Holyoak,[38] people judged the causal relationship between two words (e.g., *spark* and *fire*). Responses were faster when the cause preceded the effect than the reverse. However, when people judged whether two words were associated, and did not make a causal judgment, this pattern was not observed. Still, overall causality is central to event model construction.

Causal relations involve a temporally ordered dependency because causes precede their effects. Further, causal relations require an event model with at least two spatial–temporal frameworks. The first includes the cause and the second the effect. For example, consider 'The rocket launched the satellite into orbit'. An event model capturing this requires a spatial–temporal framework whose spatial location is the launch pad and whose temporal frame is earlier, and another whose spatial location is earth orbit and whose temporal frame is later. Thus, causal relations are a part of extended, dynamic event models.

The degree to which information is part of a causal chain can influence whether it is interpreted as being part of the same or a different event. Information that is causally unrelated is more likely to be stored in separate models. By contrast, information that can be interpreted as being part of the same causal chain is more likely to be interpreted as being part of the same event.[39]

Causal relations may also influence other aspects of a situation model. Information that is relevant to the causal chain is more likely to be encoded than other information.[34,40] For example, knowing that an open beach umbrella provides shade from the sun, or that a magnifying lens dropped in dry grass can cause a fire, increases the probability that these spatial relations will be encoded. Information that is relevant to the causal structure of an event model is said to be functionally relevant.

BASIC PRINCIPLES

In this section we outline some basic principles of event models. The aim of these principles is to provide a guide to how event models are constructed and operate in cognition. What kind of information do they contain, and what kind of information are they unlikely to contain? How do they relate to other types of knowledge?

The first of these principles is that event models capture important characteristics of events in a ways that are functionally parallel to a real situation. 'Functionally parallel' means there is a second-order isomorphism[41] between the representation and the thing represented. For a second-order isomorphism, there is a one-to-one mapping between modeled events and possible events, and this mapping preserves similarity relations. For example, if one reads that two poker players sat down at a table one might construct an event model with tokens for the players. One might then read that either one more player or two more players joined the first two. The number of players is represented in the event model in a functionally parallel fashion with the resulting representation being more similar for the one-more-player case than the two-more-player case. The nature of the similarity is not determined: the number of players could be represented, e.g., by the firing rate of a population of neurons or by the particular population of cells that are activated. Second-order isomorphism can be contrasted with a first-order isomorphism in which a representation is a physical reproduction of the actual stimulus. In that case, there would have to be a physical token for each poker player in the nervous system, and adding a player would require adding a token. Second-order isomorphism also can be contrasted with representations based on arbitrary relations. For example, if the number of poker players were represented using English numerals, then the representation of two poker players ('2') would be no more similar to the representation of three players ('3') than the representation of four poker players ('4').

The second principle is that although event models are isomorphic to the events they represent, they are not full-blown and complete replicas of them. Instead, a model contains information relevant to understanding the basic structure of the event and little more. That is, they may include components that are sketchy in a similar way that a drawing may be incomplete or have elements that are occluded. Because event models are partial representations, this gives them a degree of flexibility in terms of the number of possible event configurations that they could conform to. For example, when reading a novel and then seeing a movie adaptation, the incompleteness and flexibility of the event models created during reading allow one to map this information into one's experience of the movie and allow one to recognize it as the same basic story. Perhaps one reason many of us find that movie adaptations do not live up to the books on which they are based is that the event models created while reading retain a greater element of flexibility. (In addition, there are details that can be conveyed by text that are difficult, if not impossible, to convey by film.) Note well that incompleteness is different from abstraction, although it can facilitate abstraction. For example, if one has visited a certain bar on a number of occasions, on returning one may be reminded of those previous occasions. Sketchier event models may facilitate this reminding if they capture the features common to the repeated instances, such as the location and the selection of beers on tap, while omitting the features that vary from instance to instance such as one's clothes and companions.

Event models are not holistic representations. Instead, they are componential. One can see this by studying how event models are constructed. It appears they must be constructed piece by piece, and this process can be time consuming and effortful. This was clearly shown in a study by Zwaan.[42] In this study, people read narratives. Half of them were told that these were newspaper articles (describing real events) and half were told that these were works of literature (where the focus is on the language itself). Later testing measured memory for three levels of representation: the surface form of the text (the exact wording), the specific facts or propositions asserted by the text, and the situation described by the text, including implications and assumptions not explicitly asserted. As can be seen in Table 1, people in the literary condition remembered the surface form better than those in the newspaper condition, suggesting they processed the details of the wording more thoroughly. People in the newspaper condition, on the other hand, had stronger situation model memory, suggesting a more thorough encoding of the events.

TABLE 1 | The 'd' Scores for the Various Levels of Representations in the Study by Zwaan[42]

	Surface Form	Propositions	Situation Model
Newspaper	−0.05 (0.54)	1.20 (0.74)	0.66 (0.80)
Literary	0.27 (0.42)	1.46 (0.90)	0.14 (0.71)

Standard deviations in parentheses.

Different groups of comprehenders may have habitually different model construction habits. For example, Stine-Morrow, Gagne, Morrow, and DeWall[43] (see also Radvansky and Copeland[26]) found that on initially reading a story, older adults had relatively better memory for the situation model compared to the surface structure of the text, whereas younger adults had relatively better memory for the surface structure. On a second reading, older adults filled in more of the surface structure, whereas younger adults filled in more of the situation model. This shows that the richness and completeness of one's event model may vary depending on one's momentary comprehension goals and based on one's habits of understanding.

Referent Specific Knowledge

Events are unique. As such, each event model should, for the most part, stand for a single event. However, despite the ultimate uniqueness of events, there are elements that transcend a single event. Our knowledge about the commonalities across a set of events is a form of semantic knowledge, and the knowledge structures that store it are referred to as *event schemas*. Whereas event models represent particular events (*instances*), event schemas represent classes of events (*types*). The concept of an event schema[44,45] is related to the more specific notions of *script*[46] and *structured event complex*.[47] Unless we are referring to the specifics of one of these constructs we use the more general term 'event schema'.

Event schemas are helpful when a person needs knowledge to understand stereotypical aspects of an event to either fill in unmentioned, but highly likely components for which a person has a large knowledge base. Our view is similar to that of the schema-pointer-plus-tag model of schema processing.[48–52] In our view, pan-event information about specific referents is stored in a referent-specific memory apart from the event model. Moreover, nonrelevant aspects of an event are not directly represented in a model, but the model may contain a pointer indicating that information is stored elsewhere. This is especially true for the minor and peripheral aspects of an event. For example, for a model of the events in a house, rather than containing information about all of details of each room, the model may contain pointers to such knowledge.

SEGMENTATION

Suppose that you work in an office, and on arriving one morning you stop by the mailroom where you have a conversation with a colleague, then walk down the hall to the lounge and pour yourself a cup of coffee. Most theories would suppose that you formed two event models, one for the mail-checking event and one for the coffee-serving event. People naturally segment ongoing activity into events at such points, producing segmentations that are reliable across observers and within observers over time.[47,53] But how are these events individuated? Gibson[54] held that events are individuated in virtue of invariant spatiotemporal patterns that persist throughout the duration of an event. This works nicely for simple physical events—e.g., the event of a ball being dropped and bouncing is unified by the dynamics of bouncing under gravity that persists throughout the bouncing.[55] However, to date, no one has offered an account of the individuation of everyday events—those involving people and goal-directed actions—in these terms, and it is not clear that such an account is possible. For events of this sort, event segmentation theory (EST)[56] offers an account of how continuous ongoing activity is segmented into events. Here, we give a brief summary of EST. Other introductions are presented by Kurby and Zacks,[57] Zacks and Sargent,[58] and the full theoretical account is given by Zacks et al.[56]

EST begins with the presupposition that observers attempt to predict the near future as an ongoing part of perception. Perceptual processing transforms sensory inputs to elaborated representations that include predictions. This processing is influenced by event models. Event models bias the perceptual processing stream, allowing the comprehender to fill in missing information and disambiguate ambiguous information. For event models to be effective, they must for the most part be shielded from the vicissitudes of the sensory input, holding a stable state in the face of missing, ambiguous, or partially conflicting information. Thus, while checking mail one's event model continues to represent the location of your mailbox, even if it is temporarily occluded by a colleague leaning down to retrieve her or his own mail.

To be effective, event models also must be updated from time to time. If not, one will be afflicted by perseveration—a mail-checking event model that held over into a coffee-serving event could lead one to misinterpret a drawer of silverware as a mail slot. But how to update event models at just the right time, without an external signal cuing what the events are, or even that a new event has begun? EST proposes that event models are updated in response to transient increases in prediction error. At some point as you leave the mailroom and head toward the lounge, things are likely to become less predictable. You walk through doors, encounter different people and objects, and see new causal sequences and goals in progress. EST claims that when prediction error increases comprehenders update their event models based on the currently available sensory and perceptual information. In most cases, the new event model will be more effective than the old one, and the prediction error will decrease as the system settles into a new stable state. This process is illustrated in Figure 2.

In the example of arriving at work, checking mail, and getting coffee, the comprehender experiences the activity in the flesh and is also a participant. However, one may hypothesize that the same mechanisms apply when the comprehender is a passive observer, and apply whether the events are experienced through sight, sound and touch or through reading. EST further proposes that people simultaneously maintain event models on multiple timescales. At a fine temporal grain, prediction error is integrated over a relatively brief temporal window, and brief increases in prediction error lead to updating. At coarser grains, prediction error is integrated over longer windows, and more sustained error spikes are needed to produce updating. Updating will tend to be hierarchical, such that coarse-grained updating rarely occurs without simultaneous fine-grained updating.

What is the evidence for these proposed mechanisms? Behavioral and neurophysiological data provide evidence that people construct a string of event models during perception and reading. In perception, passive viewing of events produces transient increases in brain activity at event boundaries.[59,60] Similar results have been observed during reading.[61,62] Converging results come from the finding that observers tend to slow down at event boundaries[63], and at those points at which changes that may produce event boundaries occur.[22,64] There is also evidence

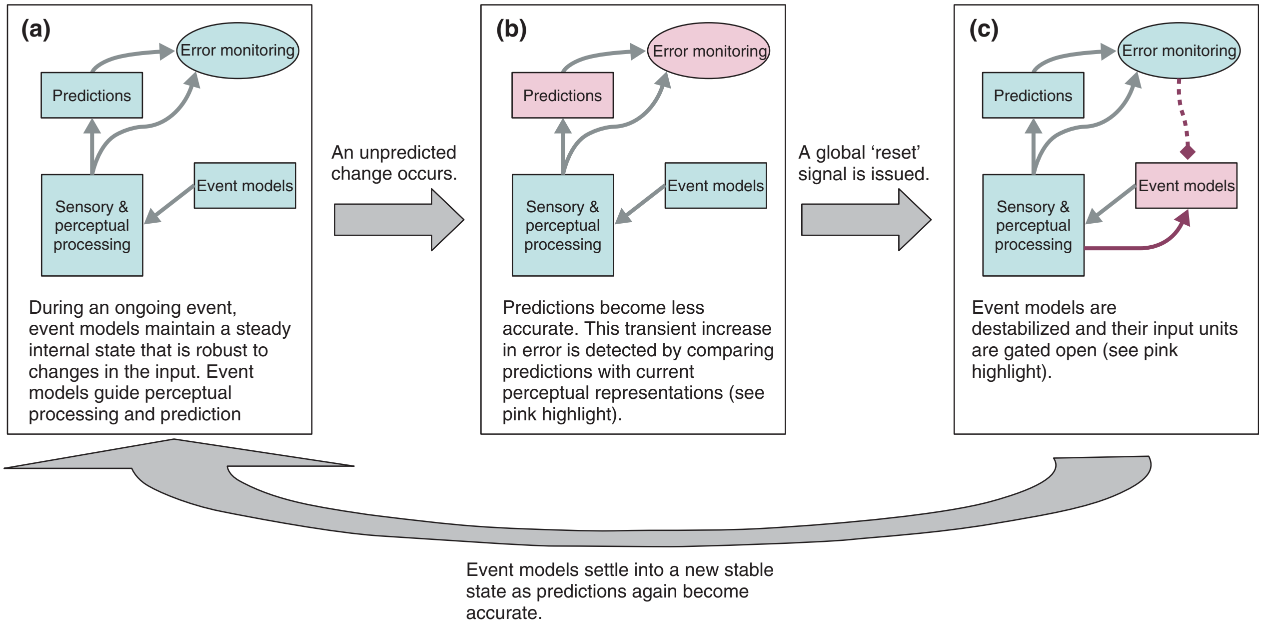

FIGURE 2 | A schematic depiction of how event segmentation emerges from perceptual prediction and the updating of event models. (a) Most of the time, sensory and perceptual processing leads to accurate predictions, guided by the event models that maintain a stable representation of the current event. Event models are robust to moment-to-moment fluctuations in the perceptual input. (b) When an unexpected change occurs, prediction error increases and this is detected by error monitoring processes. (c) The error signal is broadcast throughout the brain. The event models' states are reset based on the current sensory and perceptual information available; this transient processing is an event boundary. Prediction error then decreases and the event models settle into a new stable state. (Reprinted with permission from Ref 57. Copyright 2008 Elsevier).

that viewers[65] and readers[63,66] update memory representations at event boundaries. For example, in one recent study Swallow, Zacks and Abrams[65] presented viewers with clips from professional cinema that had been segmented by a previous group of viewers. From time to time the clips were interrupted and the viewers' memory for recently presented objects was probed. In all cases the objects had last been seen exactly 5 seconds previously. However, when a new event had begun during those intervening 5 seconds, information in working memory was rendered much less available. (There was also evidence that information which had been encoded into long-term memory was slightly *more* available.)

These memory findings support the general proposal that event models are updated at event boundaries. They also raise the more general question of how event models are updated, to which we now turn.

UPDATING

There are four general ways the contents of an event model can be updated. The first is that a new model is constructed (*model creation*). Second, information may be incorporated into an existing event model (*model elaboration*). A third possibility is that an existing model is altered to accommodate new information (*model transformation*). A fourth possibility is that information is stored in two or more models, but then it becomes apparent that they refer to the same situation. This information is combined into a single model (*model blending*).

Model creation is the simplest case. For example, in the *structure building framework* of discourse comprehension,[67] model creation occurs because the reader detects that incoming information in a text is relatively unrelated to the previous information. This could occur, e.g., if the protagonist moves from one location to another or if a new character enters the scene. In EST,[56] model creation occurs as the result of a transient increase in prediction error.

Event model elaboration is also relatively straightforward. There are two basic types of model elaboration. The first is the addition of new information that was not included in a prior version of the model, but that does not involve any change other than the addition of new components. For example, suppose a person creates a model to represent the event 'George was sitting on a bench in the park'. Later the person learns that 'George was wearing green overalls'. This information can be added to the model without changing any of the prior contents. In this case, property information is added to the token for George to indicate that he is wearing green overalls. Alternatively, elaboration may also occur when information is removed from a model and this removal does not alter other contents of the model. This information would not be involved in the functional, temporal, causal, or intentional relations among the entities in that event.

Model transformation occurs when some aspects of a prior model need to be altered because of the new information. It is probably the most complex type of updating. The changes involved are not minor in that there is likely to be some change in the functional, temporal, causal, or intentional relations among entities. These changes are not so drastic that a new model is needed, but more processing is needed than with model elaboration because the structure of the situation has undergone some meaningful change. Model transformation occurs when there has been a major change in an event, but it is still interpreted as being part of the same course of events. For example, a shift to a new spatial–temporal framework along a causal chain would correspond to model transformation. Alternatively, the introduction of a new goal or subgoal for a person would also correspond to model transformation.

The structure building framework[67] assigns a major role to model elaboration. As a reader proceeds through a text, new information is continuously mapped into the current situation model. Similarly, in Kintsch's *construction-integration* theory,[68] new information is incorporated into a model if it is associated to current information in the model. Importantly, associations allow information not explicitly mentioned to be incorporated. EST[50] departs significantly from these accounts in assigning a smaller role to model elaboration. Specifically, EST proposes that model elaboration occurs primarily immediately after model creation. Here is why: when a spike in prediction error occurs, the input gates on an event model are opened briefly, and then close over a brief window. As a result, elaboration ceases quickly and the content of an event model is dominated by the information encountered at the beginning of the event. EST makes this proposal for the sake of parsimony; whereas the structure building framework has two distinct mechanisms for shifting and model elaboration, EST makes do with only shifting. At first glance, it may seem counterintuitive that model elaboration should cease shortly after event model creation. However, patterns that appear like continuous event model elaboration can be approximated if the comprehender updates finer-grained event models while maintaining a coarse-grained event model.

The final kind of updating process, blending, occurs when information has been stored across two or more models, and the person realizes that it pertains to the same event.[4] In this case, these separate models are blended together to form a new, integrated model. This blending occurs through an alignment of information along the relevant dimensions. For example, if a reader reads that the protagonist of a story has crossed from one room to another, the reader may form two separate models. If the reader then learns that the activities in the two rooms form one unified causal sequence, the two models may be unified into a single model. This blending process is critically important in theories of logical reasoning.[4]

Model creation, model elaboration, model transformation, and model blending all should have unique behavioral signatures. Specifically, elaborating a model should lead to high availability of the elements just added to the model, but should not lead to large increments in processing time or decreases in the availability of prior information. Model transformation should lead to major increases in processing time because more of the event model is being changed. Furthermore, elements of the event model that were relevant prior to the updating may become less available. Finally, model blending should reveal behavioral signatures of two events now being treated as one. For example, this may result in a reduction in interference that was present before, with the person moving from having to coordinate two event models, to having the information integrated in one. To date, almost no research has been done to assess whether and when each of these forms of updating occurs.

In sum, there are a number of ways an event model may be updated. At one extreme, event models may be formed from whole cloth and rebuilt each time an event boundary is encountered. This is parsimonious but could be inefficient, insufficient, or both. At the other extreme, event models may be incrementally updated, revised, and combined. This is flexible and powerful—but perhaps too flexible and powerful, robbing event models of their explanatory force in cognitive theories. An important challenge for future research is to pin down what sorts of updating event models undergo.

CONCLUSION

In this article, we have tried to describe the nature of event models, the sort of structure they may have, and how they may be created and updated during online processing. The major components of an event model include the spatial–temporal framework, the entities involved in the event, as well as important or salient properties they may have, structural relations among entities in an event, and linking relations among events. New event models are created when the components of the current model no longer apply to the experienced situation. A possible mechanism for identifying such points is the monitoring of prediction error. Event models may be updated after they are created or simply replaced, and the relative role of each of these is currently unclear.

Events are a fundamental aspect of cognition. They give our thoughts and action purpose and are the basis of our intelligent understanding of the world. How we conceive of and represent events allows us not only to predict the likely future, and plan effectively, but also to imagine new possibilities. By better understanding how people conceive of and use their knowledge of events, we can be better positioned to identify when prediction, planning, and imagination will be more effective, and when they will have difficulty.

ACKNOWLEDGEMENTS

G.A.R. would like to acknowledge that this work was supported, in part, by funding through Sandia National Laboratories through Rob Abbott. J.M.Z. would like to acknowledge the support of the National Institute of Aging (RO1-MH70674). The ideas presented in this article have been developed through conversations with a number of people over the years, including David Copeland, Chris Kurby, Jackie Curiel, Katinka Dijkstra, Nancy Franklin, Art Graesser, Sabine Krawietz, Joe Magliano, Windy McNerney, Nicole Speer, Khena Swallow, Andrea Tamplin, Tom Trabasso, Barbara Tversky, Bob Wyer, Rose Zacks, and especially extended conversations with Rolf Zwaan. We would like to thank Hanisha Manickavasagan for her assistance with the preparation of the manuscript.

REFERENCES

1. Rosch E, Mervis C, Gray W, Johnson D, Boyes-Braem P. Basic objects in natural categories *Cogn Psychol* 1976, 8:382–439.
2. Barwise J, Perry J. *Situations and Attitudes*. Cambridge, MA: MIT Press; 1983.
3. Morrow DG, Bower GH, Greenspan SL. Updating situation models during narrative comprehension. *J Mem Lang* 1989, 28:292–312.
4. Johnson-Laird PN. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge MA: Harvard University Press; 1983.
5. van Dijk TA, Kintsch W. *Strategies of Discourse Comprehension*. New York: Academic Press; 1983.
6. Radvansky GA, Copeland DE, Zwaan RA. A novel study: the mental organization of events. *Memory* 2005, 13:796–814.
7. Wyer RS, Radvansky GA. The comprehension and validation of social information. *Psychol Rev* 1999, 106:89–118.
8. Zwaan RA, Radvansky GA. Situation models in language comprehension and memory. *Psychol Bull* 1998, 123:162–185.
9. White KG, Ruske AC, Colombo M. Memory procedures, performance and processes in pigeons. *Cogn Brain Res* 1996, 3:309–317.
10. Radvansky GA. Spatial directions and situation model organization. *Mem Cognit* 2009, 37:796–806.
11. Perrig W, Kintsch W. Propositional and situational representations of text. *J Mem Lang* 1985, 24:503–518.
12. Taylor HA, Tversky B. Spatial mental models derived from survey and route descriptions. *J Mem Lang* 1992, 31:261–292.
13. Borghi AM, Glenberg AM, Kaschak, MP. Putting words in perspective. *Mem Cognit* 2004, 32:863–873.
14. Nigro G, Neisser U. Point of view in personal memories. *Cogn Psychol* 1983, 15:467–482.
15. Graesser AC, Zwaan RA. Inference generation and the construction of situation models. In: Weaver CA, Mannes S, Fletcher CR, eds. *Discourse Comprehension: Essays in Honor of Walter Kintsch*. Hillsdale NJ: Erlbaum; 1995, 117–139.
16. Epstein R, DeYoe EA, Press DZ, Rosen AC, Kanwisher N. Neurological evidence for a topographical learning mechanism in parahippocampal cortex. *Cogn Neuropsychol* 2001, 18:481–508.
17. Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature* 1998, 392:598–601.
18. Anderson A, Garrod SC, Sanford AJ. The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Q J Exp Psychol* 1983, 35A:427–440.
19. Newtson, D. Foundations of attribution: the perception of ongoing behavior. In: Harvey JH, Ickes WJ, Kidd RF, eds. *New Directions in Attribution Research*, Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1976, 223–248.

20. Conway MA. Autobiographical memory. In: Bjork EL, Bjork RA, eds. *Memory*. San Diego: Academic Press; 1996.
21. Radvansky GA, Copeland DE, Berish DE, Dijkstra K. Aging and situation model updating aging. *Neuropsychol Cogn* 2003, 10:158–166.
22. Zwaan RA. Processing narrative time shifts. *J Exp Psychol Learn Mem Cogn* 1996, 22:1196–1207.
23. Barsalou LW. The content and organization of autobiographical memories. In: Neisser U, Winograd U, eds. *Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory*. Cambridge: Cambridge University Press; 1988.
24. Black JB, Bower GH. Story understanding as problem solving. *Poetics* 1980, 9:223–250.
25. Omanson RC. An analysis of narratives: identifying central supportive, and distracting content. *Discourse Processes* 1982, 5:195–224.
26. Radvansky GA, Copeland DE. Working memory and situation model updating. *Mem Cognit* 2001, 29:1073–1080.
27. Trabasso T, Secco T, van den Broek PW. Causal cohesion and story coherence. In: Mandl H, Stein NL, Trabasso T, eds. *Learning and Comprehension of Text*. Hillsdale NJ: Erlbaum; 1984, 83–111.
28. Gernsbacher MA, Goldsmith HH, Robertson RR. Do readers mentally represent characters' emotional states? *Cogn Emotion* 1992, 6:89–111.
29. Gernsbacher MA, Robertson RR. Knowledge activation versus sentence mapping when representing fictional characters' emotional states. *Lang Cogn Processes* 1992, 7:353–371.
30. Komeda H, Kawasaki M, Tsunemi K, Kusumi T. Differences between estimating protagonists' emotions and evaluating readers' emotions in narrative comprehension. *Cogn Emotion* 2009, 23:135–151.
31. Komeda H, Kusumi T. The effect of a protagonist's emotional shift on situation model construction. *Mem Cognit* 2006, 34:1548–1556.
32. Lutz MF, Radvansky GA. The fate of completed goal information in narrative comprehension. *J Mem Lang* 1997, 36:293–310.
33. Magliano JP, Radvansky GA. Goal coordination in narrative comprehension. *Psychon Bull Rev* 2001, 8:372–376.
34. Radvansky GA, Copeland DE. Functionality and spatial relations in situation models. *Mem Cognit* 2000, 28:987–992.
35. Franklin N, Tversky B. Searching imagined environments. *J Exp Psychol Gen* 1990, 119:63–76.
36. Bryant DJ, Tversky B, Franklin N. Internal and external spatial frameworks for representing described scenes. *J Mem Lang* 1992, 31:74–98.
37. Wasserman EA. Attribution of causality to common and distinctive elements of compound stimuli. *Psychol Sci* 1990, 1:298–302.
38. Fenker DB, Waldmann MR, Holyoak KJ. Accessing causal relations in semantic memory. *Mem Cognit* 2005, 33:1036–1046.
39. Zacks JM, Speer NK, Reynolds JR. Segmentation in reading and film comprehension. *J Exp Psychol Gen* 2009, 138:307–327.
40. Garrod S, Sanford AJ. On the real-time character of interpretation during reading. *Lang Cogn Processes* 1989, 1:43–59.
41. Shepard RN, Chipman S. Second-order isomorphism of internal representations: shapes of states. *Cognit Psychol* 1970, 1:1–17.
42. Zwaan RA. Effect of genre expectations on text comprehension. *J Exp Psychol Learn Mem Cogn* 1994, 20:920–933.
43. Stine-Morrow EAL, Gagne DD, Morrow DG, DeWall BH. Age differences in rereading. *Mem Cognit* 2004, 32:696–710.
44. Mandler JM. A code in the node: the use of a story schema in retrieval. *Discourse Processes* 1978, 1:14–35.
45. Rumelhart DE, Ortony A. The representation of knowledge in memory. In: Anderson RC, Spiro RJ, Montague WE, eds. *Schooling and the Acquisition of Knowledge*. Hillsdale NJ: Erlbaum; 1977.
46. Abelson RP. Psychological status of the script concept. *Am Psychol* 1981, 36:715–729.
47. Grafman J, Partiot A, Hollnagel C. Fables of the prefrontal cortex. *Behav Brain Sci* 1995, 18:349–358.
48. Graesser AC, Gordon SE, Sawyer JD. Recognition memory for typical and atypical actions in scripted activities: tests of a script pointer + tag hypothesis. *Journal of Verbal Learning & Verbal Behavior* 1979, 18:319–332.
49. Graesser AC, Nakamura GV. The impact of schemas on comprehension and memory. In: Bower GH, ed. *The Psychology of Learning and Motivation*, Vol. 16. New York: Academic Press; 1982.
50. Graesser AC, Woll SB, Kowalski DJ, Smith DA. Memory for typical and atypical actions in scripted activities. *J Exp Psychol [Hum Learn]* 1980, 6:503–515.
51. Nakamura GV, Graesser AC. Memory for script-typical and script-atypical actions: a reaction time study. *Bull Psychon Soc* 1985, 23:384–386.
52. Trafimow D, Wyer RS Jr. Cognitive representation of mundane social events. *Bull Psychon Soc* 1993, 64:365–376.
53. Speer NK, Swallow KM, Zacks JM. Activation of human motion processing areas during event perception. *Cogn Affect Behav Neurosci* 2003, 3:335–345.
54. Gibson JJ. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin; 1979.

55. Bingham GP, Wickelgren EA. Events and actions as dynamically molded spatiotemporal objects: a critique of the motor theory of biological motion perception. In: Shipley TF, Zacks JM, eds. *Understanding Events: From Perception to Action*. 2008, 255–286.
56. Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR. Event perception: a mind/brain perspective. *Psychol Bull* 2007, 133:273–293.
57. Kurby CA, Zacks JM. Segmentation in the perception and memory of events. *Trends Cogn Sci* 2008, 12:72–79.
58. Zacks JM, Sargent JQ. Event perception: A theory and its application to clinical neuroscience. In: Ross BH, ed. *The psychology of learning and motivation: Advances in research and theory, The psychology of learning and motivation: Advances in research and theory; 0079-7421*. Vol. 53. San Diego, CA US: Elsevier Academic Press; 2010, 253–299.
59. Zacks JM, Braver TS, Sheridan MA, Donaldson DI, Snyder AZ, Ollinger JM, Buckner RL, Raichle ME. Human brain activity time-locked to perceptual event boundaries. *Nat Neurosci* 2001, 4:651–655.
60. Zacks JM, Swallow KM, Vettel JM, McAvoy MP. Visual movement and the neural correlates of event perception. *Brain Res* 2006, 1076:150–162.
61. Speer NK, Reynolds JR, Zacks JM. Human brain activity time-locked to narrative event boundaries. *Psychol Sci* 2007, 18:449–455.
62. Whitney C, Huber W, Klann J, Weis S, Krach S, Kircher T. Neural correlates of narrative shifts during auditory story comprehension. *NeuroImage* 2009, 47:360–366.
63. Speer NK, Zacks JM. Temporal changes as event boundaries: processing and memory consequences of narrative time shifts. *J Mem Lang* 2005, 53:125–140.
64. Rinck M, Weber U. Who when where: an experimental test of the event-indexing model. *Mem Cognit* 2003, 31:1284–1292.
65. Swallow KM, Zacks JM, Abrams RA. Event boundaries in perception affect memory encoding and updating. *J Exp Psychol Gen* 2009, 138:236–257.
66. Ditman T, Holcomb PI, Kuperberg GF. Time travel through language: temporal shifts rapidly decrease information accessibility during reading. *Psychon Bull Rev* 2008, 14:750–756.
67. Gernsbacher MA. *Language comprehension as structure building*. Hillsdale: Elbaum; 1990.
68. Kintsch W. The role of knowledge in discourse comprehension: a construction-integration model. *Psychol Rev* 1988, 95:163–182.

FURTHER READING

Shipley TF, Zacks JM. eds. *Understanding events: From perception to action*. New York: Oxford University Press; 2008.