Event Perception: A Mind/Brain Perspective

Jeffrey M. Zacks, **Nicole K. Speer**, **Khena M. Swallow**, **Todd S. Braver**, and **Jeremy R. Reynolds**
Washington University in Saint Louis

Abstract

People perceive and conceive of activity in terms of discrete events. Here we propose a theory according to which the perception of boundaries between events arises from ongoing perceptual processing and regulates attention and memory. Perceptual systems continuously make predictions about what will happen next. When transient errors in predictions arise, an event boundary is perceived. According to the theory, the perception of events depends on both sensory cues and knowledge structures that represent previously learned information about event parts and inferences about actors' goals and plans. Neurological and neurophysiological data suggest that representations of events may be implemented by structures in the lateral prefrontal cortex and that perceptual prediction error is calculated and evaluated by a processing pathway including the anterior cingulate cortex and subcortical neuromodulatory systems.

Keywords

events; attention; memory; prefrontal cortex; orienting response

What is an Event?

The world as presented to human sense organs is continuous, dynamic, and fleeting. Yet people seem to perceive events as stable entities, to identify parts of events and their relations to other parts. For example, one might describe baking cookies by listing the following parts: "Preheating the oven, mixing the ingredients in a bowl, putting the dough on a cookie sheet…" This could reflect mere happenstance or accidents of linguistic structure, but a growing body of research suggests that talk of discrete events reflects a deeper psychological reality, that people perceive activity in terms of discrete events, that ongoing processing resources are devoted to this perceptual process, and that the on-line perception of events determines how episodes are encoded in memory (Zacks & Tversky, 2001). Thus, events are key components of perception, attention, and memory.

In this paper, we present a theory of the perception of everyday events, review psychological data that have informed the theory, and discuss possible neural substrates of the theory's components. We begin with the formal definition of an *event*: "a segment of time at a given

location that is perceived by an observer to have a beginning and an end" (Zacks & Tversky, 2001). This definition is useful, but surely does not exhaust the common conception of an event. The everyday notion of an event probably has a family resemblance structure, with some highly typical members such as weddings and breakfasts, and some atypical members such as the decay of a radioactive atom or the melting of a pond. Typical events seem to share some common features. They range from a few seconds (eating a strawberry) to a few hours (going for a hike). They are directed toward a goal; the goal of a wedding is to formalize a union, and the goal of breakfast is to sate one's hunger. Events involve animate agents, often human. These features, however, are neither necessary nor sufficient. Some events are quite short (cutting a ribbon) or quite long (World War II). Events that are natural occurrences, such as landslides, may lack both goals and animate agents. So, the taxonomic boundaries of the category "event" are fuzzy. The spatial and temporal boundaries of events also can be fuzzy – it is sometimes difficult to say where or when one event ends and another begins. Neither taxonomic fuzziness nor boundary fuzziness is a particular problem for the psychology of events, and both are exactly analogous to the psychology of objects. Here, we are concerned with the core of the category "event:" events that involve goal-directed human activity and are of modest duration (seconds to tens of minutes). (For comparative reviews of conceptions of events in psychology, see Stränger & Hommel, 1996; Shipley, in preparation.) In the next section, we present a theory of how human observers segment continuous activity into discrete events. In subsequent sections, we discuss evidence in support of this theory and its implications.

A Theory of Event Segmentation

Perception can be described as a roughly hierarchical process in which sensory information is successively transformed into representations that form the basis for action. Particularly important are representations of states of the world in the near future, which may be called *perceptual predictions*. Perceptual predictions are valuable because they allow an organism to anticipate the future and to plan appropriate actions rather than merely react to incoming stimuli. Such representations are critical for avoiding interception by predators, intercepting prey, and coordinating behavior with others. To the extent that information processing is hierarchical, perceptual predictions arise late in the processing hierarchy because incoming sensory information is transformed to generate predictions. In addition to being hierarchical, perception can be described as recurrent: Later processing stages affect the flow of processing in earlier stages. Finally, perception can be described as cyclical: Perceptual predictions are compared constantly to what actually happens and these comparisons are used to guide ongoing processing. These three notions – hierarchy, recurrence, and cyclicality – are working assumptions in many different theories of perception (Neisser, 1967), neurophysiology (Fuster, 1991; Carpenter, & Grossberg, 2003), and language processing (van Dijk, & Kintsch, 1983). They have been developed perhaps most fully in recurrent neural network models, which have been applied to word learning (Elman, 1990), to action learning (Jordan & Rumelhart, 1992), and to event perception (Hanson & Hanson, 1996).

The theory presented here, which we call *Event Segmentation Theory* (EST), shares these three properties. In this section, we describe the theory in information processing terms. Later in the paper, we recast the theory in terms of the neural systems that may implement these information processing components.

Architecture and Principles

EST proposes that event segmentation arises from the perceptual processing stream depicted in Figure 1. Its core is a pathway whose input is a set of sensory representations and whose output is a set of perceptual predictions. The sensory inputs correspond to the information conveyed by the peripheral nervous system to the cortex. In the visual modality, for example, this corresponds to basic information about brightness, color, and possibly some preliminary

edge extraction. Sensory inputs are transformed by perceptual processing to produce multimodal representations with rich semantic content, encoding information such as object identity and location, motion trajectories, and the identities and attitudes of other people. According to the theory, processing is oriented in time such that it results in predictions about the future state of perceptual representations. For example, extracting a motion contour leads to predictions about the future locations of objects and inferring the goals of a person leads to predictions about his or her future movements.

We propose that perceptual processing is guided by a set of representations called *event models* that bias processing in the perceptual stream. An event model is a representation of "what is happening now," which is robust to transient variability in the sensory input. The stability of event models over time is a source of perceptual constancy; an ongoing event is a single entity despite potential disruptions in sensory input such as occlusion or distraction. In this regard, event models are similar to the *object files* proposed to mediate object constancy in visual perception (Kahneman, Treisman, & Gibbs, 1992) or the short-term action representations proposed to mediate perceptual constancy in biological motion (Stränger & Hommel, 1996, see section 5). However, event models are hypothesized to be active over much longer time frames than object files or biological motion representations. In terms of the current theory, object files and biological motion representations are hypothesized to be part of high-level perceptual processing components.

Event models are working memory representations, which are implemented by transient changes in neural activation rather than long-term changes in synaptic weights. They are not necessarily accessible to consciousness, though people may have partial awareness of their contents under some circumstances. Event models are multimodal, integrating information from visual, auditory, and the other sensory modalities. In these regards they are akin to the representations recently proposed by Baddeley as forming an *episodic buffer* (Baddeley, 2000). Most of the time, the contents of event models are insensitive to immediate sensory and perceptual input, providing a stable representation of the current event to guide perceptual processing. This is indicated by the gated arrow terminating on the event models component in Figure 1. Event models also receive input from semantic memory representations that capture shared features of previously encountered events: *event schemata*. Event schemata contain previously learned information about the sequential structure of activity. Unlike event models, event schemata are implemented by permanent synaptic changes. The information they store includes distinctive physical features such as object and actor movement, statistical information about which patterns of activity are likely to follow a given pattern, and information about actors' goals.

The quality of perceptual prediction depends critically on whether one's current event models are a good fit to what is actually happening. Prediction quality is evaluated by an error detection mechanism that compares the perceptual processing stream's predictions to what actually happens in the world. Most of the time, the event models represent the current state of events well and perceptual prediction is easy and accurate. From time to time, however, activity becomes less predictable and the current contents of the event models become less useful for perceptual prediction. At these points, prediction error increases. EST proposes that at these points a gating mechanism detects these transient increases in prediction error and reacts to them by updating the event models. Updating consists of (1) resetting the current representations (indicated by the dashed black line in the figure) and (2) transiently increasing the influence of the pathway from sensory inputs to the event models. Together, these operations drive the event models into a new stable state. As the event models are updated, prediction error typically improves and the influence of sensory inputs on the event models diminishes. The signal pathways into the event models can be thought of as controlled by a gate that swings open in response to increases in error prediction and then quickly swings shut

again. Thus, the system alternates between long periods of stability and brief periods of change. Periods of stability are perceived by observers as events and periods of change are perceived as the boundaries between events.

The contents of event models are determined by a combination of bottom-up and top-down processing. When the event models' sensory inputs are transiently opened, they receive information from the current state of sensory and perceptual representations in a bottom-up fashion. Event schemata affect event models in a top-down fashion. (We propose that the influence of event schemata on event models is continuous and unaffected by the gating mechanism. However, this claim is based largely on parsimony and may need to be revised in the future.) At the same time that schemata influence the current contents of the event models, the event models' contents update the event schemata through a slow incremental learning process.

Event models are active and accessible representations of the events that are currently underway, yet the amount of information contained in these models almost certainly exceeds what can be actively maintained in a limited capacity working memory system. The effective capacity of working memory can be augmented by the efficient use of previously stored knowledge representations. Ericsson and Kintsch (1995) proposed the construct of *long-term working memory* to describe these interactions. Information that is needed to perform skilled activities is rapidly encoded into long-term memory using knowledge structures that anticipate future retrieval demands. This information remains readily accessible as long as some part of it is available in short-term memory and can be rapidly reinstated once some part of the structure is retrieved. Importantly, developing effective long-term working memory for information in a particular domain is a skill that requires repeated practice with that domain. We presume that most human adults have had extensive experience with a wide range of events. Thus, event schemata help expand the effective capacity of event models by storing predictive information about the future relevance of certain aspects of events.

Here is an example of how the mechanism in EST might work when one observes an everyday activity: Imagine watching a man wash dishes. He takes one plate from a pile next to the sink, scrapes food from it, and then places it in the sink. He does the same with a second plate. At this point, a number of cues make it likely that he will continue to scrape the plates. First, continuing to scrape would maintain a coherent movement pattern. Second, it would be consistent with previous observations in which it was statistically likely that scraping a plate was followed by more scraping of plates. Third, the observer might infer that the man had the goal of scraping all of the plates. Thus, for the duration of the plate-scraping activity, affairs would be predictable. However, when the man scraped the last plate, things would become less predictable. The coherent movement pattern would cease, the statistical dependency would be broken, and the inference of the actor's goal to scrape all of the plates would no longer have predictive value. At this point, perceptual prediction would decline, leading to the activation of the gating mechanism and updating of the event model.

A Partial Computational Implementation

Reynolds and colleagues have developed a neural network simulation that implements the core features of EST: perceptual prediction, activation-based event models, and error-based gating of those models (Reynolds, Zacks, & Braver, in press). The network is presented with animations of a human actor performing simple actions (e.g., jumping jacks), represented as the three-dimensional location of 18 points on the actor's body (see Figure 2). At each timepoint, the network attempts to predict the actor's body position at the next timepoint. The network is trained on a corpus of stored events, such that each event is presented from start to finish, but each event can be followed by any randomly chosen event from the corpus. The network has pools of units corresponding to sensory inputs, perceptual processing, predicted

future inputs, and event models as described in EST (see Figure 1). (Event schemata are not implemented in the simulation.) The pathway from sensory inputs to predicted future outputs is fully connected in a feed-forward fashion. The event model units have connections from sensory inputs that are gated; they open only at transient increases in prediction error. Simulations using this network provide support for the basic architecture of EST: Throughout training, boundaries between individual events are associated with larger prediction errors than within-event timepoints. Further, the network uses these transient spikes in prediction error to update stable representations differentiating among possible events. These event representations improve performance on the prediction task. Importantly, with appropriate gating, the event model representations can self-organize without any explicit labeling or categorization of the events. Thus, the proper gating and maintenance of event models aids perception.

Multiple Timescales

Note that the dish-scraping example and the Reynolds et al. (in press) network focus on one timescale, in which activity becomes less predictable at the end of the dish-scraping or at the end of the individual events in the network's corpus. However, variations in predictability are to be expected on finer and coarser timescales as well. For example, it is likely that as each dish-scraping comes to an end there is a small, brief increase in prediction error, and that as the dishwashing activity comes to an end there is a larger, longer increase in prediction error. We hypothesize that the architecture in EST is implemented simultaneously on a range of timescales, spanning from a few seconds to tens of minutes. For each timescale, the error signal is integrated to provide a reset signal tuned to the appropriate grain. Fine-grained representations are tuned such that they can be updated in response to small, brief increases in prediction error. For coarse-grained representations, the error signal is integrated with a longer time constant such that resets happen only in response to larger, more sustained increases in error.

In addition to simultaneous parsing on multiple timescales, it is possible that event segmentation sometimes tracks simultaneous activities in parallel. For example, when attending a child's birthday party one might simultaneously segment the actions of children playing a birthday game, and those of parents having a conversation at the same time. It is an empirical question whether such parallel processing occurs or whether observers are limited to processing one activity stream at time.

Cognitive Control

According to EST, the segmentation of activity into events happens on an ongoing basis and plays two general and central roles in regulating perception and cognition. First, event segmentation controls the allocation of cognitive resources over time. During periods of low prediction error, the pathway from sensory inputs to event models is inactivated and the event models are stable, which conserves resources. More intensive processing occurs transiently when prediction error increases and the event models are reset. This regulation of resources over time can be viewed as a form of attention, focusing processing resources adaptively at those moments when incoming sensory information is most behaviorally relevant (Coull & Nobre, 1998; Nobre, 2001). In other words, event segmentation does not itself require attention; rather, it implements a mechanism of attention. Second, event segmentation controls the updating of information in working memory by resetting the event models. The term *cognitive control* has been used to describe the control of attention and working memory in a variety of task domains (Cohen, Braver, & O'Reilly, 1996; Posner & Snyder, 1975). In our view, EST's proposal that event segmentation controls resource allocation and updates memory is a claim that event segmentation is a core, domain-general mechanism of cognitive control.

Related Approaches

Our proposal that event models maintain stable representations that influence perceptual processing and are updated in response to errors of prediction is similar to proposals of several neural network models. In these networks, transient updating based on failures of prediction is a means of balancing stability and flexibility: Representations need to be stable across moment-to-moment fluctuations in perceptual input, but need to be updated when they are no longer appropriate. These considerations have played a large role in the development of *adaptive resonance theory* (ART, Grossberg, 1999; Carpenter & Grossberg, 2003). In ART networks, perceptual input is "cleaned up" by recurrent interactions with a stable high-level representation. This high-level representation forms an abstraction of the perceptual input over some period of time. When the cleanup process leads to large distortions of the perceptual input, the current, high-level representation is deemed no longer appropriate and a search is initiated for a new high-level representation. Our proposal differs from the one instantiated in ART networks most significantly in that error comparison is made between a predicted perceptual state and the actual perceptual input. In ART networks, there is no intrinsic orientation of the processing stream with respect to time.

Similar concerns motivate the architecture of several recent models of prefrontal cortex (PFC). In these models, PFC represents current goals for action and the means to achieve them. Goal representations must be stable until the goal is achieved (or blocked) in order to be effective; however, once a goal is no longer relevant, a new goal representation is desirable. In a broad theoretical review, Miller and Cohen (2001) suggest that updating of PFC representations could be gated by phasic dopamine signals from midbrain dopamine neurons (Braver & Cohen, 2000), triggered by encountering unexpected rewards; other models address the possibility that unexpected *lack* of reward may also trigger the updating of memory representations (Rougier & O'Reilly, 2002; O'Reilly, Noelle, Braver, & Cohen, 2002; Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005). EST differs from this family of models in two regards. First, stable representations contain information about the state of the world rather than goals and the means to achieve them. Second, gating in this family models is based on failure to predict the reward value of a situation, whereas gating in EST is based on failure of perceptual prediction.

Finally, a model of frontal cortex proposed by Frank and colleagues also includes stable representations that are occasionally updated via a gating mechanism (Frank, Loughry, & O'Reilly, 2001; Frank, Seeberger, & O'Reilly, 2004). However, in this model, the gating mechanism is implemented by loops through frontal cortex, the basal ganglia and thalamus, with subcortical structures modulating the excitability of frontal cortex and thereby determining whether it is stable or plastic. This approach, like the Miller and Cohen (2001) model and related proposals, focuses on representations required for action rather than for perception. Unlike dopamine accounts, this model provides a natural mechanism for selectively updating some representations without disrupting others. However, in the context of event perception, a more global signal might be more appropriate. In EST, we propose that the complete representation of an event at a given timescale is updated based on a relatively global signal, such as could be provided by the catecholamine neurotransmitters dopamine and norepinephrine. By varying the time constant over which this signal is integrated, selective updating of fine-grained or coarse-grained event models can be accomplished. EST thus differs from the model proposed by Frank and colleagues in three substantial ways: first, by basing the gating mechanism on the failure of perceptual prediction rather than on reward; second, by characterizing perception rather than action; and third, by adopting a less selective updating mechanism more compatible with implementation by neuromodulatory neurotransmitters than by corticothalamic loops.

Implications

The principal novel features of EST are that event models maintain stable representations of "what is happening now" and are updated based on transient increases in perceptual prediction error. The theory has several implications for perception and cognition:

1. Most importantly, the theory implies that the segmentation of ongoing activity into discrete events is a spontaneous concomitant of ongoing perception and does not require conscious attention.

2. Event segmentation is a mechanism of cognitive control. The gating mechanism resets event models and thus is the means by which the cognitive system exerts control over the disposition of processing resources and the updating of working memory.

3. Event segmentation happens simultaneously on multiple timescales, though an observer may attend to a particular timescale.

4. Event segmentation incorporates information from multiple senses. This results from the fact that the mechanisms in EST are general across sensory modalities and incorporate information from multiple modalities.

5. Event segmentation depends on change. When the world is static, prediction is easy.

6. Event segmentation depends on prior knowledge. Event models are constructed via the interaction of sensory input with stored knowledge (including the knowledge stored in event schemata).

In the following two sections, we review research on the cognitive and neural correlates of event segmentation with these implications in mind. The first section ("Causes and Consequences of Event Segmentation") describes behavioral data that provide support for the model. The second section ("Neural Correlates of Event Perception") describes proposals for how the nervous system may implement the information processing model and a review of the relevant neuropsychological and neurophysiological data that motivate these proposals.

Causes and Consequences of Event Segmentation

Perceptual Measures of Event Segmentation

The first question that comes up in asking how people perceive temporal structure in events is: How can one measure it? Newtson (1973) introduced a simple and surprisingly powerful solution to this problem, which he dubbed *unitization* (for reviews, see Newtson, 1976; Stränger & Hommel, 1996; Zacks & Tversky, 2001). Participants are asked to watch movies of everyday activities (such as a person filling out a questionnaire) and to press a button whenever they judge that a boundary between successive events has occurred. We will use the term "unitization" to refer to this task and to distinguish it from event segmentation, which we hypothesize is an ongoing perceptual process that is independent of any intentional task. When performing a unitization task, participants may be asked to identify the largest events they find meaningful (coarse-grained unitization) or the smallest (fine-grained unitization). With a little bit of practice, people generally have no problem following the instruction and this simple task has produced rich and replicable phenomena. By two measures, reliability of unitization is good. First, participants show good agreement regarding the location of event boundaries (Newtson, 1976). Second, the differences that do exist among observers can be attributed, in part, to stable individual differences rather than to noise: test-retest studies have found good reliability both in the length of the units people identify (Newtson, 1976) and in the particular locations they mark as event boundaries (Speer, Swallow, & Zacks, 2003). The reliability of the unitization procedure provides support for EST's claim that event segmentation is a spontaneous concomitant of ongoing perception; the procedure corresponds intuitively to replicable aspects of observers' ongoing experience.

The original interest in event unitization concerned social cognition: How does the grain at which people segment an activity affect observers' attributions about why actors perform particular actions? Newtson (1973) asked participants to segment movies of a man filling out a questionnaire or building a model molecule into either fine-grained or coarse-grained events and then to make judgments about the activity that had been performed. When observers were asked to segment activity into fine-grained events, they were more likely to draw conclusions about actors' permanent traits (e.g., personality characteristics). They also formed impressions of the actors' traits that were more differentiated and were more confident in their judgments. This association between fine-grained units and dispositional attributions was supported correlationally in one subsequent study (Wilder, 1978b), but not in another (Wilder, 1978a). The grain at which observers segment events also affects the dispositional attributions observers make. In one study, participants judged an actor as more likable if they had unitized her actions into fine-grained units rather than coarse-grained units (Lassiter, 1988).

If people form systematically different impressions when they attend to fine-grained than coarse-grained events, one reasonable possibility is that observers adaptively modulate the grain at which they segment activity in response to the needs of the situation. Newtson (1976) proposed that fine-grained unitization is more resource-demanding than coarse-grained unitization and that observers unitize at the coarsest grain they can sustain while maintaining a coherent representation of the ongoing activity. Activity that is relatively coherent and predictable can be parsed at a coarse grain, whereas activity that is confusing or surprising must be parsed at a finer grain. This assumes that observers can only perceive boundaries at one grain at any time, in which case it is intuitive that identifying few units would be less resource-demanding than identifying many units. Evidence for this proposal has accrued from studies in which a single surprising action was inserted into an otherwise predictable activity (Newtson, 1973) and from studies in which the overall predictability of the activity was manipulated more systematically (Wilder, 1978a;Wilder, 1978b).

EST provides an alternative account of these findings. The theory asserts that people do not perceive event boundaries on only one timescale. Rather, they *perceive* event boundaries on multiple timescales simultaneously, but selectively *attend* to one timescale in response to instructions or other experimental manipulations. According to this view, when activity is coherent participants segment it at multiple timescales, and can choose to attend to finer or coarser grains. Attending to coarser grains may be preferred because it reduces the frequency of decision-making and button-pressing. However, when activity is less coherent coarse-grained event segmentation may break down, because prediction error will be uniformly high on coarse timescales, leaving only fine-grained segmentation intact. Data from experiments in which people segment the same activities multiple times at different timescales support this hypothesis (Zacks, Tversky, & Iyer, 2001; Speer, Swallow, & Zacks, 2003; Lozano, Hard, & Tversky, in press; Hard, Lozano, & Tversky, in press; Hard, Tversky, & Lang, in press). In these experiments, participants segmented activities at both a coarse and a fine grain, on different viewings. If viewers segmenting at a fine grain were spontaneously grouping fine-grained events into larger units, one would expect coarse-grained event boundaries to be a subset of fine-grained event boundaries. As a result, one would expect coarse and fine event boundaries to be aligned such that each coarse boundary would be placed close to a fine boundary. This pattern has been observed in several studies (Zacks, Tversky, & Iyer, 2001; Speer, Swallow, & Zacks, 2003; Hard, Tversky, & Lang, in press). One also would expect that coarse-grained events would tend to enclose a set of fine-grained events, such that coarse boundaries would follow rather than precede the fine boundary to which they were closest; this has also been observed (Hard, Lozano, & Tversky, in press; Lozano, Hard, & Tversky, in press). The presence of relations in segmentation that span time-scales indicates that the coarse-grained segmentation occurred even when participants were attending to fine-grained segmentation.

Consequences of Event Segmentation for Long-Term Memory

According to EST, working memory representations (the event models) are updated selectively at those points in time that correspond to perceptual event boundaries. This means that perceptual information at those times receives more extensive processing than perceptual information from other points in time. This extra processing should result in better long-term memory for this information. Several converging measures suggest that perceptual information from event boundaries is preferentially accessible in long-term memory. In one experiment, participants watched a movie and then saw still pictures from event boundaries, points in between boundaries, and from a similar movie they had not seen. They then reported which pictures came from the movie they had watched (Newtson & Engquist, 1976). Accuracy was higher for pictures taken from event boundaries. Another study found that descriptions of event boundaries from memory were richer and more detailed than descriptions of nonboundaries (Schwan, Garsoffky, & Hesse, 2000). Also, the overall amount of information that can be recalled is affected by the grain of segmentation: Multiple studies have reported that fine-grained unitization of movies leads to more detailed recall of the events depicted than does coarse-grained unitization (Lassiter, Stone, & Rogers, 1988; Lassiter, 1988; Hanson & Hirst, 1989). However, it is currently a matter of debate whether fine-grained unitization also improves recognition memory for depictions of events (Hanson & Hirst, 1991; Lassiter & Slaw, 1991). These effects are consistent with the view that observers can selectively attend to one timescale while segmenting simultaneously at several timescales.

According to the theory, when the surface structure of an event depiction aligns with its underlying event structure the gating mechanism should operate efficiently, resulting in enhanced long-term recall for the events. Conversely, surface cues that conflict with the event structure of an activity may lead to poorer memory. Three studies have used film editing techniques to test these hypotheses. In the first study (Schwan & Garsoffky, 2004), participants viewed short movies of everyday events and then recalled them. The movies were presented either intact, with deletions that corresponded to intervals surrounding event boundaries, or with deletions of intervals in between event boundaries. Memory for the edited movies with preserved event boundaries was as good as memory for intact movies, but memory for edited movies with deleted intervals around event boundaries was poorer. The second study examined longer events and manipulated cues to event structure by marking the event boundaries rather than deleting actions (Boltz, 1992). In this series of experiments, participants viewed a feature film with no commercial breaks or a film with commercial breaks. The breaks corresponded to or conflicted with event boundaries. Commercials at event boundaries improved later recall of the activity, whereas commercials at nonboundaries impaired memory. Commercial breaks at event boundaries also improved memory for the temporal order of events. One other study (Schwan, Garsoffky, & Hesse, 2000) manipulated the placement of cuts between different camera positions in short movies. Cuts at event boundaries improved memory for those points in time, but had little effect on overall memory performance. It is notable that the placement of cuts in the Schwan et al. study produced relatively weak effects compared to the placement of commercials in the Boltz study. One possibility is that simple cuts are not sufficiently salient to affect the operation of the gating mechanism – at least for participants used to viewing movies and television. Another possibility is that the effects of event segmentation on memory are more pronounced at longer timescales. Together, these results converge in supporting a role for event segmentation in long term memory encoding.

Patterns of individual differences suggest a strong relation between one's ability to segment activity during learning and one's ability to recall it later. In one study (Zacks, Speer, Vettel, & Jacoby, in press), older adults with and without dementia of the Alzheimer's type (DAT) segmented movies of everyday events and then performed a recognition memory test for visual details from the movies. An individual's ability to properly segment the movies was evaluated

by comparing her or his segmentation to the segmentation of the group as a whole. Group-typical segmentation was found to be a unique predictor of later memory, even after the presence of dementia and overall cognitive level were controlled. This is consistent with EST's proposal that event boundaries receive richer processing and thus that identifying the proper boundaries results in more effective encoding for long-term memory.

The theory's proposal that event boundaries receive differential processing during perception has implications for the ability to use perceptual experiences when learning new skills. The ability to segment an activity into the right units should be valuable in learning how to perform the actions that occur during the activity. Several studies of procedural learning from events support this hypothesis. In one series of experiments, participants learned from a movie how to assemble a TV cart or a construction-block model (Hard, Lozano, & Tversky, in press). During the learning phase, they segmented the movie into events; afterwards, they built the TV cart or the model. A number of features of the situation were varied: whether the participants segmented at a coarse grain, a fine grain, or both (and, if so, in what order); whether the participants described the activity while segmenting, and whether they were explicitly instructed to group fine-grained events into larger structures. Across experimental conditions, a consistent pattern was observed: Those individuals whose segmentation was more hierarchically structured were better at assembling the TV cart than those whose segmentation was less hierarchically structured. This was true across experimental conditions, and across individuals within conditions. Participants' descriptions suggested that the benefit of hierarchical encoding was mediated by a tendency for those participants who segmented hierarchically to simulate the events from the actor's point of view. A subsequent series of experiments experimentally manipulated perspective-taking and supported the hypothesis that hierarchical encoding facilitated simulating the actor's perspective (Lozano, Hard, & Tversky, in press).

In a final procedural learning study (Zacks & Tversky, 2003), participants learned about an everyday procedure such as putting together a musical instrument using a computer program. The program was designed to visually depict a set of pre-supposed coarse-grained event boundaries in the procedure or to conflict with those boundaries. In one experiment, event boundaries were identified by asking a separate group of participants to segment a movie of the activity to be taught. In this experiment, visual structure that coincided with these boundaries improved memory for the order of events. In another experiment, event boundaries were taken from manufacturer's instructions, which may have conflicted with perceptually natural event boundaries. In this experiment, reinforcing the event boundaries *reduced* memory for the order of events. These results suggest that memory is facilitated when surface cues support the intrinsic event structure of an activity, and memory is impaired when surface structure conflicts with the intrinsic event structure.

Together, studies of long-term memory for events and studies of learning procedures from events suggest that event boundaries are used to structure memory encoding. The fact that event boundaries are remembered better than non-boundaries supports EST's claim that event segmentation is a cognitive control mechanism. Further support for this claim is that individuals who are good at segmenting events remember them better than individuals who are poor at segmenting. Moreover, cuing the appropriate event structure facilitates memory, whereas miscuing event structure impairs memory. Finally, people consistently segment activity hierarchically, making connections that span timescales. This supports the claim that event segmentation proceeds simultaneously on multiple timescales.

Features That Correlate With Perceptual Event Segmentation

According to EST, event segmentation depends on changes in the environment and on prior knowledge. Physical changes in the environment can drive event segmentation in a bottom-up

fashion by increasing the potential for prediction error. When sensory inputs are constant, a well-tuned perceptual prediction system will adhere to the old adage about weather forecasting – "tomorrow will be the same as today" – and will be correct in this prediction. When changes occur there is more opportunity for error. At the same time, prior knowledge can influence event segmentation in a top-down fashion by biasing the system's interpretation of what a change portends. One type of prior knowledge can be tied to domain-specific reasoning about goals and plans. When watching goal-directed human activity, observers may infer an actor's goal based on past experience or explicit instructions and use this information to bias perceptual prediction. For example, baseball fans can anticipate events that might come as a surprise to baseball novices (e.g., all the players running off the field at the end of an inning) and this is likely to influence event segmentation. Another type of prior knowledge involves domain-general statistical learning. Humans have powerful mechanisms for learning sequential dependencies in streams of information, which can influence motor performance (Seger, 1994) and language learning (e.g., Brent, 1999; Saffran, Aslin, & Newport, 1996). Similar mechanisms may guide perceptual prediction. According to the model, prior knowledge about goals and plans and prior statistical learning are encoded in event schemata. A small body of evidence indicates that both physical changes and prior knowledge are correlated with the perception of event boundaries.

One distinctive physical change that correlates with the perception of event boundaries is movement. In one study, participants unitized brief movies and the experimenters identified those points that most observers agreed were event boundaries (Newtson, Engquist, & Bois, 1977). The movies then were coded using a dance notation that provided a discrete representation of which actor joints had changed position by more than 45° during each one-second interval. The number of joints that changed between successive intervals is a rough index of the amount of motion at that point in the movie. The number of changing joints was averaged at each grain for transitions into and out of event boundaries and for successive intervals within an event. Transitions into and out of event boundaries had statistically greater numbers of changes (i.e., more movement) than successive within-event intervals. Converging evidence for a relation between movement features and segmentation comes from a recent study in which participants segmented animations of geometric objects playing simple games (Hard, Tversky, & Lang, in press). The animations were coded by eye for movement change such as stops, direction changes, and changes in speed. These movement changes were associated with increases in both fine-grained and coarse-grained segmentation.

A pair of studies (summarized in Baird & Baldwin, 2001) suggests that event segmentation is correlated with actors' goals and plans. In one study, one group of participants was asked to segment movies of everyday activities based on when the actor completed a goal. A second group of participants watched sequences taken from these movies with brief tones placed either at points when goals were completed or at midpoints between goal completions. After each sequence, they were asked to watch it again and press a button to mark where the tone had occurred. Tones were remembered more accurately when they were placed at goal completions, whereas midpoint tones were remembered as having occurred closer to completions than they actually did. These effects on memory suggest that participants perceived the activity in terms of intention-based units and this affected how the locations of tones were stored during the initial encoding phase. However, an alternative explanation is that knowledge structures representing goals biased performance only during the retrieval phase of the task. More direct evidence that goal completions are perceived as event boundaries comes from a second study, conducted with 10–11 month old infants. In this experiment, the infants were repeatedly exposed to a sequence taken from a movie of an everyday activity. They then were tested with one of two altered versions of the sequence. In one version, a pause was inserted at a moment that had been identified as the completion of a goal; in the other, a pause was inserted in the middle of two completions. The experimenters hypothesized that if the infants encoded the

activity in terms of the actor's intentions, a pause in the middle of accomplishing a goal would be more surprising than a pause at the end of completing a goal. Thus, they should look longer at the version with a pause in the middle. This was exactly what the experimenters found: Infants looked longer at the altered sequences when those sequences contained a pause that interrupted an ongoing goal compared to when the pause occurred at a goal completion.

A recent study provides evidence that both physical changes and goals are systematically related to event segmentation (Zacks, 2004). In three experiments, participants viewed animations that showed the movements of a pair of objects (see Figure 3a), and unitized them to mark coarse-grained or fine-grained units. Two types of animation were shown: *goal-directed activity* stimuli were generated by recording the actions of two people controlling the objects in a simple video game and *random activity* stimuli were generated by a random process designed to match the velocity and acceleration of the video game stimuli. Rather than hand-coding movement features of interest to the experimenters (Hard, Tversky, & Lang, in press; Newtson, Engquist, & Bois, 1977), a quantitative movement analysis was performed. For each animation, an exhaustive set of movement features was calculated, including the position, velocity, and acceleration of each object, the distance between the objects, their relative velocity, and their relative acceleration. For each experimental condition, the relation between the movement features and the probability that a participant would identify an event boundary was characterized using linear models. Movement features were significantly related to event unitization for all conditions, providing evidence that distinctive physical features play a role in event segmentation. Two further results suggested that the processing of movement information interacted with goal processing to determine event segmentation. First, the relation between movement features and unitization was consistently stronger for fine-grained unitization than coarse-grained unitization (see Figure 3b). This suggests that movement features may play a particularly strong role in identifying the smallest units of activity, but that other features may be important in identifying which of those low-level event boundaries are also boundaries between larger units of activity. Second, the relation between movement features and unitization was weaker for goal-directed activity than for random activity. This suggests that prior knowledge may play a role in modulating how physical characteristics are processed in order to identify event boundaries.

Knowledge about actors' goals is one source of prior knowledge that can affect event segmentation. Simple statistical information is another complementary source of prior knowledge, which is also related to the perceptual segmentation of events. One series of studies indicated that similar mechanisms may play a role in the identification of events and their boundaries (Avrahami & Kareev, 1994). In three experiments, participants viewed movies made by concatenating short clips from cartoon films. In the first experiment, the experimenters constructed a sequence of clips that included a perceptually salient change. These changes were identified by untrained participants as a part boundary. A second group of participants viewed a stream of clips that contained repeated presentations of the sequence. They were then shown the sequence in a short stream of clips and were asked where they thought it should be divided. These participants placed the boundary not at the salient change, but at the end of the sequence. A second experiment showed that participants were more likely to recognize a sequence that repeated throughout a movie if the clips appearing before and after the sequence were varied. A final experiment showed a similar effect using a recall paradigm. These data suggest that the repetition of an arbitrary sequence across different contexts is sufficient for it to be conceived of as a coherent unit. However, it is important to note that in all three experiments the tasks depended heavily on memory, so the extent to which these data speak to perception, as such, is not clear.

To summarize, studies of the unitization of events provide evidence consistent with EST's claims that event segmentation depends on change and that event segmentation depends on

prior knowledge. The fact that event boundaries are correlated with changes in movement features in two quite different paradigms indicates that event segmentation depends on change. The fact that event boundaries are correlated with goals and with sequential statistical structure indicates that event segmentation depends on prior experience.

Causes and Consequences of Segmentation in Story Understanding

Narrative stories depict events through a medium that is surely not equivalent to the actual experience of those events; however, there is good reason to think that narrative comprehension may share representations and processes with the comprehension of live events. Recent theories of narrative comprehension claim that readers and listeners mentally simulate an experience based on the description provided in the text (e.g., Glenberg & Kaschak, 2002; Zwaan, Stanfield & Yaxley, 2002). From a researcher's point of view, narratives offer opportunities for quantification and control over some features that may be important for event segmentation, so the available data provide important insights into event understanding.

In narrative comprehension, the effects of prior knowledge on segmentation have been studied under the rubrics of schemata (Rumelhart, 1975), scripts (Schank & Abelson, 1977), and situation models (van Dijk & Kintsch, 1983). At least since Bartlett's (1932) classic studies in the early 20th century, experimental psychologists have argued that readers understand narratives by constructing representations of the events described in the text and later remember these constructions rather than simply remembering the text itself. Schemata, scripts, and situation models differ in many particulars, but they share a common assumption: Readers use prior knowledge to segment a narrative into discrete events. Prior knowledge may include information about actors' goals and purely conventional (statistical) sequential dependencies. Our use of the term "event schema" is closely related to the usage described here and our use of the term "event model" is analogous to the term "situation model" used in the narrative comprehension literature.

The term "schema" was originally introduced in neurology to describe the coordination of movement in goal-directed action and adopted by Bartlett (1932) to describe structured representations of the actions that typically happen in stories. The notion of schema in cognitive psychology was developed more fully by Rumelhart and colleagues (e.g., Rumelhart, 1980). Rumelhart (1977) argued that stories consist of episodes; in each episode, a protagonist is confronted with a situation that causes him or her to desire some goal and the protagonist tries to achieve that goal. Thus, each episode is interpreted in terms of a schema for trying to achieve a goal. This schema has a recursive structure, such that components of trying to achieve a goal (e.g., buying an ice cream) include subgoals that the protagonist tries to achieve in order to fulfill the larger goal (e.g., getting money). When comprehending a story, readers or listeners encode it in terms of a series of events that are segmented and hierarchically arranged in accord with this goal-based structure. According to schema theory, summarizing a story consists of pruning low levels in the hierarchical representation. The theory also predicts that schematic representation can lead to two kinds of distortion in delayed recall. First, pruning can occur because high levels in the schema are better represented than low levels, leading recall to increasingly resemble summarization. Second, distortions can occur that normalize recall to the participant's schema for that type of activity. Experimental tests of these proposals supported the theory (Rumelhart, 1977).

The term "script" was introduced in computer science and psychology by Schank and his colleagues (Schank & Abelson, 1977), who used it to refer to a structured representation of an activity that has predictable relations among settings, actors, props, and actions. As with event schemata, scripts are knowledge structures with a nested organization such that the contents of a script include other knowledge structures. A script specifies a list of the events that are typically part of the activity represented by the script. For example, a script for visiting the

doctor might include signing in with the receptionist, filling out forms, waiting in the waiting room, and being examined by the doctor. Scripts may include information about the order in which these events occur (Abelson, 1981). Script theory motivated a number of studies of reading comprehension and memory. One set focused on readers' knowledge about the everyday activities that might be represented by scripts (Bower, Black, & Turner, 1979). These experiments established that people agree about the typical actions performed in common everyday activities and about the boundaries between events in a script-based narrative. Moreover, after reading a story, people remember the events and their order as having been more similar to the standard script for the activity than was actually described in the story. Another set of studies provided evidence for the hierarchical organization of stories in memory (Abbott, Black, & Smith, 1985). Participants read stories based on scripts and their memories were tested. In one experiment, reading about low-level (fine-grained) events primed retrieval of high-level (coarse-grained) events, but not vice versa. In another experiment, participants received sentences that were not in the story but could be inferred from it. They were more likely to report that these sentences, which were never presented, had occurred in the story if they corresponded to high-level events than if they corresponded to low-level events. This suggests that the high-level information had been inferred based on the underlying hierarchical representation. Similar effects have been obtained with filmed narratives, including direct comparisons between narratives presented as texts and as films (Lichtenstein & Brewer, 1980; Brewer & Dupree, 1983). In sum, research on event schemata and scripts supports the view that event segmentation in narrative is correlated with inferences about the goals of actors and with statistical dependencies between events.

Several current theories of narrative comprehension claim that understanding a story involves the construction of a situation model. Unlike event schemata and scripts, which are representations of classes of events, situation models are representations of particular events that are described in a narrative. These theories build on Johnson-Laird's (1989) characterization of mental models as simulations of actual situations and on the reading comprehension models of Kintsch and colleagues (Kintsch, 1994), who introduced the term "situation model" into the literature. One theory of comprehension that incorporates a situation model representation, called the event indexing model, is particularly helpful for thinking about the relation between event perception and narrative understanding (Zwaan & Radvansky, 1998; Zwaan, Magliano, & Graesser, 1995; Zwaan, 1999). The event indexing model proposes that readers and listeners construct a model that represents the situation described in a text. To do so, they track indices representing five dimensions of events: the current *time,* the current location in *space,* the *objects* and *characters* currently present and relevant, the *causes* of events, and the *intentions* of protagonists. A change in any dimension requires that the reader update the current model, deactivating one representation and activating a new one or reactivating a previous representation. Considering the conceptual similarity between situation models and event models, it is a small step to propose that readers segment a narrative into events based on changes in these features. According to the theory proposed here, features such as time, space, objects, and characters affect segmentation in narrative via bottom-up processing. However, according to the theory, features such as causes and intentions depend also on inferences based on prior knowledge.

Until recently, there was little evidence for the proposal that changes in the dimensions of events as proposed by the event indexing model are perceived by readers as boundaries between events. Two sets of recent experiments provide strong support for this proposal. One series of studies (Speer, Zacks, & Reynolds, under review) took a correlational approach, employing narratives from pre-existing descriptions of a child's activities over the course of a day (Barker & Wright, 1966). Each story was divided into clauses (defined as a verb plus its argument structure). Raters coded each clause for changes in the five dimensions of situation models described previously. Readers then segmented the activity in the stories. Readers tended to

identify event boundaries at those clauses in which one of the situation model dimensions was changing. The probability of identifying an event boundary increased parametrically with the number of dimensions that changed. These findings support the model's claim that event segmentation depends on change. In line with the model's claim that prediction becomes more difficult at event boundaries, clauses following changes in a dimension were rated as being less predictable than other clauses. A second series of studies used narratives in which shifts in time were manipulated using temporal references (Speer & Zacks, 2005). An example is provided in Figure 4. Participants were asked to segment the activity in these stories using the unitization procedure described previously. Sentences containing a temporal reference were likely to be identified as the onsets of new events and this was especially true when the temporal reference indicated a long interval (an hour as opposed to a moment).

According to the theory, the perception of an event boundary leads to a cascade of processing that resets the event models' contents. Thus, information presented prior to an event boundary should be difficult to retrieve following that boundary. In the Speer and Zacks (2005) study, this hypothesis was tested by measuring reading time and recognition memory (see Figure 5). Reading times for sentences containing a shift in narrative time were longer than nearly identical control sentences, consistent with the hypothesis that increased processing occurred. References to information presented just prior to a temporal reference were read more slowly when the temporal reference indicated a time shift than when no time shift occurred. This provides indirect support for the hypothesis that memory is diminished following an event boundary. Direct support for this hypothesis came from an experiment in which participants' recognition memory was tested immediately after reading a temporal reference sentence. Information presented just before the temporal reference was recalled less well when the temporal reference indicated a time shift than when no time shift occurred (See also Zwaan, 1996, Rinck & Bower, 2000, and Bower & Rinck, 2001, and see Zwaan & Radvansky, 1998 for a review of related effects). Together, these results support the theory's claim that event segmentation is a form of cognitive control, modulating cognitive processing and memory maintenance.

In sum, research on narrative comprehension provides support for four implications of the theory. First, the fact that reading or listening to stories leads to robust event segmentation supports the implication that event segmentation is multi-sensory – it does not depend on specific visual or auditory features present in live activity. Second, event segmentation depends on changes, particularly on changes in dimensions identified by the event indexing model (Zwaan, 1999). Third, event segmentation depends on prior knowledge, particularly knowledge about goals and statistical dependencies. Finally, event segmentation is a mechanism of cognitive control, regulating the contents of memory.

Neural Correlates of Event Segmentation

Research examining the neural mechanisms by which events are perceived and conceived is still in its infancy. However, studies of patients with disorders of event understanding and action planning and the few available neuroimaging studies of event understanding provide surprisingly strong support for the theory proposed here. As will be seen in this section, neuroimaging data support the theory's claims that event segmentation is a spontaneous concomitant of ongoing perception, happens simultaneously on multiple timescales, and depends on change. Patient data support the theory's claims that event segmentation incorporates information from multiple senses, depends on prior knowledge, and is a mechanism of cognitive control. Patient studies and neuroimaging data also provide critical information about how the brain may implement the information-processing theory proposed here. In this section we propose a tentative mapping between the processing components described earlier (see "A Theory of Event Segmentation"), and review the relevant

neurophysiological data with this mapping in mind. In particular, we focus on the possible neural substrates of the unique components of the theory, event models and event schemata.

Sensory Inputs and Perceptual Processing

Figure 6 reproduces Figure 1, annotated with the brain regions that are hypothesized to correspond to each of the computational units. Recall that the core of the segmentation mechanism in EST is the perceptual processing stream illustrated in the left of the figure. Early sensory representations are transformed in a perceptual processing stream leading to representations that predict the state of the world a short time hence. The early-stage representations correspond to the outputs of primary sensory areas, including primary auditory cortex (A1), primary visual cortex, (V1), and primary somatosensory cortex (S1). The nature of these representations is relatively well understood, particularly for V1. Representations in these areas are topographically organized maps of the distal environment, representing its spatial structure as well as modality-specific features (Kolb & Whishaw, 2003). In the case of vision, the mid-stage representations are also fairly well understood. After V1, visual processing is segregated into a dorsal and ventral stream, though there is some communication between the two (Felleman & Van Essen, 1991;Ungerleider & Mishkin, 1978). Regions within these streams represent specialized aspects of the visual world (and also integrate some information from other modalities via recurrent projections). Regions in inferotemporal cortex (IT) in the ventral stream respond selectively to properties of objects, including properties that are invariant over changes in orientation, lighting, etc. (Tanaka, 1996). Regions in the dorsal stream corresponding to areas MT and MST in the monkey, dubbed "the MT complex" (MT +), selectively represent features of object and observer movement (Tootell, et al., 1995). Adjacent areas in the posterior superior temporal sulcus (pSTS) are selectively activated by nonrigid biological motion (Grossman, et al., 2000;Beauchamp, Lee, Haxby, & Martin, 2003). These perceptual processing streams are oriented in time such that they not only represent the current state of the world, but represent predictions about what is likely to happen a short time later.

Two lines of research directly address the neurophysiology of perceptual processing during ongoing activity. In one study, participants passively viewed 30 min of a Hollywood movie while whole-brain activity was recorded with functional MRI (fMRI) (Hasson, Nir, Levy, Fuhrmann, & Malach, 2004). The fMRI data were warped into a common atlas space and analyzed to identify regions in which the brain responses were similar across observers. The areas identified included a region in the inferior temporal lobe (fusiform gyrus) that had previously been found to be sensitive to the presentation of faces; it responded when close-ups of faces were on the screen. Another temporal region (parahippocampal gyrus) that had previously been found to be sensitive to pictures of buildings and spaces responded selectively when establishing shots or other wide-angle views were presented. This paradigm established that these areas, which previously had been explored using highly controlled picture judgment tasks, responded selectively in the context of realistic dynamic movies (see also Bartels & Zeki, 2004).

Studies of neural processing during the passive viewing of event boundaries specifically address the theory's claim that event segmentation controls the deployment of processing resources over time. In one experiment (Zacks, et al., 2001), participants passively viewed movies of everyday events while brain activity was recorded with fMRI. Following passive viewing they watched the movies again, this time segmenting the activity by pressing a button to mark boundaries between events. A network of brain regions showed transient increases in activity at perceptual event boundaries during passive viewing (see Figure 7). In other words, the activity of regions in this network correlated with participants' later identification of segment boundaries. Evoked responses were greater for coarse-event boundaries than for fine-

event boundaries, consistent with the behavioral finding that participants encode activity hierarchically (Zacks, Tversky, & Iyer, 2001). Activated areas included right posterior frontal cortex (BA 6) and a collection of regions in extrastriate visual cortex, including temporal, occipital, and parietal areas (BAs 19, 31, 37, 39). In particular, the strongest activity was identified near the human MT complex (MT+), an area in posterior visual cortex that is selectively responsive to visual motion. A second study directly localized MT+ in individual participants who also completed the event viewing and segmentation tasks (Speer, Swallow, & Zacks, 2003). Evoked responses in the individually identified MT+ regions were large and statistically reliable. Another study (Zacks, Swallow, Vettel, & McAvoy, 2006) localized MT + in a set of observers who then went on to watch simple two-object animations of the sort described previously (see "Features That Correlate with Perceptual Event Boundaries"). MT + responded selectively when the objects were moving faster, and also selectively responded to event boundaries. These data indicate that extrastriate visual areas and right frontal cortex either perform computations that contribute to the detection of event boundaries or are up-regulated as part of boundary detection. One possibility is that the activity in extrastriate visual cortex reflects the transient opening of event models to sensory and perceptual information.

One recent study used a narrative reading paradigm to look at the relation between on-line processing of event features and processing of event boundaries (Speer, Reynolds, Swallow, & Zacks, under review). Participants read extended narratives that had been coded for changes in event dimensions (Speer, Zacks, & Reynolds, under review, see "Segmentation in Story Understanding," above) while brain activity was measured with fMRI. After scanning, participants re-read the narratives and segmented them into events. The results indicated that neural systems that are specialized for processing particular features of live action were transiently activated when reading about changes in those features. For example, reading that a character began interacting with a new object activated readers' somatomotor cortex, whereas reading that the primary character in a story moved from one place to another activated the medial temporal cortex, near regions activated during spatial changes in navigation in virtual reality (Shelton & Gabrieli, 2002; Burgess, Maguire, & O'Keefe, 2002). This indicates that changes in the narrated situation were being processed in real time. Further, the data suggested that the processing of these changes led to the perception of event boundaries: An overlapping network, including posterior parietal, and right anterior temporal and frontal cortex transiently increased in activity at event boundaries and these increases were fully mediated by activity associated with event changes. This pattern of results supports EST's claim that changes lead to transient increases in prediction error, which in turn lead to the detection of event boundaries and the updating of event models.

Perceptual Prediction and Error Detection

EST claims that perceptual predictions are constantly compared with actual sensory input, providing an evaluation of how well perception is functioning. As indicated in Figure 6, we hypothesize that predicted future inputs are represented in the anterior cingulate cortex (ACC), and that the error detection function is implemented by neuromodulatory nuclei in the midbrain – the substantia nigra, ventral tegmental area, and locus ceruleus. We begin by reviewing data that bear on the error detection mechanism and then work back to the perceptual prediction representations.

Several mechanisms have been identified that could compute prediction error in event-structure perception (Schultz & Dickinson, 2000). Increases in prediction error lead to a cascade of processing that has been characterized as an *orienting response* (Sokolov, Spinks, Naeaetaenen, & Lyytinen, 2002). According to the theory, the resetting of event models and the transient increase in sensitivity to sensory input are two components of that response. We hypothesize that this resetting is implemented by midbrain neuromodulatory systems. These

systems appear to broadcast error signals – including prediction errors – through widespread projections to the cortex. In particular, circuits based on dopamine and norepinephrine have received substantial attention. Dopamine neurons in the substantia nigra and ventral tegmental area are sensitive to differences between actual and predicted rewards (Schultz, 1998). Norepinephrine neurons in the locus ceruleus appear to track performance in attention-demanding tasks, and have been proposed to regulate the sensitivity of an organism to external stimuli (Usher, Cohen, Servan-Schreiber, Rajkowski, & Aston-Jones, 1999). The ACC is sensitive to both sorts of error and may play a role in adaptively modulating behavior in response to prediction error (Botvinick, Braver, Barch, Carter, & Cohen, 2001). More specifically, different subpopulations within the ACC respond when a monkey is learning a new sequential structure than when it has discovered the sequence and is using the learned information to guide performance (Procyk, Tanaka, & Joseph, 2000). ACC and nearby regions have been proposed to underlie learning of sequential structure in simple motor tasks and in cognitive domains (Koechlin, Danek, Burnod, & Grafman, 2002). One possibility is that the ACC computes the discrepancy between perceptual predictions and actual inputs (Cohen, Botvinick, & Carter, 2000) and that in when prediction error spikes this triggers the nuclei of the catecholamine neurotransmitter systems. These subcortical nuclei have diffuse projections throughout cortex (Schultz & Dickinson, 2000) and receive inputs from the ACC (Holroyd & Coles, 2002). The resetting of event models may be mediated by projections from the substantia nigra or locus ceruleus to the striatum, which modulates activity in frontostriatal circuits, or by specific reciprocal connections with lateral PFC (Picard & Strick, 1996). We believe that participants' identification of event boundaries corresponds to the resetting of event models when they perform event segmentation tasks. This account is consistent with a recent theory of locus ceruleus norepinephrine function, which proposes that norepinephrine serves as a general reset signal, allowing neural networks to move out of one stable state and settle into a new stable state based on the current network input (Bouret & Sara, 2005). We hypothesize that this reset is approximately hierarchically structured, such that resets of representations with longer timescales are generally a subset of resets of representations with shorter timescales. This could be implemented by evaluating prediction error at a range of timescales in the ACC, with subcomponents of the ACC projecting to distinct targets. Alternatively, ACC could signal only prediction errors on short timescales and computations within lateral PFC could compute which of these fine-grained breaks are also coarse-grained breaks.

There are a small number of neurophysiological data that bear on the relation between error detection and event segmentation. These include studies using EEG evoked response potential (ERP) methods in combination with sentence processing paradigms. In sentence processing, an ERP component called the N400 has been associated with the processing of local breakdowns in predictability (Kutas & Hillyard, 1980). One study of narrative comprehension contrasted sentence-level breakdowns in predictability with discourse-level breakdowns, by presenting participants with sentences that were locally coherent but whose meaning conflicted with the larger narrative (van Berkum, Hagoort, & Brown, 1999). These breakdowns in narrative predictability produced N400 responses that were very similar to those produced by sentence-level breakdowns. A follow-up study found similar results with auditory presentation rather than reading (van Berkum, Zwitserlood, Hagoort, & Brown, 2003). Breakdowns in predictability also have been studied with pictures and movies. In one study, participants viewed a series of grayscale images that told a brief story, with the final image being either congruent or incongruent with the preceding ones (West & Holcomb, 2002). Incongruous final images produced N400s and an earlier negative response at approximately 325 ms. In another study, participants viewed short movies in which an object appeared that was either expected or unexpected given the context of the movie (Sitnikova, Kuperberg, & Holcomb, 2003). Unexpected objects led to a frontal-medial N400 and a large late positive response at lateral electrodes. The N400 may have reflected a mismatch between the semantic features of the final object and its context, whereas the late positive component may have reflected a lack of fit of

the final object into the causal/logical structure of the activity (Sitnikova, Holcomb, & Kuperberg, in preparation). Together, these data suggest that drops in predictability in text, pictures, or movies, produce a reliable brain response that is maximal approximately 400 ms after presentation of the unexpected stimulus. A recent source-localization ERP study (Frishkoff, Tucker, Davey, & Scherg, 2004) localized the earliest correlates of semantic incongruity in a sentence-processing paradigm to the ACC, with activity then spreading to prefrontal sites (and, to a lesser degree, posterior sites). These results suggest that discrepancies between the current event model and incoming information can lead to error signals that originate in the ACC and then propagate widely throughout the brain.

Event Models and Event Schemata

Data from several paradigms suggest that both event models (representations of what is happening now) and event schemata (representations of semantic knowledge about events in general) may be implemented by anterior lateral PFC (BA 45/46). Many of the experimental tasks that have been used do not permit one to distinguish between schemata and models; therefore, we focus here on neuropsychological and neurophysiological data that establish the importance of PFC for tasks that require representations of events.

Before proceeding to neuropsychological data that *do* appear to constrain the neural substrate of event representations, we discuss two disorders in which the pattern of deficits *does not* appear to provide constraints with respect to which brain areas represent events. The first of these is dementia of the Alzheimer's type (DAT). On its face, DAT is a reasonable candidate for a disorder affecting event representations because patients with the disorder sometimes fail to remain oriented with regard to the time, location, and people present in a given situation. Two studies examined the ability of patients with DAT to generate scripts for a particular activity or to verify the order in which two actions typically occurred in an activity (Weingartner, Grafman, Boutelle, Kaye, & Martin, 1983; Grafman, et al., 1991). Such tasks should depend on the presence of intact event schemata and may involve the construction of event models. Both studies found evidence that patients with DAT had impaired script processing; however, this impairment was associated with other disorders of semantic knowledge and did not appear to be a specific impairment of event representations. A third study assessed the ability of patients with very mild DAT to segment and to remember everyday activities (Zacks, Speer, Vettel, & Jacoby, in press). Those with dementia were poorer at segmenting than neurologically healthy older adults and had poorer memory for the events. However, these deficits were again part of a general pattern of cognitive decline. Within both groups of participants, event segmentation was predictive of later memory even after controlling for overall level of cognitive function. This suggests that event understanding plays a unique role in memory; however, these data do not indicate that event understanding is selectively impaired in DAT.

Another patient population in which script processing has been investigated is autism. In one study (Trillingsgaard, 1999), high-functioning autistic children (with IQ in the normal range) and control participants matched for mental age were asked to describe what typically happens in a set of everyday activities. The controls almost always gave lists of events that met minimal criteria for conforming to scripts, but the autistic individuals did so only half of the time. However, the autistic children also were impaired on a test of second-order theory of mind (knowing what someone else knows about what you know), which does not, on its face, appear to depend specifically on event representations. Thus, DAT and possibly autism may produce disordered event representations, but this appears to be the result of general cognitive disturbances rather than damage to event representations in particular.

Although studies of DAT and autism do not appear to constrain the localization of event representations, studies of patients with focal damage to PFC do. PFC has been associated with

the processing of temporally structured information and the maintenance of information over long delays (for a review, see Fuster, 1997). The specific association of PFC with script processing has been proposed by Grafman and colleagues (e.g., Grafman, 1995). They argued that PFC is the storage site for event representations call managerial knowledge units (MKUs), a type of script. They pointed out that the cytoarchitecture of PFC is comparable to the rest of cortex, suggesting that it processes information in ways similar to other cortical areas. However, subdivisions of PFC are uniquely positioned to integrate multimodal information about the typical unfolding of everyday events. In support of this view, Grafman and colleagues have built on the previous clinical literature on action disorders, collecting new lesion and neuroimaging data to test their proposal directly.

One study compared patients with prefrontal lesions to patients with posterior lesions and non-brain-damaged controls (Sirigu, et al., 1995). When participants were asked to list the events that are typical of everyday activities, the prefrontal group was more likely than the other groups to end their lists early (before the activity was complete) or late (including extra actions outside the activity). When the participants were asked to place the events in a script into the correct temporal order, all but one of the prefrontal patients made errors, whereas the other groups performed without error. Other studies have found that patients with prefrontal lesions are less able than controls to identify violations of normal sequential structure in scripts and to identify events that are not part of a script (Sirigu, et al., 1996; Allain, Le Gall, Etcharry-Bouyx, Aubin, & Emile, 1999).

Converging with the data from patients with focal PFC lesions are data from patients with neurological diseases that include PFC pathology, which suggest that PFC dysfunction impairs access to either event models or event schemata. Patients with schizophrenia have relatively intact abilities to identify fine-grained event boundaries, but are selectively impaired at identifying the correct location of coarse-grained boundaries (Zalla, Verlut, Franck, Puzenat, & Sirigu, 2004). Patients with Parkinson's disease (PD), which is characterized by degeneration of frontostriatal dopamine circuits, also are impaired on tasks that require ordering events within a script and detecting events from outside the script (Zalla, et al., 1998). The authors of this study proposed that the PD patients' deficits did not reflect damage to the underlying representations in PFC, but to switching mechanisms implemented by frontostriatal circuits; a similar logic could be applied to the schizophrenia data, as frontal dopamine projections are compromised in schizophrenia (Knable & Weinberger, 1997).

Several neuroimaging studies of event understanding have used script judgment tasks. In one study, participants were asked to imagine the sequence of actions involved in dressing and preparing for an emotionally neutral event (dinner) or a sad event (their mother's funeral) while brain activity was recorded using positron emission tomography (PET, Partiot, Grafman, Sadato, Flitman, & Wild, 1996). Thinking about the sequence of activities in both the neutral and sad contexts led to increased activity in PFC (including the superior, medial, and middle frontal gyri) compared to a set of loose control tasks. An fMRI study of script processing used an order verification task in which participants read a list of words describing events and were asked to indicate whether the events were listed in the order in which they typically occurred (Crozier, et al., 1999). A tight control task was administered: Verifying whether a list of words formed a syntactically valid sentence. The script task led to increases in four areas not activated in the sentence task: the left and right middle frontal gyri in PFC [Brodmann's area (BA) 8], the left supplementary motor area, the posterior frontal cortex (BA 6), and the left angular gyrus in the parietal lobe (BA 39). Activation in these areas has subsequently been replicated in another fMRI study (Knutson, Wood, & Grafman, 2004). A recent PET study focused on the temporal grain of script processing (Ruby, Sirigu, & Decety, 2002). Participants judged whether three-event sequences (depicted as pictures or two-word descriptions) were shown in their typical order. The sequences either showed events at a long timescale (e.g., growing a

crop) or a short timescale (e.g., brushing one's teeth). Long-term order verification was associated with stronger activity bilaterally in the angular gyrus (BA 39), the precuneus, and the medial superior frontal gyrus. Short-term order verification was associated with stronger activity in the left middle frontal gyrus, inferior temporal gyrus, middle occipital gyrus (BA 19), and supramarginal gyrus (BA 40). Particularly intriguing was a pattern suggesting a topographic organization of short-term and long-term events in parietal cortex, with short-term event structure apparently represented more anteriorly. However, this design does not permit conclusions about which areas the order verification task activated overall. Together, these results suggest that PFC and perhaps the lateral parietal cortex are selectively activated by thinking about the sequential structure of events. These data also provide some indication that this activation is stronger in the left hemisphere; however, further research is needed to clarify the role and lateral organization of these areas.

PFC does not generally increase in activity at event boundaries during passive viewing (Zacks, et al., 2001; Zacks, Swallow, Vettel, & McAvoy, 2006; Speer, Reynolds, & Zacks, in press). However, increases in fMRI activity were observed in lateral and medial PFC during an event segmentation task and electroencephalography indicated that a negative wave of activity moved from prefrontal regions to posterior regions prior to the decision to identify a segment boundary (Hanson, Negishi, & Hanson, 2001). One possibility is that this PFC activity reflects the activation of new event models and/or event schemata at event boundaries. However, another possibility is that it reflects task-specific decision-making or response planning rather than the activation of event representations. EST does not make strong predictions about whether PFC should be transiently activated at event boundaries; the theory predicts that the content of event models will change at this point, but such changes may not produce global increases or decreases in overall neural activity.

In addition to data from event comprehension, studies of patients with difficulty sequencing actions support the hypothesis that PFC is specialized to maintain event representations. Schwartz and colleagues have characterized *action disorganization syndrome* as a selective impairment of the ability to sequence goal-directed actions in the face of an intact ability to perform individual actions and an intact understanding of the objects and movements involved (Schwartz, Reed, Montgomery, Palmer, & Mayer, 1991; Schwartz, et al., 1995; Schwartz, 1995; Schwartz, Segal, Veramonti, Ferraro, & Buxbaum, 2002). For example, when one patient with a bilateral lesion to frontal cortex tried to brush his teeth, he showed disturbances of sequential ordering that included a dramatic tendency to perseverate, repeating individual actions such as rinsing the brush under the faucet or repeating whole sequences of actions (Schwartz, Reed, Montgomery, Palmer, & Mayer, 1991). Direct evidence that similar representations support the sequence of events in comprehension and in action comes from a study of frontal lesion patients with and without action disorganization syndrome (Humphreys & Forde, 1998). In this study, patients with frontal lesions who had difficulty performing sequentially structured activities also performed poorly at judgments about event structure including the script-listing task described previously. However, it is important to note that the question of whether action disorganization is the result of specific damage to event representations in PFC remains a matter of debate (Schwartz, et al., 1998).

In sum, data from neuropsychological and neurophysiological studies support the view that representations of events are subserved by the lateral PFC. Our characterization of PFC representations has some similarity to that of Miller & Cohen (2001); however, as noted previously we hypothesize that PFC represents events rather than goals and the means to achieve them. Our characterization of event representations in PFC also draws heavily on the arguments of Grafman and colleagues (Grafman, Partiot, & Hollnagel, 1995; Grafman, 1995; Wood & Grafman, 2003) as indicated previously. The present theory augments those theoretical proposals in two ways: first, it proposes a mechanism by which event models are

reset, and second, it proposes that working memory representations can be dissociated from long-term memory representations.

Toward Refining the Theory

The theory in its current form has some clear limitations. First, the mechanism by which event models are reset is an important aspect of the account, but it needs more development. More formal modeling, particularly aimed at exploring the relationship between the reset mechanism proposed here and that employed in ART models (Grossberg, 1999; Carpenter & Grossberg, 2003), would be very productive.

A second limitation is that the theory in its current form is a purely passive perceiver. In our view, perception and prediction are tightly interleaved with motor simulation. We hypothesize that the event schemata described in EST are used not just to guide perception of new activities based on past experience, but also to plan actions. We recognize that perceivers are usually also actors and so their event models include information about their own goals and allow them to anticipate the consequences of their actions as well as the actions of others. The general proposal that common representations support perception and action has broad empirical support (for reviews, see Gallese, 2001; Hommel, Muesseler, Aschersleben, & Prinz, 2001; Prinz, 1997). Moreover, there are hints that the specific representations posed by event models may play important roles in guiding action planning (see also Zacks & Tversky, 2001). In infant and child development, parsing activity into goal-based events may be critical for learning by imitation (Baldwin & Baird, 1999; Meltzoff, 1995). In the procedural learning studies described previously (see "Consequences of Event Segmentation for Long-Term Memory"), higher-quality segmentation was associated with better learning of an action sequence (Hard, Lozano, & Tversky, in press; Lozano, Hard, & Tversky, in press; Zacks & Tversky, 2003). Finally, in the neural network modeling we have reviewed (see "A Theory of Event Segmentation"), the gating of memory representations based on prediction error has been found to be a powerful technique for action control. Other models of everyday action sequencing have stable event representations that are similar in spirit to those of EST, though they do not share its commitment to a gating mechanism (Botvinick & Plaut, 2002; Botvinick & Plaut, 2004; Cooper & Shallice, 2000). We believe an important goal for future research should be to extend EST to account for (a) how event models explicitly represent one's own goals and (b) how perceptual prediction can include predicting the consequences of one's actions.

The theory makes one prediction that does not square well with the available data: If sudden increases in prediction error detected in the ACC lead to the perception of an event boundary, one would expect that increased ACC activity would be observed at event boundaries during passive viewing. This has not been the case (Speer, Reynolds, & Zacks, in press; Speer, Swallow, & Zacks, 2003; Zacks, et al., 2001). However, as was the case for prefrontal cortex, ACC activity has been reported during the active detection of event boundaries (Hanson, Negishi, & Hanson, 2001). It is possible that the failure to detect ACC activity at event boundaries during passive viewing reflects a lack of power. The passive viewing studies, to date, have used relatively predictable and stereotyped events; thus, these signals may have been especially weak. Alternatively, it may be that the error calculation is performed in the ACC, but the fMRI signal is detected at the downstream afferents from the ACC. However, this is at odds with findings that the fMRI signal in the ACC is modulated by conflict in other domains (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Further research is needed to verify whether the ACC responds selectively at event boundaries; if it does not, research should focus on evaluating other candidate mechanisms for the error detection mechanism.

Limitations notwithstanding, the theory in its current form provides a heuristic framework to guide future research. First, the theory predicts that the perception of an event boundary is

associated with an orienting response. This could be tested with measures of eye behaviors (eye movements, blinks, and pupil diameter), postural responses, and peripheral physiological responses (galvanic skin response, heart rate).

Second, EST predicts that the functional connectivity between sensory processing pathways and right PFC increases at event boundaries. This could be tested using electrophysiology or functional MRI by measuring whether the states of these regions become transiently more correlated at event boundaries.

Third, EST claims that event models and event schemata are dissociable. If so, are they localized to different parts of the brain? Current data on this question are mixed. One proposal comes from the text processing literature. In a recent review, Gernsbacher and Kaschak (2003) argued that the processes by which isolated elements are connected to build a coherent representation of the situation described by a text depend particularly on brain areas in the right hemisphere. Right PFC has been associated with the active maintenance of multimodal information over time (Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000; Gruber & von Cramon, 2003). Conversely, there is evidence that the left PFC is specialized for the storage of semantic knowledge (Cabeza & Nyberg, 2000). Together, such results suggest that event models may be right lateralized and event schemata may be left lateralized within PFC. However, the data regarding the specialization of the right hemisphere for establishing narrative coherence are mixed (Robertson, et al., 2000; St. George, Kutas, Martinez, & Sereno, 1999; Ferstl & von Cramon, 2002; Ferstl & von Cramon, 2001; Maguire, Frith, & Morris, 1999; Speer, 2005), and there are no direct comparisons of working memory and semantic memory representations of events in the neuroimaging or neuropsychological literatures. An alternative possibility is that event models and event schemata are both implemented by bilateral PFC, but by different regions within PFC or different populations of neurons within a region.

A fourth line of research suggested by the theory concerns the acquisition of event schemata. Schema acquisition is a difficult and under-studied problem, but the current framework offers some points of traction. In particular, it may be valuable to adapt procedures from the sequence learning literature for studying the acquisition of schemata for everyday events, because EST claims that prefrontal event schemata are the same representations that are involved in some kinds of sequential learning. EST also makes specific neuroanatomic predictions (based primarily on proposals by Hazeltine & Ivry, 2003, mentioned previously). For example, the cerebellum has been implicated particularly strongly in the early stages of motor learning when temporal coordination is effortful. Thus, activity in the cerebellum should be substantial when one is exposed to new sequentially structured events and should diminish with learning. Also, the supplementary motor area and lateral premotor cortex, both posterior to PFC in the frontal lobe, have been associated with implicit learning of sequential relations in motor sequences. One possibility is that these regions implement part of the sequential learning mechanisms described previously, and therefore should be active during the learning of new sequential relations in everyday events, even if no overt motor output is required.

Conclusion

Events are natural kinds, just like objects: Everyday life consists of picnics and meetings just as it consists of chairs and birds. The perception of event structure has only recently emerged as an independent scientific problem. It draws on a broad body of research in linguistics, psychology, and neuroscience as well as a small but growing literature on the perception of event structure *per se*. We are entering a stage in which these diverse findings can be integrated into theories that provide comprehensive accounts of event perception and point the way for future research.

Acknowledgments

References

Abbott V, Black JH, Smith EE. The representation of scripts in memory. Journal of Memory and Language 1985;24:179–99.

Abelson RP. Psychological status of the script concept. American Psychologist 1981;36:715–29.

Allain P, Le Gall D, Etcharry-Bouyx F, Aubin G, Emile J. Mental representation of knowledge following frontal-lobe lesion: dissociations on tasks using scripts. Journal of Clinical & Experimental Neuropsychology 1999;21(5):643–65. [PubMed: 10572284]

Avrahami J, Kareev Y. The emergence of events. Cognition 1994;53(3):239–61. [PubMed: 7842635]

Baddeley A. The episodic buffer: a new component of working memory? Trends in Cognitive Sciences 2000;4(11):417–23. [PubMed: 11058819]

Baird, JA.; Baldwin, DA. Making sense of human behavior: action parsing and intentional inference. In: Malle, BF.; Moses, LJ.; Baldwin, DA., editors. Intentions and intentionality: Foundations of social cognition. Cambridge, MA: MIT Press; 2001. p. 193-206.

Baldwin, DA.; Baird, JA. Action analysis: A gateway to intentional inference. In: Rochat, P., editor. Early social cognition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1999. p. 215-40.

Barker, RG.; Wright, HF. One boy's day; a specimen record of behavior. Hamden, CT: Archon Books; 1966.

Bartels A, Zeki S. Functional brain mapping during free viewing of natural scenes. Human Brain Mapping 2004;21:75–85. [PubMed: 14755595]

Bartlett, FC. Remembering: A study in experimental and social psychology. New York: The Macmillan Company; 1932.

Beauchamp MS, Lee KE, Haxby JV, Martin A. fMRI responses to video and point-light displays of moving humans and manipulable objects. Journal of Cognitive Neuroscience 2003;15:991–1001. [PubMed: 14614810]

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

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. Psychological Review 2001;108(3):624–52. [PubMed: 11488380]

Botvinick MM, Plaut DC. Representing task context: insights from a connectionist model of action. Psychological Research 2002;66:298–311. [PubMed: 12466927]

Botvinick MM, Plaut DC. Doing without schema hierarchies: A recurrent connectionist approach to routine sequential action and its pathologies. Psychological Review 2004;111(2):394–429.

Bouret S, Sara SJ. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. Trends in Neurosciences 2005;28(11):574–82. [PubMed: 16165227]

Bower GH, Black JB, Turner TJ. Scripts in memory for text. Cognitive Psychology 1979;11:177–220.

Bower GH, Rinck M. Selecting one among many referents in spatial situation models. Journal of Experimental Psychology: Learning, Memory, & Cognition 2001;27(1):81–98.

Braver, TS.; Cohen, JD. On the control of control: The role of dopamine in regulating prefrontal function and working memory. In: Monsell, S.; Driver, J., editors. Attention and Performance XVIII. Cambridge, MA: MIT Press; 2000. p. 713-38.

Brent MR. An efficient, probabilistically sound algorithm for segmentation and word discovery. Machine Learning Mach Learn 1999;34(1–3):71–105.

Brewer WF, Dupree DA. Use of plan schemata in the recall and recognition of goal-directed actions. Journal of Experimental Psychology: Learning, Memory, and Cognition 1983;9(1):117–29.

Burgess N, Maguire EA, O'Keefe J. The human hippocampus and spatial and episodic memory. Neuron 2002;35(4):625–41. [PubMed: 12194864]

Cabeza R, Nyberg L. Imaging cognition II: An empirical review of 275 PET and fMRI studies. Journal of Cognitive Neuroscience 2000:1–47. [PubMed: 10769304]

Carpenter, GA.; Grossberg, S. Adaptive resonance theory. In: Arbib, MA., editor. The handbook of brain theory and neural networks. 2. Cambridge, MA: MIT Press; 2003. p. 87-90.

Cohen JD, Botvinick M, Carter CS. Anterior cingulate and prefrontal cortex: who's in control? Nature Neuroscience 2000;3(5):421–23.

Cohen JD, Braver TS, O'Reilly RC. A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. Philosophical Transactions of the Royal Society, Series B 1996;351(1346):1515–27.

Cooper R, Shallice T. Contention scheduling and the control of routine activities. Cognitive Neuropsychology 2000;17(4):297–338.

Coull JT, Nobre AC. Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. Journal of Neuroscience 1998;18 (18):7426–35. [PubMed: 9736662]

Crozier S, Sirigu A, Lehericy S, Moortele P-Fvd, Pillon B, Grafman J, et al. Distinct prefrontal activations in processing sequence at the sentence and script level: An fMRI study. Neuropsychologia 1999;37:1469–76. [PubMed: 10617267]

Elman JL. Finding structure in time. Cognitive Science 1990;14(2):179–211.

Ericsson KA, Kintsch W. Long-term working memory. Psychological Review 1995;102(2):211–45. [PubMed: 7740089]

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex 1991;1(1):1–47. [PubMed: 1822724]

Ferstl EC, von Cramon DY. The role of coherence and cohesion in text comprehension: an event-related fMRI study. Cognitive Brain Research 2001;11(3):325–40. [PubMed: 11339984]

Ferstl EC, von Cramon DY. What does the frontomedian cortex contribute to language processing: coherence or theory of mind? Neuroimage 2002;17(3):1599–612. [PubMed: 12414298]

Frank MJ, Loughry B, O'Reilly RC. Interactions between frontal cortex and basal ganglia in working memory: a computational model. Cognitive, Affective and Behavioral Neuroscience 2001;1(2):137–60.

Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 2004;306(5703):1940–43. [PubMed: 15528409]

Frishkoff GA, Tucker DM, Davey C, Scherg M. Frontal and posterior sources of event-related potentials in semantic comprehension. Cognitive Brain Research 2004;20(3):329–54. [PubMed: 15268912]

Fuster JM. The prefrontal cortex and its relation to behavior. Progress in Brain Research 1991;87:201–11. [PubMed: 1907745]

Fuster, JM. The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe. 3. Philadelphia: Lippincott-Raven; 1997.

Gallese V. The 'shared manifold' hypothesis. Journal of Consciousness Studies 2001;8(5–7):33–50.

Gernsbacher MA, Kaschak MP. Neuroimaging studies of language production and comprehension. Annual Review of Psychology 2003;54:91–114.

Glenberg AM, Kaschak MP. Grounding language in action. Psychonomic Bulletin & Review 2002;9(3):558–65. [PubMed: 12412897]

Grafman J. Similarities and distinctions among current models of prefrontal cortical functions Structure and Functions of the Human Prefrontal Cortex. Annals of the New York Academy of Sciences 1995;769:337–68. [PubMed: 8595036]

Grafman J, Partiot A, Hollnagel C. Fables of the Prefrontal Cortex. Behavioral and Brain Sciences 1995;18(2):349–58.

Grafman J, Thompson K, Weingartner H, Martinez R, Lawlor BA, Sunderland T. Script generation as an indicator of knowledge representation in patients with Alzheimer's disease. Brain and Language 1991;40(3):344–58. [PubMed: 2054591]

Grossberg S. The link between brain learning, attention, and consciousness. Consciousness and Cognition 1999;8:1–44. [PubMed: 10072692]

Grossman ED, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, et al. Brain areas involved in perception of biological motion. Journal of Cognitive Neuroscience 2000;12(5):711–20. [PubMed: 11054914]

Gruber O, von Cramon DY. The functional neuroanatomy of human working memory revisited - Evidence from 3-T fMRl studies using classical domain-specific interference tasks. NeuroImage 2003;19(3):797–809. [PubMed: 12880808]

Hanson C, Hanson SJ. Development of schemata during event parsing: Neisser's perceptual cycle as a recurrent connectionist network. Journal of Cognitive Neuroscience 1996;8(2):119–34.

Hanson C, Hirst W. On the representation of events: A study of orientation, recall, and recognition. Journal of Experimental Psychology: General 1989;118(2):136–47. [PubMed: 2525593]

Hanson C, Hirst W. Recognizing differences in recognition tasks: A reply to Lassiter and Slaw. Journal of Experimental Psychology: General 1991;120(2):211–12.

Hanson, SJ.; Negishi, M.; Hanson, C. Connectionist neuroimaging. In: Wermter, S.; Austin, J.; Willshaw, D., editors. Emergent neural computational architectures based on neuroscience: Towards neuroscience-inspired computing. Berlin: Springer; 2001. p. 560-77.

Hard BM, Lozano SC, Tversky B. Hierarchical encoding of behavior: Translating perception into action. Journal of Experimental Psychology: General. in press.

Hard BM, Tversky B, Lang D. Segmenting abstract events: Building event schemas. Memory & Cognition. in press.

Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R. Intersubject synchronization of cortical activity during natural vision. Science 2004;303(5664):1634–40. [PubMed: 15016991]

Hazeltine, E.; Ivry, RB. Neural structures that support implicit sequence learning. In: Jimenez, L., editor. Attention and implicit learning. Philadelphia: John Benjamins; 2003. p. 73-107.

Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychological Review 2002;109(4):679–709. [PubMed: 12374324]

Hommel B, Muesseler J, Aschersleben G, Prinz W. The Theory of Event Coding (TEC): A framework for perception and action planning. Behavioral & Brain Sciences 2001;24(5):849–937. [PubMed: 12239891]

Humphreys GW, Forde EME. Disordered action schema and action disorganisation syndrome. Cognitive Neuropsychology 1998;15(6–8):771–811.

Johnson-Laird, PN. Mental Models. In: Posner, MI., editor. Foundations of cognitive science. Cambridge, MA: MIT Press; 1989. p. 469-500.

Jordan MI, Rumelhart DE. Forward models: Supervised learning with a distal teacher. Cognitive Science 1992;16(3):307–54.

Kahneman D, Treisman A, Gibbs BJ. The reviewing of object files: object-specific integration of information. Cognitive Psychology 1992;24(2):175–219. [PubMed: 1582172]

Kintsch W. Text comprehension, memory, and learning. American Psychologist 1994;49(4):294–303. [PubMed: 8203801]

Knable MB, Weinberger DR. Dopamine, the prefrontal cortex and schizophrenia. Journal of Psychopharmacology 1997;11(2):123–31. [PubMed: 9208376]

Knutson KM, Wood JN, Grafman J. Brain activation in processing temporal sequence: an fMRI study. NeuroImage 2004;23(4):1299. [PubMed: 15589094]

Koechlin E, Danek A, Burnod Y, Grafman J. Medial prefrontal and subcortical mechanisms underlying the acquisition of motor and cognitive action sequences in humans. Neuron 2002;35(2):371–81. [PubMed: 12160754]

Kolb, B.; Whishaw, IQ. Fundamentals of human neuropsychology. 5. New York: Worth Publishers; 2003.

Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 1980;207(4427):203–5. [PubMed: 7350657]

Lassiter GD. Behavior perception, affect, and memory. Social Cognition 1988;6(2):150–76.

Lassiter GD, Slaw RD. The unitization and memory of events. Journal of Experimental Psychology: General 1991;120(1):80–82.

Lassiter GD, Stone JI, Rogers SL. Memorial consequences of variation in behavior perception. Journal of Experimental Social Psychology 1988;24(3):222–39.

Lichtenstein ED, Brewer WF. Memory for goal-directed events. Cognitive Psychology 1980;12:412–45.

Lozano SC, Hard BM, Tversky B. Perspective-taking promotes action understanding and learning. Journal of Experimental Psychology: Human Perception & Performance. in press.

Maguire EA, Frith CD, Morris RG. The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. Brain 1999;122(Pt 10):1839–50. [PubMed: 10506087]

Meltzoff AN. Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. Developmental Psychology 1995;31(5):838–50.

Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Annual Review of Neuroscience 2001;24:167–202.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts; 1967.

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

Newtson, D. Foundations of attribution: the perception of ongoing behavior. In: Harvey, JH.; Ickes, WJ.; Kidd, RF., editors. New directions in attribution research. Hillsdale, NJ: Lawrence Erlbaum Associates; 1976. p. 223-48.

Newtson D, Engquist G. The perceptual organization of ongoing behavior. Journal of Experimental Social Psychology 1976;12:436–50.

Newtson D, Engquist G, Bois J. The objective basis of behavior units. Journal of Personality and Social Psychology 1977;35(12):847–62.

Nobre AC. Orienting attention to instants in time. Neuropsychologia 2001;39(12):1317–28. [PubMed: 11566314]

O'Reilly RC, Noelle DC, Braver TS, Cohen JD. Prefrontal cortex and dynamic categorization tasks: Representational organization and neuromodulatory control. Cerebral Cortex 2002;12:246–57. [PubMed: 11839599]

Partiot A, Grafman J, Sadato N, Flitman S, Wild K. Brain activation during script event processing. Neuroreport 1996;7(3):761–66. [PubMed: 8733740]

Picard N, Strick PL. Motor areas of the medial wall: a review of their location and functional activation. Cerebral Cortex 1996;6(3):342–53. [PubMed: 8670662]

Posner, MI.; Snyder, CRR. Attention and cognitive control. In: Bower, GH., editor. Information Processing and Cognition: The Loyola Symposium. Hillsdale, NJ: Lawrence Erlbaum Associates; 1975. p. 55-85.

Prabhakaran V, Narayanan K, Zhao Z, Gabrieli JDE. Integration of diverse information in working memory within the frontal lobe. Nature Neuroscience 2000;3(1):85–90.

Prinz W. Perception and action planning. European Journal of Cognitive Psychology 1997;9(2):129–54.

Procyk E, Tanaka YL, Joseph JP. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. Nature Neuroscience 2000;3(5):502–08.

Reynolds JR, Zacks JM, Braver TS. A computational model of event segmentation from perceptual prediction. Cognitive Science. in press.

Rinck M, Bower G. Temporal and spatial distance in situation models. Memory and Cognition 2000;28(8):1310–20.

Robertson DA, Gernsbacher MA, Guidotti SJ, Robertson RR, Irwin W, Mock BJ, et al. Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. Psychological Science 2000;11(3):255–60. [PubMed: 11273413]

Rougier NP, O'Reilly RC. Learning representations in a gated prefrontal cortex model of dynamic task switching. Cognitive Science 2002;26(4):503–20.

Rougier NP, Noelle DC, Braver TS, Cohen JD, O'Reilly RC. Prefrontal cortex and flexible cognitive control: rules without symbols. Proceedings of the National Academy of Sciences 2005;102(20):7338–43.

Ruby P, Sirigu A, Decety J. Distinct areas in parietal cortex involved in long-term and short-term action planning: a PET investigation. Cortex 2002;38(3):321–39. [PubMed: 12146659]

Rumelhart, DE. Notes on a schema for stories Language, thought, and culture. In: Bobrow, DG.; Collins, A., editors. Representation and understanding; studies in cognitive science. New York: Academic Press; 1975. p. 211-36.

Rumelhart, DE. Understanding and summarizing brief stories. In: Laberge, D.; Samuels, SJ., editors. Basic processes in reading: Perception and comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates; 1977. p. 265-303.

Rumelhart, DE. Schemata: The building blocks of cognition The Psychology of reading. In: Spiro, RJ.; Bruce, BC.; Brewer, WF., editors. Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education. Hillsdale, NJ: L. Erlbaum Associates; 1980. p. 33-58.

Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. Science 1996;274(5294): 1926–28. [PubMed: 8943209]

Schank, RC.; Abelson, RP. Scripts, plans, goals, and understanding: an inquiry into human knowledge structures. Hillsdale, NJ: L. Erlbaum Associates; 1977.

Schultz W. Predictive reward signal of dopamine neurons. Journal of Neurophysiology 1998;80(1):1–27. [PubMed: 9658025]

Schultz W, Dickinson A. Neuronal coding of prediction errors. Annual Review of Neuroscience 2000;23:473–500.

Schwan S, Garsoffky B. The cognitive representation of filmic event summaries. Applied Cognitive Psychology 2004;18(1):37–55.

Schwan S, Garsoffky B, Hesse FW. Do film cuts facilitate the perceptual and cognitive organization of activity sequences? Memory & Cognition 2000;28(2):214–23.

Schwartz MF. Re-examining the role of executive functions in routine action production. Annals of the New York Academy of Sciences 1995;769:321–5. [PubMed: 8595035]

Schwartz MF, Montgomery MW, Buxbaum LJ, Lee SS, Carew TG, Coslett HB, et al. Naturalistic action impairment in closed head injury. Neuropsychology 1998;12(1):13–28. [PubMed: 9460731]

Schwartz MF, Montgomery MW, Fitzpatrick-DeSalme EJ, Ochipa C, Coslett HB, Mayer NH. Analysis of a disorder of everyday action. Cognitive Neuropsychology 1995;12(8):863–92.

Schwartz MF, Reed ES, Montgomery M, Palmer C, Mayer NH. The quantitative description of action disorganisation after brain damage: A case study. Cognitive Neuropsychology 1991;8(5):381–414.

Schwartz MF, Segal M, Veramonti T, Ferraro M, Buxbaum LJ. The Naturalistic Action Test: A standardised assessment for everyday action impairment. Neuropsychological Rehabilitation 2002;12(4):311–39.

Seger CA. Implicit learning. Psychological Bulletin 1994;115(2):163–96. [PubMed: 8165269]

Shelton AL, Gabrieli JDE. Neural correlates of encoding space from route and survey perspectives. Journal of Neuroscience 2002;22(7):2711–17. [PubMed: 11923436]

Shipley, TF. An invitation to an event. In: Shipley, TF.; Zacks, JM., editors. Understanding events: How humans see, represent, and act on events. New York: Oxford University Press; in preparation

Sirigu A, Zalla T, Pillon B, Grafman J, Agid Y, Dubois B. Encoding of sequence and boundaries of scripts following prefrontal lesions. Cortex 1996;32(2):297–310. [PubMed: 8800616]

Sirigu A, Zalla T, Pillon B, Grafman J, Agid Y, Dubois B. Selective impairments in managerial knowledge following pre-frontal cortex damage. Cortex 1995;31(2):301–16. [PubMed: 7555008]

Sitnikova, T.; Holcomb, P.; Kuperberg, G. Neurocognitive mechanisms of human comprehension. In: Shipley, TF.; Zacks, JM., editors. Understanding events: How humans see, represent, and act on events. New York: Oxford University Press; in preparation

Sitnikova T, Kuperberg G, Holcomb PJ. Semantic integration in videos of real-world events: An electrophysiological investigation. Psychophysiology 2003;40(1):160–64. [PubMed: 12751813]

Sokolov, EN.; Spinks, JA.; Naeaetaenen, R.; Lyytinen, H. The orienting response in information processing. Mahwah, NJ: Lawrence Erlbaum Associates; 2002.

Speer, NK. Unpublished doctoral dissertation. Washington University in Saint Louis; Saint Louis, MO: 2005. Text comprehension processes in the brain.

Speer, NK.; Reynolds, JR.; Swallow, KM.; Zacks, JM. Reading stories activates neural representations of perceptual and motor experiences. Manuscript submitted for publication

Speer NK, Reynolds JR, Zacks JM. Human brain activity time-locked to narrative event boundaries. Psychological Science. in press.

Speer, NK.; Zacks, JM.; Reynolds, JR. Situation changes predict the perception of event boundaries, reading time, and perceived predictability in narrative comprehension. Manuscript submitted for publication

Speer NK, Swallow KM, Zacks JM. Activation of human motion processing areas during event perception. Cognitive, Affective & Behavioral Neuroscience 2003;3(4):335–45.

Speer NK, Zacks JM. Temporal changes as event boundaries: processing and memory consequences of narrative time shifts. Journal of Memory and Language 2005;53:125–40.

St George M, Kutas M, Martinez A, Sereno MI. Semantic integration in reading: engagement of the right hemisphere during discourse processing. Brain 1999;122(Pt 7):1317–25. [PubMed: 10388797]

Stränger, J.; Hommel, B. The perception of action and movement. In: Prinz, W.; Bridgeman, B., editors. Handbook of Perception and Action: Volume 1. San Diego, CA: Academic Press; 1996. p. 397-451.

Tanaka K. Inferotemporal cortex and object vision. Annual Review of Neuroscience 1996;19:109–39.

Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, et al. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. Journal of Neuroscience 1995;15(4):3215–30. [PubMed: 7722658]

Trillingsgaard A. The script model in relation to autism. European Child & Adolescent Psychiatry 1999;8 (1):45–49. [PubMed: 10367740]

Ungerleider, LG.; Mishkin, M. Two cortical visual systems. In: Ingle, D.; Goodale, MA.; Mansfield, RJW., editors. Analysis of visual behavior. Cambridge, MA: MIT Press; 1978. p. 549-86.

Usher M, Cohen JD, Servan-Schreiber D, Rajkowski J, Aston-Jones G. The role of the locus coeruleus in the regulation of cognitive performance. Science 1999;283(5401):549–54. [PubMed: 9915705]

van Berkum JJA, Hagoort P, Brown CM. Semantic integration in sentences and discourse: evidence from the N400. Journal of Cognitive Neuroscience 1999;11(6):657–71. [PubMed: 10601747]

van Berkum JJA, Zwitserlood P, Hagoort P, Brown CM. When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. Cognitive Brain Research 2003;17(3):701–18. [PubMed: 14561457]

van Dijk, TA.; Kintsch, W. Strategies of discourse comprehension. New York: Academic Press; 1983.

Weingartner H, Grafman J, Boutelle W, Kaye W, Martin PR. Forms of Memory Failure. Science 1983;221 (4608):380–82. [PubMed: 6867715]

West WC, Holcomb PJ. Event-related potentials during discourse-level semantic integration of complex pictures. Cognitive Brain Research 2002;13(3):363–75. [PubMed: 11919001]

Wilder DA. Effect or predictability on units of perception and attribution. Personality and Social Psychology Bulletin 1978a;4(2):281–84.

Wilder DA. Predictability of behaviors, goals, and unit of perception. Personality and Social Psychology Bulletin 1978b;4(4):604–07.

Wood JN, Grafman J. Human prefrontal cortex: processing and representational perspectives. Nature Reviews Neuroscience 2003;4:139–47.

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

Zacks JM, Braver TS, Sheridan MA, Donaldson DI, Snyder AZ, Ollinger JM, et al. Human brain activity time-locked to perceptual event boundaries. Nature Neuroscience 2001;4(6):651–55.

Zacks, JM.; Speer, NK.; Vettel, JM.; Jacoby, LL. Event understanding and memory in healthy aging and dementia of the Alzheimer type. in press

Zacks JM, Swallow KM, Vettel JM, McAvoy MP. Visual movement and the neural correlates of event perception. Brain Research 2006;1076(1):150–62. [PubMed: 16473338]

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

Zacks JM, Tversky B. Structuring information interfaces for procedural learning. Journal of Experimental Psychology: Applied 2003;9(2):88–100. [PubMed: 12877269]

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

Zalla T, Sirigu A, Pillon B, Dubois B, Grafman J, Agid Y. Deficient in evaluating pre-determinated sequences of script events in patients with Parkinson's disease. Cortex 1998;34(4):621–28. [PubMed: 9800095]

Zalla T, Verlut I, Franck N, Puzenat D, Sirigu A. Perception of dynamic action in patients with schizophrenia. Psychiatry Research 2004;128(1):39. [PubMed: 15450913]

Zwaan RA. Processing narrative time shifts. Journal of Experimental Psychology: Learning, Memory, & Cognition 1996;22(5):1196–207.

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

Zwaan RA, Magliano JP, Graesser AC. Dimensions of situation model construction in narrative comprehension. Journal of Experimental Psychology: Learning, Memory, & Cognition 1995;21(2):386–97.

Zwaan RA, Radvansky GA. Situation models in language comprehension and memory. Psychological Bulletin 1998;123(2):162–85. [PubMed: 9522683]

Zwaan RA, Stanfield RA, Yaxley RH. Language comprehenders mentally represent the shape of objects. Psychological Science 2002;13(2):168–71. [PubMed: 11934002]
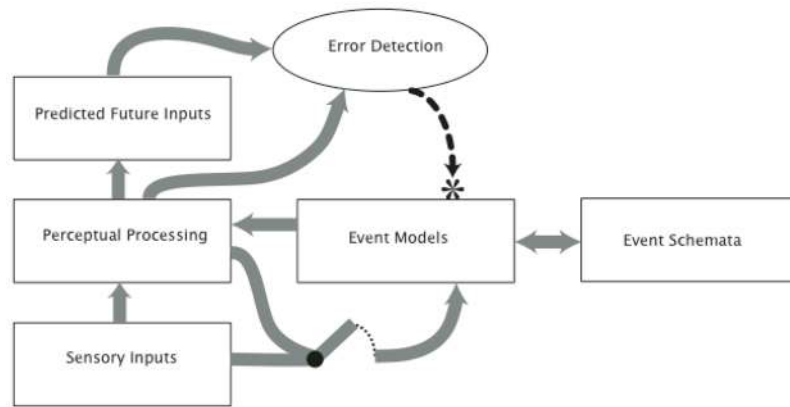
Figure 1.
Schematic depiction of the theory. Thin gray arrows indicate the flow of information between processing areas. Dashed lines indicate projections that lead to the resetting of event models. The connection from sensory inputs to event models is gated, such that event models receive sensory input only during the reset phase.
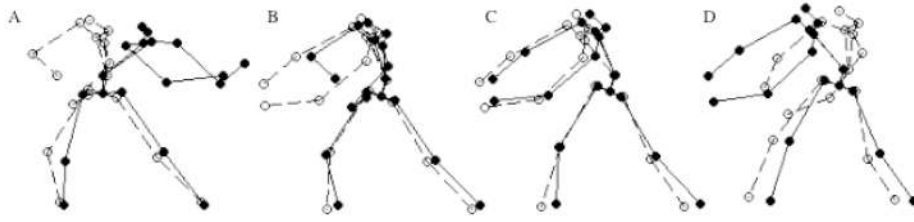
Figure 2.
Graphical representation of the inputs and outputs to the Reynolds et al. (in press) neural network model. A–D show four consecutive frames of an actor chopping down a tree. The target output (dashed lines) on each frame is the model input (solid lines) on the subsequent frame. Frames C and D have similar inputs but dissimilar target outputs, illustrating the need to represent necessitating the representation of long-term sequential dependencies to achieve accurate prediction. *Note*. From "A computational model of event perception from perceptual prediction," by J. R. Reynolds, J. M. Zacks, and T. S. Braver, in press, *Cognitive Science*. Copyright 2006 by the authors. Reprinted with permission.
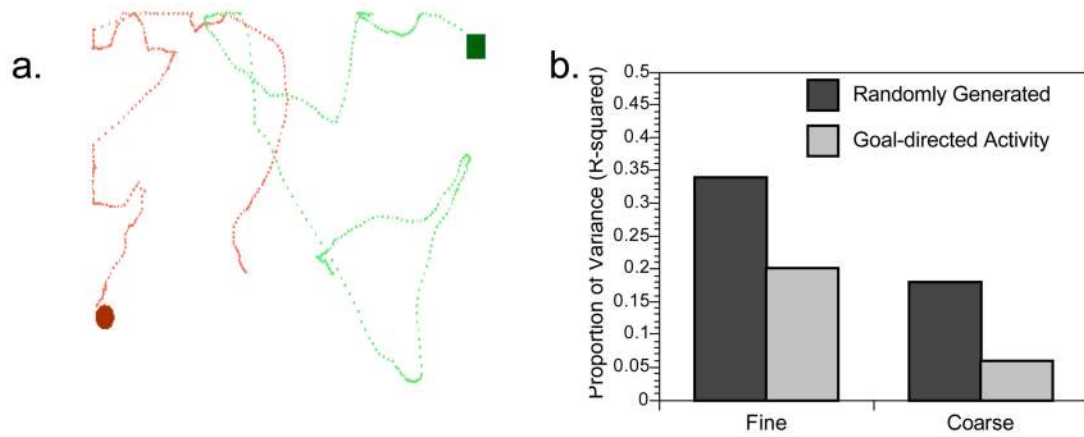
Figure 3.
(a) Trace of two randomly moving objects. From the time-varying x and y locations of the object, an exhaustive set of movement features was calculated, which included the objects' position, speed, and acceleration, the distance between the objects, and their relative speed and relative acceleration. (b) Movement features were quite predictive of where observers segmented the activity when the activity was randomly generated and was unitized at a fine grain. The dependence of unitization on movement features was reduced for goal-directed human activity, and for coarse-grained unitization. *Note*. Adapted from "Using movement to understand simple events," by J. M. Zacks, 2004, *Cognitive Science, 28*, p. 979–108. Copyright 2006 by the author. Adapted with permission.

...

Introduction	She had just bought a new camera, and she hoped the pictures would turn out well.
Object	She could hear water running, and figured there must be a **creek** nearby.
Time shift	*A moment later/An hour later*, she was collecting wood for a fire.
Anaphor	She heard a noise near the stream.

...

Figure 4.
Excerpt from a narrative describing a woman on a backpacking trip. Sequences such as this were embedded throughout the narratives. A critical object (in **bold**) is introduced, followed by a sentences that either contains a time shift (*an hour later*) or a control phrase (*a moment later*). Finally an anaphor (underlined) refers back to the critical object. (See Speer, et al., 2005.)
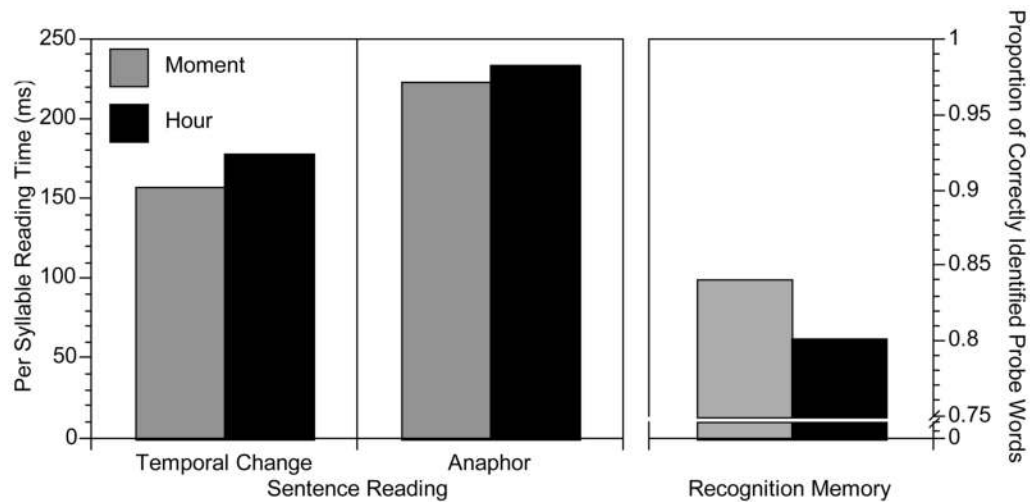
Figure 5.
The left panel shows that readers were slower to process sentences if they contained a time shift ("an hour later") than a control phrase (" a moment later"). As is seen in the middle panel, they were also slower to read a following sentence that contained a reference to a previously mentioned object. Finally, the right panel shows that after reading a time shift sentence, critical objects were correctly recognized less often. (See Speer, et al., 2005.)
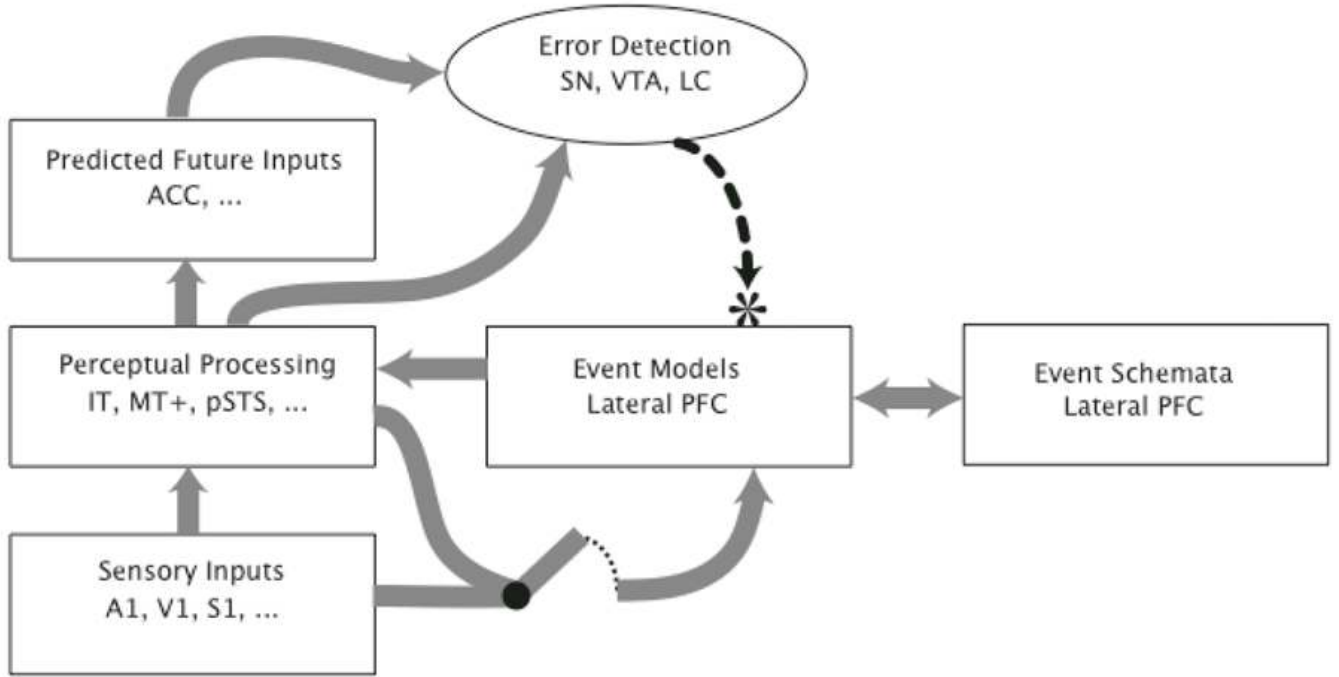
Figure 6.
Schematic depiction of the model, with hypotheses about the neurophysiological structures corresponding to the different components of the model. Thin gray arrows indicate the flow of information between processing areas, which are proposed to be due to long-range excitatory projections. Dashed lines indicate projections that lead to the resetting of event models. Abbreviations: PFC = prefrontal cortex; IT = inferotemporal cortex; MT+ = human MT complex; pSTS = posterior superior temporal sulcus, ACC = anterior cingulate cortex, SN = substantia nigra, VTA = ventral tegmental area, LC = locus ceruleus. Gray arrows indicate the flow of information between processing areas.
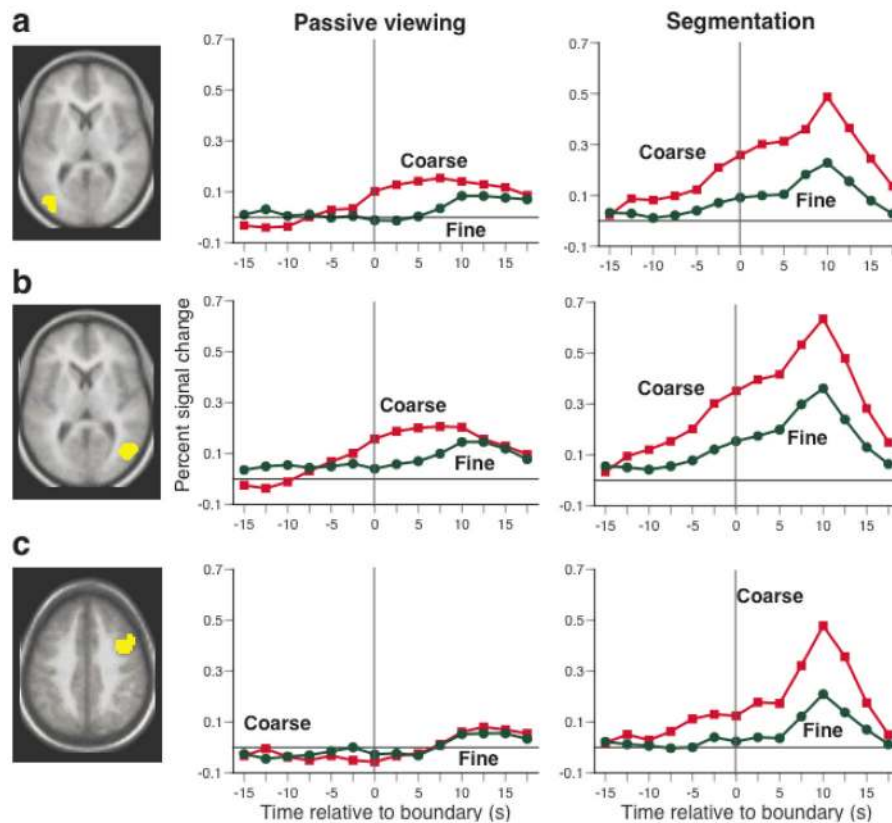
Figure 7.
Focal brain activity in three regions, showing transient changes at event segment boundaries (identified by vertical lines). Panels **a–c** depict the bilateral extrastriate and right frontal clusters of activated voxels. All regions showed reliable responses to event boundaries during passive viewing, and larger responses during active segmentation. The left image shows the extent of the cluster, superimposed on an averaged anatomical image for the 16 participants. The two graphs to the right show the evoked response during coarse and fine event units for passive viewing and active segmentation (see Zacks, et al., 2001).