Running head: EVENT STRUCTURE

Event Structure in Perception and Conception

Jeffrey M. Zacks[1,2] and Barbara Tversky[1]
[1]Stanford University, [2]Washington University

Please address communications to:

Jeffrey M. Zacks

Washington University

Psychology Department

St. Louis, MO 63130-4899

314-935-8454

jzacks@artsci.wustl.edu

IN PRESS, PSYCHOLOGICAL BULLETIN

Abstract

Events can be understood in terms of their temporal structure. Here, we draw on several bodies of research to construct an analysis of how people use event structure in perception, understanding, planning, and action. Philosophy provides a grounding for the basic units of events and actions. Perceptual psychology provides an analogy to object perception: Like objects, events belong to categories and, like objects, events have parts. These relationships generate two hierarchical organizations for events: taxonomies and partonomies. Event partonomies have been studied by looking at how people segment activity as it happens. Structured representations of events can relate partonomy to goal relationships and causal structure; such representations have been shown to drive narrative comprehension, memory and planning. Computational models provide insight into how mental representations might be organized and transformed. These different approaches to event structure converge on an explanation of how multiple sources of information interact in event perception and conception.

> "I claim not to have controlled events, but confess plainly that they have controlled me." —Abraham Lincoln

> "Temporal aspects of behavior are among the most compelling in experience and among the most easily measured of all of behavior's unnumbered characteristics. Despite the saliency of the time dimension however, little is known about the actual arrangement of behavior along its temporal axis." (Barker, 1963)

The human mind has a gift for bringing order to chaos. The world presents nothing but continuity and flux, yet we seem to perceive activity as consisting of discrete events that have some orderly relations. This ability guides our understanding of what is happening, helps control our actions in the midst of it, and forms the basis of our later recollection of what took place. An inability to perceive events as such would be even more debilitating than an inability to perceive objects—a hyperbole of blindsight. This paper addresses the processes of event structure perception and the knowledge structures on which it relies.

Here we consider the following archetype for an event: a segment of time at a given location that is conceived by an observer to have a beginning and an end. We will call the process by which observers identify these beginnings and endings, and their relations, <u>event structure perception.</u> This conception of event has several implications which we will elucidate in the following pages. First, events are in the minds of beholders, yet tied to actions in the world. Next, the temporal dimension of events leads to an inherent asymmetry in event boundaries and organization. Moreover, perception of temporal sequence is elementary to perception of causality.

Of course, everyday usage of the term "event" is more general than this archetype. While ordinary usage doesn't distinguish between types and tokens, this is a characterization of tokens. Some research on knowledge representations concentrates on event types; these can be thought of as categories of event tokens. Events can be perceived on different temporal scales, spanning the evolution of the universe and the collision of subatomic particles. They can have imprecise spatial or temporal boundaries, as might a large outdoor party. Events can even have spatial or temporal discontinuities, as in the celebration of New Year's or a meeting interrupted by a fire alarm. Our aim is to give an account of the modal case, of the everyday events that people commonly discuss. The research reviewed here and the proposals we make will apply to varying degrees to the deviations from the archetype.

Event structure perception and the conceptual processes on which it depends have played a tangential role in a large number of research programs. We believe there is considerable leverage to be gained by examining human understanding of event structure directly. As will emerge, this investigation has implications for language processing, memory, planning, and action.

In the sections that follow we take a series of excursions into the several literatures that address mental representations of events. The first takes up the philosophy of actions and event descriptions, providing a logical foundation for events. The second surveys the literature on object perception, contributing heuristic analogies—and dis-analogies—to event perception. The third goes into real-time methods for studying event structure, providing direct evidence of how people understand events as they happen. The fourth reviews how events can be characterized in terms of their qualitative and quantitative features, pointing to how structure within event units contributes to structured relations between them. The fifth describes theories of the mental representations that capture such relations between components of events. The sixth examines how artificial intelligence models have used structured event representations to understand stories and plan action.

These excursions will point to a number of points on which these disparate approaches converge. These convergences do not constitute a theory of event structure perception and conception per se, but they do provide strong constraints on any such theory. First, event structure manifests pervasive covariation between multiple sources of information. Second, human performance indicates the presence of representations that relate events on a fine temporal grain to events on a coarse temporal grain. Third, both perception and conception of events show evidence of interactions between effects of natural selection and effects of experience-driven learning. All three constrain the class of models that can account for the perceptual and conceptual phenomena.

The Logical Structure of Events

A science of event perception needs its atoms. Recent philosophy has made foundational contributions to two ontological issues: First, "What are the basic units of action?" In other words, what are the basic units into which intentional human action can be analyzed. Second, "What are events?" There has been a vigorous debate concerning how events should be conceptualized relative to objects, properties and propositions. Our discussion of both issues relies heavily on an excellent collection of papers edited by Casati and Varzi (1996). Readers more interested in the psychological issues than the philosophical debates may wish to skip to the last paragraph of this section.

Atomic Components of Events: Basic Actions

Contemporary philosophy has provided an instructive analysis of the atomic units of actions, which may be of substantial value to psychological theory. However, two caveats are in order. First, “action” is not synonymous with “event.” Many of the events we observe are actions (hitting a foul ball, buying a car) but many are not (a candle blowing out on a windy day, a partial eclipse). Actions are performed intentionally by actors, so they are less general than events. Also, actions happen objectively in the world, while for our purposes as psychologists events arise in the perception of observers. That said, Danto (1963) defines a basic action as follows: “B is a basic action of a if and only if (i) B is an action and (ii) whenever a performs B, there is no other action A performed by a such that B is caused by A.” Goldman (1970) offers a critique of this definition and a replacement. In Goldman’s definition, the notion of causal dependence (in ii) is replaced by a more general kind of supervenience called “level-generation.” Examples of basic actions for most people under both definitions include “raising one’s hand,” as in signaling to answer a question, and “moving one’s head,” as in craning one’s neck to see better. Note that while both raising one’s hand and moving one’s head are basic actions, neither signaling to answer a question nor craning one’s neck to see better is. The latter are non-basic actions derivative of the two respective basic acts. Basic actions identify a class of actions that are logical as well as behavioral primitives: All basic actions are importantly physically primitive, but not all physically primitive actions are basic actions (e.g., craning one's neck to see better).

What are events?

Regarding the second question, contemporary analytic philosophy has made a number of proposals as to how to characterize events. One is to treat events as logical particulars, that is as constants that can bind variables in first-order predicate calculus (Davidson, 1966/1996). Another way to put this is that events should be represented as primitives, just as are objects. Sentences like "Phoebe fed a coelacanth" are thus treated as "There exists an event x such that Fed(Phoebe, coelacanth, x)." On this account, events are individuated with regard to their causal properties; if one event can substitute for another in all its causal relations, the two are identical. This logical argument can be taken to support an ontological conclusion that events have the same ontological status as objects such as Phoebes and coelacanths (Davidson, 1970/1996b). However, this conclusion has been vigorously disputed (Hacker, 1982/1996). Horgan (1978/1996) has argued against events as particulars based on parsimony, suggesting instead that talk about events be reduced to talk about facts. Horgan argues that while Davidson's (1970/1996a) proposal works technically, if one rejects the ontological claims the reductive strategy is simpler and therefore a better theory.

An alternative is that events are not particulars, but exemplifications of a property P by some substance (particular) x at time t (Kim, 1975/1996). On this view, events are complexes of primitives and relations rather than primitives. This has the psychologically problematic feature of being ontologically promiscuous. It provides no constraints on the number and range of properties a single object can exemplify at a single time, so the number of events in which it participates is unbounded. However, this is not necessarily a fault when considered from the viewpoints of ontology or logic. Bennett (1996), in attempting to constrain the proliferation of events allowed under this view, argues that what might be unbounded are not property-exemplifications per se, but descriptions of property-exemplifications. Bennett also considers the possibility that events might be temporal parts of objects, but rejects it on the grounds that it fails to cover some well-known cases. For example, if a ball is both heating and rotating, one may wish to refer to these as separate events, though they involve the same thing over the same time-period. This apparent failing may represent a real break between the goals of philosophy and of psychology. The fact that the temporal-part proposal fails to treat these as separate events seems intuitively plausible as a description of an observer's experience, though it may be inappropriate for some aspects of formal description.

The simplest proposal as to what events are was made by Quine (1985/1996): Simply treat events as objects. That is, regard events as bounded regions of space-time. What we typically think of as objects (chairs and tables) are one family of bounded space-time regions, events are another. This is similar in spirit to Davidson's (1970/1996a) events-as-particulars account, but more radical. In psychology, Miller and Johnson-Laird (1976, p. 87) have made a related proposal, that events are dynamic objects and what we call "objects" are concrete objects. One attractive feature of Quine's proposal is that it eliminates the difficulties involved with individuating events based on their causal relations. It is also the philosophical definition closest in spirit to the psychological characterization given earlier. The latter (a segment of time at a given location conceived by an observer to have a beginning and an end) simply defines a psychological event as the perceptual experience of a Quine-event.

This excursion into philosophy provides three insights into the psychology of events. First, actions can be analyzed into primitives and relations, suggesting a similar strategy can be fruitfully applied to events. Second, though one can construct a respectable argument that logical talk of events should be reduced to talk about facts, such logical sanitization may be psychologically implausible. Third, one can also reasonably argue for treating events as one treats objects. In the next excursion, we will show there are powerful psychological reasons for doing so.

Lessons from Objects

Event perception can be regarded as the <u>temporally extended</u> analog of object perception: Events are objects in the manifold of the three dimensions of space plus the one dimension of time. Indeed, Asch (cited in Newtson, Hairfield, Bloomingdale, & Cutino, 1987) described events as <u>gestalts</u> in the stream of activity that flows through time, much in the same spirit as Quine's (1985/1996) argument for equivalent treatment of objects and events. On that view the distinction between what we call "events" and what we call "objects" in everyday language marks the presence of interesting temporal structure; formally they are the same. Observers recognize objects by their distinctive shapes, colors, textures, tactile properties, and motion. They recognize events based on these features of their component objects and configurations of objects, but also on the basis of their temporal structure. In this section we will take this analogy seriously, and will see that many attributes of object perception generalize naturally to events.

Objects have boundaries in space. A coffee cup takes up a certain amount of space of a certain shape. One can perceptually identify where it is and where it is not. Similarly, events have boundaries in time. The event "picking up a coffee cup" has a beginning and an end; it takes up a certain amount of time. Events also are psychologically bounded in space. A particular instance of picking up a coffee cup occurs at a particular location that can be spatially bounded like the coffee cup itself. (In what follows, we will mostly assume that the spatial boundedness of events is more-or-less accounted for by a theory of perception of the objects involved. This is a natural assumption, but one that could bear some empirical attention.)

Parts of Objects

Objects have parts with a particular spatial configuration. A car has parts like doors, windows, an engine, wheels, and seats. These parts can in turn be divided into sub-parts; for example, a seat generally consists of a bench, a back, a seatbelt, and a headrest. The hierarchical relationship between parts and sub-parts constitutes a partonomy (Miller & Johnson-Laird, 1976; Tversky, 1990; Tversky & Hemenway, 1984). Partonomic relationships give rise to distinctive spatial configurations that can be of use in categorizing objects. In some situations objects can be quickly classified based on their shape (Rosch, 1978). Where parts join, they give rise to distinctive physical features: contour discontinuities or maxima in local curvature (Biederman, 1985; Hoffman & Richards, 1984).

There is evidence that people depend especially heavily on those perceptual features that discriminate parts when identifying objects: When contour discontinuities are deleted from line drawings of objects, this interferes more with categorizing the objects than the deletion of smooth contours of similar size (Biederman, 1985). This fact has been taken as evidence for a process of object recognition by components (Biederman, 1987). Another possibility is that distinctive cues are used directly without recovering part structure per se; this view has received support from experiments demonstrating view-dependent object-recognition (e.g, Bülthoff, Edelman, & Tarr, 1995; Tarr, 1995). However, the relevance of the empirical finding for present purposes stands, independent of the theoretical interpretation: Contour discontinuities, associated with parts, are important for object recognition.

Taxonomies and the Basic Level

A partonomy is one common form of hierarchical structure that characterizes objects. Another hierarchy that characterizes objects is a taxonomy, which is based on "kind of" (rather than "part of") relationships. For example "Jeep Wagoneer" is a kind of "Automobile," which in turn is a kind of "Vehicle." Taxonomic hierarchies also apply to events. For example, "Frisbee Golf" is an (atypical) kind of "Golf," which in turn is a kind of "Sport." Note that it is important to avoid conflating taxonomic and partonomic hierarchies (Miller & Johnson-Laird, 1976; Tversky, 1990). Principles that have been demonstrated for the former may not generalize to the latter.

Objects and events, then, can be referred to at varying levels of abstraction. In classic studies, Rosch and her colleagues showed that one level, the basic level, is more privileged in perception, action, and language. That level, the level of "table" rather than "furniture" or "coffee table," conveys the relatively greatest amount of information (Rosch, 1978; Rosch & et al., 1976). Members of basic-level categories have more objective similarity in shape than objects at a superordinate or subordinate level, and are easier to recognize when their shapes are averaged together. They have more common motor movements directed toward them than superordinate categories (but not than subordinate categories). They are the most abstract categories for which an image can be generated. Naming instances is fastest at the basic level (Rosch, 1978). Basic-level categories are the categories favored by adults in a neutral context, and tend to be the earliest concepts learned by children.

Objects can be viewed both taxonomically and partonomically. Each organizes information differently and promotes different inferences. Taxonomic organization promotes reasoning about intrinsic properties: "Coelacanth" is a member of the taxonomic category "fish;" therefore, it is likely to have gills. Partonomic organization promotes reasoning from physical structure to function and cause. The presence of legs allows one to infer standing, whether the object in question is a chair or a chameleon. The presence of skin allows one to infer protection, whether the object is an orange or an orangutan (Tversky, 1990; Tversky & Hemenway, 1984).

Partonomies and Taxonomies of Events

Partonomy

Like objects, events can be viewed as organized into partonomic hierarchies, reflecting relations between parts and sub-parts. Barker and Wright (1954) provide a wonderful example, reproduced here as Figure 1. The figure shows how part-of relationships in activity can be traced from moments to years (with perhaps absurd consequences at the extremes). Barker and Wright also found that behavior in natural situations is in fact naturally described hierarchically. In extensive observations of children going about their daily lives, 73% of the behavior episodes recorded were at least partially coextensive with other nearby episodes. Of these, 90% of the overlaps were partonomic relationships (enclosed or enclosing). This suggests that much naturally occurring behavior is in fact perceived by observers as partonomically organized. Moreover, within relatively homogenous samples of participants there is good agreement about what make up the typical parts of everyday activities (e.g., Bower, Black, & Turner, 1979; Slackman, Hudson, & Fivush, 1986).

INSERT FIGURE 1 ABOUT HERE

Event partonomies may have a privileged level at which cognition is particularly fluent. Barker and Wright (1954) argue for such a level, which they call a behavior episode. "Behavior episodes are analogous to physical objects which can be seen with the 'naked eye.' They are the common 'things' of behavior; they correspond to the stones, chairs, and buildings of the physical world" (p. 6). Examples of behavior episodes include: a group of boys moving a crate across a pit, a girl exchanging remarks with her mother, a boy going home from school. Six characteristics tend to mark the boundaries of behavior episodes (p. 235):

1. A change in the "sphere" of the behavior between verbal, social, and intellectual.
2. A change in the predominant part of the body.
3. A change in the physical direction of the behavior.
4. A change in the object of the behavior.
5. A change in the behavior setting.
6. A change in the tempo of the activity.

At the boundaries between behavior episodes, at least one of these changes usually occurs.

Another approach to characterizing a privileged partonomic level for events has identified it with the scene level in a script theory of story understanding. For example, the scenes in the "restaurant" script are "entering," "ordering," "eating," and "exiting" (Schank & Abelson, 1977). People seem to show good agreement for what constitute the scenes in an activity (Bower et al., 1979). Further, when presented with subordinate-level actions, people tend to make inferences up to the scene level; but when presented with information at the scene level, they are relatively unlikely to make downward inferences (Abbott, Black, & Smith, 1985).

Taxonomy

Again like objects, events can be viewed as organized into taxonomic hierarchies, reflecting “kind of” relations. Events can be described at a variety of taxonomic levels, of which there seems to be a preferred basic level. As with objects, the number of features listed for events increases greatly from the superordinate to the basic level, but not very much from the basic to the subordinate level (Hemeren, 1996; Morris & Murphy, 1990; Rifkin, 1985). Following the research program of Rosch and her colleagues, Morris and Murphy (1990) applied a set of converging methods to look for characteristic features of basic-level categories in the domain of events. In one experiment, they presented participants with excerpts from event descriptions (e.g., “scream during the scary parts”) and then asked them to verify a category label at one of three levels: subordinate (“horror movie”), basic (“movie”) and superordinate (“entertainment”). (Half of the time an incorrect category label was presented.) Responses were fastest to the basic-level labels. In another experiment, participants read simple stories and then were asked to name them. They tended to use basic-level names, except when subordinate-level names were required to distinguish between stories. Morris and Murphy did find one way in which event categories differed from what is typically found for object categories: Event categories were most differentiated at the subordinate level, rather than the basic level. That is, rated within-category similarity increased more from the basic level to the subordinate level than from the superordinate level to the basic level.

The scene level of a script has also been identified as "basic-level," based on the special properties of scenes described previously (Abbott et al., 1985), and by a similar argument one might want to apply the "basic-level" label to behavior episodes (Barker & Wright, 1954). However, as noted by Abbott, et al., the hypothesis that the scene level is basic in the sense used by Rosch and her colleagues (Rosch, 1978) risks a category mistake. The scene is a level in a partonomic hierarchy, while the "basic-level" is typically used to refer to a level of abstraction in a taxonomic hierarchy. While there may be a general sense in which the notion of the basic level can be imported from taxonomies to partonomies, some aspects are bound not to transfer. It seems safer to refer to privileged partonomic levels such as behavior episodes or scenes distinctly, rather than to use "basic-level" to refer to parts as well as kinds.

There is a potential confusion between the notion of a basic-level event and of a basic action described previously, because both contain the word "basic." Basic-level events refer to a psychologically privileged taxonomic level, while basic actions refer to a foundational causal level.

Where Parts and Kinds Meet

The special features of basic-level event categories are analogous to the non-accidental convergence of perceptual and conceptual features of object categories. Basic-level objects tend to be those that have well-defined parts (Tversky, 1990; Tversky & Hemenway, 1984). Parts dominate the new features that are added in going from the superordinate level to the basic level. Different basic-level categories differ in the parts their members contain, while subordinates within a basic-level category share many parts. Knowledge about parts is contained in concepts that address the function of objects and how parts interact. Moreover, we have already seen that partonomic structure gives rise to distinctive perceptual features: Junctions of parts form perceptually distinctive inflections (Biederman, 1985; Hoffman & Richards, 1984). Parts rated high in "goodness" tend to be functionally significant as well as perceptually salient, such as the wheels of a car or the trunk of a tree (Tversky & Hemenway, 1984). Thus, conceptual information about how objects work and how parts fit together aligns naturally with perceptual characteristics of objects.

The same may be true for events. When asked to identify event boundaries, people tend to divide activity at locations that correspond to maxima in the number of physical features that are changing. In one analysis, films of human activity were first divided into natural parts by naïve viewers. The position of the actors was then coded using dance notation. The locations in time where the naïve observers had segmented the activity tended to corresponding to points at which the positions of actors' bodies were changing the most (Newtson, Engquist, & Bois, 1977). In the analogy between objects and events, these maxima in feature changes are like contour discontinuities in objects. They mark the boundaries of parts, and correspond to locations of maximal perceptual change. As with objects, there is evidence that these points are particularly important for event identification. When slide shows are made from a movie of some activity, sequences made of perceptually-identified event part boundaries are more intelligible (i.e. better categorized) than sequences of non-event part boundaries (Newtson & Engquist, 1976). Similarly, low-bandwidth transmission of sign language is improved by selectively choosing frames in which the most change occurs (Parish, Sperling, & Landy, 1990). There is evidence that young infants are able to use such perceptual event boundaries to count events, even in the midst of continuous activity (Wynn, 1996).

Thus, there is a range of time-scales within which humans are sensitive to event part structure. However, this range is surely limited, and the kinds of features that are salient at different time-scales may vary, just as the kinds of visual features to which humans are sensitive varies with spatial scale. The smallest psychologically reified events, on the order of a few seconds, may be defined primarily in terms of <u>simple physical changes</u>. Think of a person grasping another's hand, the hands going up, going down, releasing. Longer events, from about 10 to 30 seconds, may be defined in relationship to some straightforward intentional act: the events described above, seen on this time-scale, form a handshake. From a few minutes to a few hours, events seem to be characterized by <u>plots</u>, i.e. the goals and plans of their participants, or by socially conventional forms of activity. Perhaps the handshake was part of signing a treaty. On long enough time-scales it may be that events are characterized <u>thematically</u>. In this example, perhaps the treaty signing was part of an event called a "peace process." In general, it seems that as we increase the time-scale of our view, events become less physically characterized and more defined by the goals, plans, intentions and traits of their participants.

Barker and Wright (1954) give a related characterization of the nature of behavior parts at different time-scales. They argue that small units tend to be related to minor subgoals and thus may go unnoticed by the participants. Large units, which may be related to large ongoing goals, may go similarly unnoticed. Between these two extremes are behavior episodes. In Figure 1, "walking to school" is likely to be a salient goal, but neither "stepping down from the curb" nor "climbing to the top in life" seems likely to be present to the mind of a perceiver at that moment.

Time, Objects and Events

So we see that events and objects possess similar partonomic and taxonomic features, and these features interact in the same ways. This may partially explain why in everyday language we often talk about events metaphorically as objects. As Lakoff and Johnson (1980, p. 30-31) point out, expressions such as "going to the concert," "see the concert," "there was a lot of good music in the concert" are ubiquitous object metaphors for events. Because both objects and events are related by parts and kinds, our knowledge of objects can form an experiential basis for thinking metaphorically about events.

However, there is at least one failure of the analogy between objects and events that is probably important for understanding event perception. Any given instance of an object (say, a particular teacup) persists across time, while events are necessarily ephemeral. This difference is easy to see when one considers a medium in which time is represented by space: comics. Comics depend critically on the difference between how objects and events behave with respect to time, exploiting the failure of the normal limits on event segmentation for artistic effect. As McCloud (1993) has pointed out (see Figure 2), the creator of a comic can control the way time is segmented by manipulating placement and graphical features of panels. (McCloud goes on to assert that the perception of time in comics is in fact determined more by the content of the frame than by these representational devices.)

INSERT FIGURE 2 ABOUT HERE

The difference in the role played by time in object perception and event perception has several consequences. First, it means that for objects we have the ability to recognize a particular instance as such, to recognize a particular coffee cup rather than simply recognize it as an instance of a coffee cup. A given event occurs exactly once, so while we can perceptually categorize it as, say, an instance of a restaurant visit, we will never experience that particular gastronomic indulgence again. In other words, we can perceptually categorize events, but not perceptually identify them. Of course, on the other hand we can identify individual events in memory. Moreover, technologies such as theater and recordings make it possible to have experiences that may approximate repeated perception of an event—though these experiences w ere surely not part of the environment in which the human cognitive architecture evolved.)

Second, when we consider applying how people categorize, segment, or interact with objects to understanding event perception, we should focus particularly on conditions in which exposure to the object in question is fleeting. A case can be made that our perception of the parts of events is analogous to the identification of object parts based on one visual fixation. Under normal conditions for studying object perception (and most of the time in life outside the laboratory), observers are exposed to objects for some length of time, allowing for repeated viewings, multiple views, and perhaps even interactive exploration. When visually processing objects in one fixation, attention can be driven by prior expectations to one or another feature of the object, but the observer only has one chance to collect information about the stimulus. Similarly, when processing an event, attention may be guided by conceptual expectations, prior experience with that class of event, or perceptual characteristics of the part of the event that has unfolded so far, but each moment in time is (of course) experienced only once (see also Miller & Johnson-Laird, 1976, Ch. 2).

This excursion into object structure perception contributes four points to the understanding of event structure perception. First, both objects and events form partonomic structures. Second, both objects and events form taxonomic structures, with a privileged basic level. Third, for both objects and events, there are systematic relationships between partonomy and taxonomy: at the basic level, objects and events are characterized by good parts. Finally, there is an important difference between objects and events: Objects can be reexamined—and as a result, re-identified—while a given event can only be experienced once.

Event Segmentation

We have seen that the similarities between objects and events suggest fruitful methods for exploring event structure perception. So do their differences. Object part boundaries are arranged in space, so for objects, perception of part boundaries can proceed at least in part in parallel. However, if we ask observers to report the location or other features of object part boundaries, they must do so serially, making it difficult to elicit object part boundaries without disrupting the way in which they are normally processed. For events, on the other hand, part boundaries occur primarily with respect to time. This means that a simple procedure can be designed to allow observers to report the event boundaries they perceive as they happen, without overly disrupting the normal processes of event perception.

Such a procedure has been developed by Newtson and his colleagues (Newtson, 1973). Observers watch a film or videotape of an event and simply tap a key "whenever, in [their] judgment, one unit ends and another begins." The points in time at which observers tap are called "breakpoints." It has been reported that repeated testing of the same observer produces similar breakpoints, and that there is reasonably good agreement across observers as to what the breakpoints are (Newtson & Engquist, 1976). Breakpoints tend to correspond to points at which the most physical features of the action are changing. This has been shown by coding actor motion in terms of dance notation and noting a correspondence between changes in body position and segment boundary locations (Newtson et al., 1977). More recently, similar results have been obtained with simple image analysis techniques for American Sign Language (Parish et al., 1990) and with sophisticated optic flow analyses for more natural stimuli (Rui & Anandan, 2000).

This objective basis of event boundaries, combined with reasonable inter-rater reliability, argues for their psychological salience. This conclusion is strengthened by a set of studies of the role of breakpoints in event comprehension and memory (Newtson & Engquist, 1976). In the first of these experiments, brief deletions in a film were better detected when made at breakpoints than non-breakpoints. In the second, slide shows made from sequences of breakpoints were more intelligible than slide shows made from sequences of non-breakpoints, and were easier to reorder when presented out of order. In the final experiment, recognition memory was found to be better for breakpoints than for non-breakpoints. While these experiments examined relatively brief activity sequences (on the order of minutes, with segments on the order of tens of seconds), similar results have been obtained at longer time-scales (Boltz, 1992b; Boltz, 1995).

A number of results obtained with this event segmentation technique suggest that observers actively modulate the level at which they divide activity into events. Observers can be induced to subdivide activity at a range of levels. The simplest way to do this is by direct instruction: ask observers for either the largest or smallest units with which they feel comfortable (Lassiter, Stone, & Rogers, 1988; Newtson, 1973). However, viewers also spontaneously adjust their level of segmentation, depending on their knowledge or goals. People tend to divide the stream of behavior into smaller units when it is unpredictable (Newtson, 1973; Vallacher & Wegner, 1987; Wilder, 1978a; Wilder, 1978b), or when they are instructed to focus on the task rather than impressions of the actor (Cohen & Ebbesen, 1979). In one study, grade-school students were asked to segment a videotape of another child. Providing information about the subject of the videotape before viewing led to the production of larger units (Graziano, Moore, & Collins, 1988). This sort of spontaneous variation in segmentation level has been argued to reflect an impetus to maintain a coherent understanding of what is happening around us, while expending minimal perceptual/cognitive resources. When a coarse temporal grain is insufficient to achieve this understanding, we shift to a finer grain of encoding (Newtson, 1973). A similar argument has been made for perception of our own activities (Vallacher & Wegner, 1987).

However it is induced, fine-grained segmentation lead to better recognition memory for the physical characteristics of the activity (Hanson & Hirst, 1989; Lassiter et al., 1988; Newtson & Engquist, 1976). There is currently some debate about whether this holds for recall memory as well (Hanson & Hirst, 1989; Hanson & Hirst, 1991; Lassiter & Slaw, 1991). If we accept the speculative generalization given in the previous section, that events tend to become less tied to physical activity as they are viewed on longer time-scales, we should expect superior memory for the physical characteristics of events when they are encoded in shorter temporal units.

Partonomies and Event Segmentation

We have previously seen that part information is valuable for recognizing object and events. We will shortly describe evidence that partonomic hierarchies play a substantial role in text comprehension and memory, and in event comprehension. It is therefore natural to ask if and how people actively construct partonomic representations of ongoing activity as it happens. Newtson and his colleagues have argued that against this hypothesis, and more generally against the cognitive representational view of event perception (Newtson et al., 1987). They claim that the structure of event perception can be fully accounted for by the "topology" of the behavior stream, that is by its perceptual characteristics. By this view of event perception as a stimulus-driven phenomenon, the parts of events are identified based purely on perceptual characteristics such as points of maximal change in physical features. Ebbesen (1980) gives two arguments against the view that observers spontaneously construct a partonomic hierarchy of activity. First, it may be incorrect to assert that representations of event parts exert obligatory or primary effects. The cognitive structure employed is argued to depend flexibly on the task being performed. Other orthogonal structures, such as schemata for personality types, may exert primary control of many tasks. Second, there may be multiple layers of event representation, such as an abstract or impressionistic code and a detailed recording, that are not hierarchically related.

Indirect evidence on the structure of perceptual event segmentation comes from a study by Dickman (1963). In this experiment, a film sequence was divided by the experimenter into a series of fine-grained event parts. Experimental participants then grouped these parts into larger units. Considering event beginnings and endings separately, there was good agreement across participants. However, considering beginnings and endings together led to relatively poor agreement. This apparent paradox was resolved by noting that coders tended to use different unit sizes, and that a single grouping of a coder who used large units might contain several groupings of another coder who used smaller groupings. Across coders, the structure that emerged was a partonomic hierarchy. Though suggestive, this finding bears only weakly on perception, because the experimental participants started with the pre-segmented sequence, rather than something close to the perceptual event itself.

More recently, we have collected direct evidence suggesting that in fact the human perceptual system does actively encode ongoing activity in terms of partonomic hierarchies (Zacks, Tversky, & Iyer, in press). We presented observers with videotapes of four everyday, goal-directed activities: making a bed, doing the dishes, fertilizing houseplants, and assembling a saxophone. Each participant segmented all four activities twice, once under instructions to produce the smallest units that were natural and meaningful, and once under instructions to produce the largest units that were natural and meaningful. One group also described each unit after tapping to mark its end. Order was counterbalanced, and there was a delay between the two segmentation sessions. We observed a hierarchical relationship such that large-unit boundaries were disproportionately likely to also be small-unit boundaries. This pattern is similar to that observed by Newtson (1973) using a between-subjects design (but see Ebbesen, 1980). The hierarchical structure effect was somewhat more pronounced for the familiar activities (making a bed and doing the dishes) than for unfamiliar ones (fertilizing houseplants and assembling a saxophone), and more pronounced when observers described the activity as it happened than when they simply segmented the activity. These interactions imply that cognitive representations of events influence segmentation in a top-down fashion, in addition to bottom-up perceptual information.

Another approach to studying structure in event perception has been taken by Hanson and Hanson (1996). They modeled observers' event segmentation using a recurrent neural network. The pattern of hidden unit activations across time was suggestive of a hierarchical organization to the behavior. Further, the network showed two properties that may help explain the nature of the mechanisms driving event structure perception. First, it was influenced both by bottom-up and top-down processing: Bottom-up processing comes from activation of internal representations (patterns over the hidden nodes) by characteristic perceptual activation (patterns over the input nodes), while top-down processing comes from the hidden nodes' tendency to resist changes in activation (due to the recurrent connections). Second, the network's learned expectancies about event durations influenced the network's sensitivity to new perceptual information.

Together, the mass of data from the segmentation technique and related approaches argues strongly for a bottom-up component to event structure perception: event unit boundaries are conditioned on physical features of the activity. However, effects of experience, instruction, and expectation on segmentation patterns argue for top-down influences. The segmentation technique provides a method for directly observing these influences in observers' behavior.

Causality

The study of causal perception provides another window on top-down and bottom-up processing of event structure. The perception of causality may be a key feature in defining event structure, because the moments at which we perceive causal interactions taking place (one billiard ball launching another, a skidding car stopped by a lamp-post) tend be critical to the structure. In describing a comprehensive set of studies, Michotte (1946/1963) argued that physical causality is epitomized by the phenomenon of "ampliation of the motion." In this analysis, the perception of causality is a temporally limited phenomenon that is characterized by a tension between the individuality of the objects in question and the perceptual integrity of a motion that transfers from one object to another. When one object contacts another and gives the percept of "launching" it, there are clearly two spatial displacements involved. The launcher is displaced and the launchee is displaced, and these displacements are different. However, when the causal percept of launching occurs, there is a short period of time during which the motion of the launcher is projected onto the launchee, and there is only one motion perceived.

Now, such causal interactions between objects are likely to correspond to "contour discontinuities" in the temporally extended event, i.e., to maxima in the physical features of the objects. Thus, they tend to have the perceptual characteristics of event part boundaries. In the case of a simple launching effect, the moment of the causal percept occurs when the launcher stops and the launchee starts. This is precisely where the most physical features of the activity are changing. Thus, we would predict that the most causally "loaded" moments in an event would be the points at which observers would tend to subdivide it. (To the best of our knowledge this prediction has not been tested.)

INSERT FIGURE 3 ABOUT HERE

More generally, we can ask: What do causes do? Simple low-level goals are often satisfied by the occurrence of a particular physical movement that is perceived causally. As examples of "launching" effects that coincide with goal satisfactions, think of a hockey player hitting a puck or a cook flipping a pancake. Often, they satisfy goals. Thus, even for cases where

Michotte's theory of the perception of physical causality is not clearly applicable, it may be the case that moments of maximal feature changes tend to correspond to the satisfaction of goals. This might be an explanation for the importance (noted previously) of these physical event part boundaries in event identification (Newtson et al., 1977). If we categorize events in part by matching them to stored representations of goal relationships, moments of causal interaction should be especially important for identifying candidate goal relationships.

The physical parameters of perceived causal events provide information about bottom-up processing of event structure. At the same time, these perceptually significant configurations can be tied to top-down cognition about goals and intentions. Thus, Michotte's pioneering studies provide a bridge between direct, data-driven perception and complex cognition. In following sections we will see that relations between goals and causes are central to conceptual understanding of events.

This excursion into event segmentation makes three important points for the study of event structure. First, the temporal parts of events can be effectively studied using on-line segmentation techniques. Second, events appear to be perceived in terms of a partonomic structure. Third, events possess a causal perceptual structure as well as a partonomic perceptual structure, and the two may be highly correlated.

Characterizing Event Units

The previous sections have dealt with relations between event parts: partonomic relations across timescales, taxonomic relations across levels of generality, and segmentation relations within a timescale. However, within an event unit at a given timescale and taxonomic level there may be salient features that play a role in event structure perception. Qualitative features of event units can be discovered by linguistically analyzing the general patterns in how we talk about events and by experimentally examining descriptions of events. Quantitative features of event units can be discovered by looking at the statistical properties of perceptually meaningful event units. In this section we will see that both types of analysis support the view that thinking and talking about events depends on structured representations.

Qualitative Characterizations

Structural Features of Event Language

The ways in which language carves up activity provide information about our categories for events. One promising place to start is the consideration of motion events. Talmy (1975) argues that there is a paradigmatic structure to motion events: they consist of figure (which is a nominal), a motion (which is a verb), a path (which is a prepositional) and a ground (which is a nominal). All sentences that express a motion situation express this structure, though different languages particularize it in different ways. Motion situations are more general than might at first be obvious, covering sentences such as “It rained into the room,” and “The bottle was floating in the cove,” along with more obvious cases such as “The ball rolled across the border.”

One can easily generate a psychological theory from this linguistic account: The building-blocks of events should be temporal units in which the figure, motion, path and ground are constant. A change in any of those features of the situation constitutes a new atomic event. Consider a scene in which a person is riding a ski lift. As she gets off the lift and continues by skiing (a change in the motion), it seems likely this would be perceived as a new event. When she turns (a change in path) at the base of a ski jump, that also seems likely to mark a psychological event boundary. Finally, as she accelerates down the jump, and then leaves it and continues in free flight (a change in ground), that would likely be perceived as another new event.

However, it's not clear from this analysis what a change in the figure of the motion would be. It may be that changes in figure play a different role in event perception: focusing attention on a particular part of the perceptual world. (Consider the following pair of sentences: "Sonia's hand raised the poisoned cup to her lips. Her eyes looked questioningly at Jean-Luc." The second sentence doesn't describe a new event, but focuses the reader on a new set of features.) Event structure perception is then driven by changes in the motion, path or ground relative to the currently attended figure.

The relationship between the observational characterization of activity given by Barker and Wright (1954, see "Partonomies and Taxonomies of Events" above), and the linguistic characterization of motion sentences given by Talmy (1975), can be seen by comparing the basic features in each scheme. The features path, object, and ground in Talmy's theory have clear parallels in three of Barker and Wright's features: physical direction, object, and behavior setting. However, we should keep in mind that Talmy's linguistic account is much broader in time-scale, accounting for atomic events (e.g. "Rick raised his right arm") as well as molar events (e.g. "The plane took off for France").

This theoretical analysis suggests that discourse about activity is governed by structured representations of events. This position is supported by recent work on verb tense and aspect. A convincing case can be made that linguistic phenomena such as adverbial modification and argument structure for verbs can best be understood by reference to a structured representation of events (Pustejovsky, 1991). Grammatical considerations distinguish states from processes, accomplishments from non-accomplishments. Adverbial scope can be explained naturally when verbs are understood to quantify over events; event part structure provides a consistent way to resolve ambiguity. Thus, linguistic structure points to an underlying representation of events with internal structure of their own.

Further, temporal reference (tense, aspect, etc.) can be well explained in terms of a cognitive representation of goals and preparatory processes in the service of goals. Moens and Steedman (1988) argue that the representations underlying event descriptions reify a goal state or culmination as well as a preparatory process or antecedent. For example, "Harry has reached the top" marks an antecedent in which Harry wasn't at the top and a culmination, in which he is. "Harry climbed to the top" marks a preparatory process (climbing) and its culmination. Based on this characterization, they give accounts of aspect, tense, temporal focus and reference to future events. They argue that many cases that on the surface appear to be temporal references actually refer to contingency relations between antecedents and goals. In the case of "Harry climbed to the top," the tense of 'climbed' marks the contingency relation between climbing and reaching the top . Thus, event representations are (a) structured and (b) closely tied to goals and causes.

Recently, Narayanan (1997) has accounted for several features of aspect using a computational model that ties verb semantics to structured representations of events based on physical motor-control primitives such as goal, periodicity, iteration, final state, duration, and parameters such as force and effort. The model is explicitly structured, and includes recursive and iterative devices for handling partonomic event structure.

There is also evidence that observers use causal features of event representations to formulate utterances. For example, Wolff & Gentner (1995) found a difference between verb constructions that correlated with differences in causal perception. When asked to describe direct causation in simple events, viewers tended to use lexical causatives (e.g. "sunk"), whereas when asked to describe indirect causation they tended to use periphrastic causatives (e.g. "made sink"). For describing the movements of inanimate objects (marbles), viewers were unlikely to use a lexical causative unless the objects made physical contact. When the marble's movement was initiated by a person, they were substantially more likely to use a lexical causative in the absence of physical contact—though contact still increased lexicalization. Both physical contact and initiation by an animate agent are diagnostic of a unified causal sequence, and both increased the tendency to use lexical causatives.

Descriptions of Ongoing Events

These architectural features of talk about events suggest that when people speak, they rely on structured representations. We have recently observed evidence from experimentally-manipulated descriptions of ongoing activity that supports the view that people spontaneously construct partonomic event representations, and that these representations are important for speaking. In the segmentation experiment described previously (Zacks et al., in press), we asked some observers to describe each unit as they segmented the activity. Each observer produced two sets of descriptions and two sets of unit boundaries for each activity, one at the smallest natural and meaningful timescale, and the other at the largest. Overall, the descriptions tended to be telegraphic accounts of actions on objects (e.g., "moves things aside," "opening the pillowcase").

There were systematic differences between descriptions of large and small units. Large-unit descriptions tended to specify objects more precisely than small-unit descriptions: In large-unit descriptions, objects were mentioned more often, were pronominalized less often, and were more semantically precise. Small-unit descriptions tended to specify actions more precisely than large-unit descriptions: In small-unit descriptions, verbs were repeated less often and were more semantically precise. Large units tended to be divided by objects and small units by different actions on the same object.

Moreover, the partonomic structure of the activity that appeared in participants' segmentation patterns also was reflected in their descriptions of small units. Based on the location of the large-unit boundaries, we singled out the small-unit descriptions that occurred adjacent to a large-unit boundary. For almost every syntactic or semantic feature we examined, these "special" small-unit descriptions were more similar to the large-unit descriptions than were the remaining small-unit descriptions. Together, these results suggest that the same partonomic structures that influence observers' segmentation of ongoing activity affect their descriptions.

Quantitative Characterizations

Event units can be characterized in terms of their qualitative properties, or statistically in terms of where they tend to occur. In this regard they resemble units of speech such as phonemes, syllables, or words. A growing body of research suggests that humans are sensitive to the statistical properties of language such that they can learn what the units of a language are based on their statistical properties alone. Saffran and her colleagues have shown that humans can learn the words of a simple artificial language based only on the transitional properties between syllables. This holds for adults (Saffran, Newport, & Aslin, 1996b) as well as infants (Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996a), and can take place even when the learners are not attending to the language they are hearing (Saffran, Newport, Aslin, Tunick, & et al., 1997). These behavioral findings are consistent with computation results showing that there are efficient algorithms for discovering word boundaries from available transition probabilities in appropriate corpora from which the word boundaries have been deleted (Brent, 1999; Dahan & Brent, 1999).

Similar results have been found for non-linguistic events in a limited domain. Avrahami and Kareev (1994) showed adult observers sequences formed from 3-second cartoon clips. A control group viewed a random sequence and was asked to indicate where the most significant natural break occurred. A new training sequence was then constructed, in which a sub-sequence of clips centered on the major break was embedded repeatedly, with intervening random clips. An experimental group then viewed this training sequence followed by the original sequence. After training, they tended to divide the original sequence before or after the previously-repeated sub-sequence, rather than within it, as the control subjects had done. This indicates that they had learned from the training tape (without explicit instruction) to treat the sub-sequence as a unit. In a subsequent experiment, participants were exposed to sequences that contained a repeating sub-sequence. This was embedded either in a context of surrounding clips that was either fixed or variable. Recognition memory for the repeating sub-sequence was reliably better for the variable-context group. In a third experiment, a similar manipulation was applied and recall was tested. For the group that saw sub-sequences repeat in varying contexts, there was a greater tendency for clips within a sub-sequence to cue each other during recall. Together, these results indicate that participants learned the sub-sequence boundaries based solely on their transitional probabilities, that they did so without explicit instruction, and that the learned sub-sequences cohered in later cognition. Cognitive representations of real-world events may arise in the same way. However, one caveat is in order: the materials used here were film sequences constructed so that the recurring basic units were exact repetitions of the same clips. This is unrealistic in that within a basic unit the movement sequence, background, setting, and objects are all identical from instance to instance. Thus, participants may have been learning based on simple features of the film clips that are not available in real experiences.

In a similar vein, Byrne (1999) has argued that animals can learn to imitate behavior without a causal or intentional understanding of the behavior, based only on the statistical properties of the behavior. As a prerequisite, the animal must segment the stream of behavior into discrete items. Then, patterns can be identified based on related segments' tendency to co-occur. Identified patterns can be copied, and potentially related to causes and intentions.

This kind of event-unit learning has not been directly modeled, but can be captured in the abstract by a computation model of semantic associations, Latent Semantic Analysis, or LSA (e.g., Deerwester, Dumais, Landauer, Furnas, & Beck, 1988; Landauer, 1998; Landauer & Dumais, 1997). LSA learns the relatedness between any two basic units based on their pattern of occurrence. Units that tend to occur together are inferred to be related. Importantly, two units that never co-occur are also inferred to be related if they share other units with which each does co-occur. LSA can infer larger chunks in a sequence of basic units by identifying transitions between adjacent units where this relatedness is low. Thus far, the model has only been applied to situations in which the basic units are written words, but Landauer and Dumais (1997) have argued that the same model can be applied to events, provided features that individuate the basic units can be identified (no small task).

In both the language-learning and the event learning cases, the experimental data show that humans can learn to identify based on purely statistical properties—by removing all the other information available. In the real world, statistical co-occurrence is not arbitrary. Rather, sub-sequences tend to repeat because they form components of larger plans (Newell & Simon, 1972), because they constitute parts of a behavior setting (Barker, 1963; Barker & Wright, 1954), or because they are generated by a common causal mechanism (Michotte, 1946/1963). Thus statistical learning may be a powerful way for learners to bootstrap themselves into appropriate associations between bottom-up perceptual information about the world and top-down cognitive understanding.

This excursion has demonstrated that individual event units can be analyzed in terms of both qualitative and quantitative features. We have seen that both kinds of information about individual events contribute to understanding the relationships between events.

Structured Representations of Events

For event categories that recur in our lives, our perceptions and actions are shaped by this repeated experience. One proposal is that they are guided by schemata that capture the common patterns of intercorrelations. (Our characterization of how event schemata work basically follows Rumelhart and Ortony, 1977. See also Rumelhart, 1980.) Repeated exposure to an event category leads to the creation of a schema for that category. These schemata have several distinctive characteristics. They provide for variable binding, which allows for accounts of goals and roles in perceived activity. They provide for embedding, which allows for an account of the partonomic structure of activity. They provide for varied levels of abstraction, which allows for taxonomic structure in activity. (For a comprehensive discussion of the psychological reality of hierarchical representations, see Cohen, 2000.) Finally, schemata represent knowledge rather than definitions, which allows for adaptive, probabilistic perception engaging a fluid interplay of bottom-up and top-down processing. These features give event schemata considerable power to describe phenomena in attention, cognition, memory, and action.

Event schemata provide a framework for on-line effortful cognition about ongoing activity as well as attentional deployment. By embedding, they can capture the partonomic structure of events—the way many events consist of discrete parts, and those parts in turn consist of parts. They relate actors' goals to this part structure by binding variables. Most of the time, people do things with goals in mind. Interpreting action in terms of goals may also be a productive strategy for interacting with animals and even non-living things (Dennett, 1987). This tracking of relations between goals and temporal structure is a valuable feature of an event representation

According to schema theory, recognizing an event as an instance of a category consists of matching it to a schema stored in memory. Understanding what is going on consists of matching features of the perceptual world to variables in the schema. In ongoing perception, missing information is filled in by reference to the patterns of intercorrelation captured by the schema, leading to a fluid interplay of bottom-up and top-down processing.

Text Comprehension and Memory

Event schemata figure prominently in several theories of narrative comprehension, particularly those based on scripts, story grammars, or situation models (e.g., Rumelhart, 1977; Schank & Abelson, 1977; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Scripts, story grammars, and situation models are particular versions of event schemata; all are characterized by partonomic structure. This has been a rich area for research, and is relevant to event structure perception because the same knowledge structures that are important for understanding stories about events probably are important for perceiving real events.

One indication that partonomically organized event schemata play a role in story understanding comes from studies of comprehension time for questions about stories and for stories themselves. In one study, Foss and Bower (1986) had participants read about a procedure (for joining a club) that required the satisfaction of a number of hierarchically organized goals. They found that the time it took to answer true-false questions about the possible order of events in the situation depended on the distance of the events in the hierarchical structure: Events that were farther apart and had more branches between them took longer to verify. However, people apparently do not always use part structure in searching memory for texts. In a follow-up study, Franklin and Bower (1988) found that, though participants were able to produce the implicit hierarchical structure on request, their time to answer true-false questions was predicted simply by the linear distance of the events in time—and questions about events that were far apart in time were answered more quickly (an inverse distance effect). The critical difference between the two studies seems to be the nature of the testing: Franklin and Bower tested participants repeatedly on the same material. It may be that with repeated testing, participants recoded the material as a simple list of actions, ordered in time, disregarding their goal relationships.

Another piece of evidence, though a weak one, is the lack of a graded distance effect in story reading time. If a story were represented solely in terms of linear time or local relations, one would expect that when two adjacent sentences describe temporally distant sub-events, reading time would be longer than for temporally proximal sub-events. Within simple stories in which all the statements form natural sub-parts of one superordinate part, such graded distance effects have difficult to generate (Abelson, 1981; Bower et al., 1979). If the partonomic relations between events and sub-events are directly represented, this may override the local or linear relations, explaining this result.

Event schemata have also been implicated in memory for stories. Rumelhart and Ortony (1977) identify two ways in which schemata can influence memory. First, the initial encoding of an event can be distorted by the schema(ta) that are activated at the time. Second, recall is presumed to be a reconstructive process that draws on schematic structure as well as partial traces from the original event. Evidence can be found pointing to both processes, though they are often difficult to disentangle. In one experiment, Rumelhart (1977) presented readers with stories that had been analyzed in terms of partonomic schemata. After a delay, they recalled the stories and their responses were coded in terms of the same organization. Delayed recall was characterized by a pruning of the hierarchy such that fine-grained events were collapsed into larger units, suggesting that the initial fine-grained information had been encode in terms of its relationship to larger-grained events.

When events are encoded in terms of a schema, this structure can influence later reconstructive memory. In one experiment, participants were given a test of recognition memory after reading a script-based story. When lures were chosen to be plausible parts of the script, false recognition rates were high. In a following experiment, event parts were presented in the text in an atypical order; during a later memory test there was a tendency to falsely remember them in the typical order (Bower et al., 1979). In a similar study, such inference-based false alarms were found to be asymmetric. Participants read stories that explicitly named either actions (parts) or scenes (wholes). Later, they were more likely to falsely recognize sentences that named scenes based on inferences from actions that had been presented than sentences that named actions based on scenes that had been presented, though both types of lure were plausible inferences (Abbott et al., 1985).

Similar results obtain for memory of videotapes, a stimulus type that is closer to "live" events than stories. Recall of videotapes of human activity is characterized by a hierarchical pattern of recall. Memory for actions that are relevant to the event schema is better than memory for schema-irrelevant actions. As with memory for texts, the order of sub-events tends to revert with time to the schema-normal order (Lichtenstein & Brewer, 1980). Activation of a schema can lead to false recognition of actions implied by that schema. Further, the same action is better recognized and better recalled when it is part of an activated event schema than when it is not, and recall for details within an event segment tends to be all-or-none (Brewer & Dupree, 1983).

Structural features of a videotape that highlight the hierarchical structure of the activity portrayed have been shown to improve memory for the activity. Boltz (1992a; 1995) showed participants videotapes of a spy drama. Commercials were placed either at "breakpoints" between major idea units (as rated by a group of observers) or at control locations. Viewers who saw the version with commercials placed at breakpoints had superior recall and recognition memory for the story content and better recognition memory for the temporal order of the scenes (Boltz, 1992a). They also had better memory for the duration of the stimulus (Boltz, 1995). These results indicate that placing the commercial breaks to correspond with the structure in the activity portrayed made it easier for viewers to encode the activity accurately.

Hierarchical patterns of recall, and influences of goals on memory for activity, have been found for stories in young children (Hudson, 1988; Trabasso & Stein, 1994; van den Broek, Lorch, & Thurlow, 1996), and for simple events in infants as young as 15 months (Bauer & Mandler, 1989; Travis, 1997).

This research on event schemata in text comprehension and memory is closely tied to work on situation models (e.g., Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). A situation model is a representation of the activity described by a discourse, which captures some of the qualitative (and sometimes quantitative) features of the situation described. The model can then drive bridging inferences and allow for problem solving in a more psychologically plausible fashion than propositional inference. Situation model theorists have analyzed text comprehension as the on-line construction of a situation model while processing a text. A given situation model is an instantiation of a schema, that is, it is a token while the schema is a type (Zwaan & Radvansky, 1998).

Situation models, like event schemata and scripts, include partonomic relationships. In comprehending a text, a reader extracts higher-level structural information about the activity, and this "macrostructure" has a recursive organization. In complex texts, the macrostructure may be explicitly reified in the text by features such as chapter headings and subheadings (van Dijk & Kintsch, 1983). Thus, models derived from text comprehension necessarily conflate information about the structure of the situation with information about the structure of the text qua text, because authors tend to build in redundancy between the two to assist readers.

The situation-model account that is most strongly tied to a theory of event structure is the event-indexing model (Radvansky, Zwaan, Federico, & Franklin, 1998; Zwaan, Langston, & Graesser, 1995). In this model, events are individuated based on five features: time, space, intentionality, causality, and protagonist. Temporal structure in particular seems to play a powerful role in organizing memory retrieval, and is especially useful in narrative comprehension because it is marked by every sentence in the text (Talmy, 1988).

Goals and Structured Representations

Why do we do things? The very fact that this question can be asked and answered in specific instances reveals that much activity is goal-directed. Theorists have suggested that the goal-directed nature of activity is exploited by the cognitive system in understanding the actions of others (Abelson, 1981; Bower, 1982; Bower et al., 1979; Rumelhart, 1977) and of ourselves (Vallacher & Wegner, 1987). Two characteristics of goal-directed activity are particularly relevant. First, goals can be recursively decomposed into subgoals, generating a partonomy. Second, goal-directed activity exhibits recurrent intercorrelated patterns of activities which give rise to invariances in the perceptual stream (Heider & Simmel, 1944). These can be captured by schemata (Rumelhart, 1980; Rumelhart & Ortony, 1977), frames (Minsky, 1972), or scripts (Schank & Abelson, 1977). In what follows, we will refer to goal-directed hierarchical knowledge structures for representing events as "event schemata."

Goal decomposition gives event schemata the characteristics of partonomic hierarchies. If an observer (call her "Olivia") witnesses a person wearing a mask and carrying a gun enter a bank, chances are she will infer a higher-level goal: to rob the bank. She may infer a still higher-level goal: to acquire money. This default inference can be over-ridden (e.g. if Olivia then learns the robber is a performance artist who has a history of breaking into banks and taking only the rubber bands). The goal-orientedness of event schemata allows prediction of future activity and inference of missing information. For example, Olivia can predict that the robber will head for the safe. Upon exiting with a bulging sack, she can infer the sack's contents (money—or perhaps rubber bands). Thus, goal-directed event schemata are probably important for comprehending activity.

The goal-directed nature of behavior formed the basis of Newell and Simon's (1972) General Problem Solver, a pioneering computer model of problem-solving for planning. In the model, a high-level goal is recursively broken down into lower-level goals until the outstanding lower-level goals can be satisfied. For example, "getting to Boston" can be decomposed into "getting to the airport," plus "getting on a plane." "Getting to the plane" can in turn be decomposed into "getting to the bus stop" plus "getting on a bus." This decomposition continues until the outstanding goals are represented as primitives in the computer program. Rumelhart and Ortony (1977) postulate behavioral primitives that play a similar role in human behavior.

The relationship between sub-goals and sub-parts is illustrated nicely by Barker and Wright's example shown in Figure 1. Here, the recursive part structure mirrors an implicit goal structure. "Stepping down from the curb" is not only a part of "crossing street" but is also a sub-goal.

This analysis provides a complement to the philosophical analysis of basic actions, and with the linguistic analyses of motion situations and of verb aspect. As we saw earlier, philosophical treatments of action give a positive characterization of the building blocks of activity in terms of basic actions, and linguistic treatments of motion situations provide a definition for atomic events. Theories of schema-based perception, action, and memory define the recursive operation of part decomposition, leading to a hierarchical event structure, but are largely silent as to the nature of the terminals in the hierarchy. They should be the smallest events that are encoded as such. Basic actions (for planning) and atomic events (for perception and memory) make good candidates for these terminals.

Event schemata bring together temporal (partonomic) structure and goal information. The event segmentation data described previously support the view that observers actively employ such structures as they watch activity. Recall that event segmentation has been observed to be more hierarchical for familiar activities, and more hierarchical when observers described the activity while segmenting it (Zacks et al., in press, see "Event Segmentation" above). Familiarity indexes the degree to which one has a schema for an activity, which explains the first effect. The second effect suggests that describing the activity primes the relevant schema, because the schema provides the inferential and predictive power necessary for composing a description.

There is also direct support for the importance of goals in event perception. In one study, Thibadeau (1986) examined the perceptual segmentation of a simple animated film by a large number of observers. He coded frames in the film based on two properties: which character(s) in the film for which the frame showed a satisfied goal, and the degree of intentionality reflected in the action of the frame. Both properties were predictive of the likelihood that observers would perceptual segment the film at a given location. Baldwin and Baird (1999) have argued that infants possess the foundations of a system for segmenting activity by its intentional structure. By identifying characteristic recurring patterns of physical activity that are diagnostic of actors' intentions, infants are able to parse activity into meaningful structures. In one experiment, Baird and her colleagues presented infants with a videotaped sequence for which intentionally relevant boundaries (breakpoints) had been previously identified. They inserted pauses into the videotape either at these intentional breakpoints or at appropriate control intervals that interrupted the intentional structure of the activity. They found that infants looked longer at the intention-interrupting sequences than at the sequences with pauses at the breakpoints (Baird, Saylor, & Baldwin, 1999).

Action and Perception

There is some evidence that the same schema-based processes that drive perception of others' activity influence our perception of our own actions. This issue has been addressed extensively by Vallacher and Wegner (1987). As noted previously, perception of one's own activity shows a hierarchical structure. Higher levels in this structure tend to correspond to longer-term goals and to maintenance of one's self-concept, while lower-level goals address the specifics of how an action is performed. The lower the level of encoding, the more susceptible to reinterpretation an action is, because a low-level action can occur in the context of multiple higher-level goals (i.e. multiple event schemata). Vallacher and Wegner argue for three principles governing how one understands one's own acts. First, acts tend to be maintained with respect to their prepotent identity. Second, when both a lower and higher act identity are available, the higher level tends to become prepotent. Third, when an action cannot be maintained in terms of its prepotent activity, a lower level identity tends to become prepotent (Vallacher & Wegner, 1987, p. 4-5). For example, imagine Milo's friend Mary stops him in outside his office and asks him to follow her. Milo agrees. Once they start down the hall, the first principle predicts he is likely to continue to represent his activity as "walking with Mary." As they walk, Mary explains that she's locked herself out of her office and would like him to open her door. Now, according to the second principle, a new, higher act identify can become prepotent: "helping Mary open her door." Now suppose they are repeatedly interrupted, or an obstacle prevents them from reaching Mary's office. According to the third principle, action identification would revert to a lower-level description.

The hysteresis provided by the first principle, and the tendency toward abstraction provided by the second, can lead to distortions both in current self-perception and (especially) in memory for what we've done and why we did it. On the other hand, under challenging circumstances, when self-perception tends to shift to a lower level of action identification, the resulting fine-grained coding may impair one's ability to relate what one is doing to larger-scale goals and self-concepts. For example, in one experiment forcing participants to describe their actions at a fine grain made them more susceptible to bogus feedback about a coarser grain, namely their cooperativeness in the experiment (Wegner, Vallacher, Kiersted, & Dizadji, 1986).

As well as influencing people's perception of their own actions, event schemata can undoubtedly guide those actions. One way this may happen is through variable binding in a schema: Imagine being pulled to the side of the road while driving late at night. Your behavior toward the person responsible will presumably depend on whether you instantiate a "traffic ticket" schema and assign this person the role of "police officer," or instantiate the "carjacking" schema and assign this person the role of "car thief." Another way schemata can influence action is through prediction and inference. Standing in the batter's box, a baseball player generally predicts there will be a pitch arriving shortly and behaves appropriately: crouching, eyeing the pitcher, etc..

The way an event schema drives action will depend on the level at which the current action is encoded. Changes in level of encoding are often adaptive. For example, if uncertainty or unpredictability causes a cyclist to switch from encoding her activity as "riding down Alpine Road" to "avoiding an upcoming stump," this is likely to improve her chances of safely navigating the obstacle. On the other hand, the pitfalls of concentrating on the wrong level are potentially serious. If a noise coming from the bicycle chain were to shift our cyclist's attention all the way to the level of "turning the pedal," this could lead to her "missing the big picture," perhaps with disastrous consequences. (See also Vallacher & Wegner, 1987.)

There is neurophysiological evidence for a tight relationship between the perception and performance of actions, at least on a fine temporal grain. Rizzolatti and his colleagues have shown that there are neurons in the premotor cortex of the monkey that fire both when a monkey performs a goal-directed action, and when the monkey observes another monkey (or human experimenter) perform the same action (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). The cells can be quite specific. For example, cells have been identified that respond only to a reaching gesture with a particular style of hand grip. One may speculate that there are neural representations of larger-scale events whose activity mediates both the perception and organized performance

Acquiring Schemata

In principle, the development of schemata can accommodate two sources of influence: evolutionary selection and learning. It seems likely that over evolutionary time some goals have remained salient, particularly the general goals that can also be characterized as drives: to eat, to procreate, to avoid harm. However, for much interesting behavior, the relevant goals and plans are likely to be almost entirely culturally transmitted. Thus, the most important legacy of the evolutionary forces shaping our cognitive architecture is probably a general mechanism or set of mechanisms for inferring and reasoning about causes and goals. As noted previously, this reasoning ability may have a perceptual analog in the form of a sensitivity to physical patterns that are diagnostic of intentional structure (Baldwin & Baird, 1999; Michotte, 1946/1963). Direct evidence for this view comes from a study in which 18-month-olds were given the opportunity to copy actions of an adult. When the adult attempted to achieve a goal but failed, the infants tended to produce the successful action rather than copy the failure. However, when shown similar sequences performed by an inanimate object, they tended to copy the literal motion sequence (Meltzoff, 1995). Further support comes from a recent studies by Woodward (1998), in which infants observed sequences in which an adult reached for an object. For infants as young as five months, changes to the object grasped were attended more than changes to the path of the grasping arm. In these sequences, the goal of the sequence is presumably to retrieve the object, and the path followed by the arm is incidental to that goal, so the result suggests that infants selectively attend to the goals of the sequences.

Some indication of how knowledge about event schemata is acquired comes from developmental studies of children's understanding of everyday events. A number of results have suggested that when children as young as three years old are queried about everyday activities, their responses reflect the influence of hierarchically-organized event schemata that include information about the goal structure of activity (Nelson, 1986; Slackman et al., 1986). Familiarity, age, and event salience all lead to more elaborated structures (Nelson & Gruendel, 1986). One study of recall for television stories provides a detailed picture of the features of activity that drive memory organization in development (van den Broek et al., 1996). In this experiment four-year-olds, six-year-olds, and adults were presented with clips from a children's television show. Events in the sequences had been coded in terms of four features: their number of causal connections, whether they belonged to a causal chain, their content category (based on a story grammar), and their position in the story's hierarchical goal structure. All four features predicted probability of recall for all three groups. Moreover, all four were correlated. Regression analyses indicated that the number of causal connections may drive the other effects, especially for the adults. Importantly, adults were more likely than children to describe the remembered activity in terms of goals and the events that initiated them.

Goals play a role in children's storytelling as well as in their memory for events. In a series of studies of narrations of picture-books, Trabasso and Stein and their colleagues found that as children developed they made increasing use of hierarchically organized plans in their narrations (Trabasso & Stein, 1994). From ages three to five there was an increase in children's propensity to identify characters and their relations. From three to nine there were large increases in the representation of goals and their relationships to actions and outcomes. By age nine, the narrations captured as much of the goal/plan structure as did adults' narrations.

Even for toddlers, hierarchically organized event representations have been shown to guide event perception and memory. In one study (Travis, 1997), two-year-olds were shown brief event sequences that contained two precursor actions and one of two possible goal actions example. A control group was shown both possible goals. When reproducing the actions, both groups tended to cluster actions by goal. The experimental group produced more goal-relevant actions. Both patterns suggested goal-based encoding. In a second experiment, 20-month-old and 15-month-old infants were presented with sequences that contained a given action as either a precursor of a goal, or as a sequent of that goal. For both groups, the identical action was more likely to be produced when a precursor than when a sequent. In another study (Bauer & Mandler, 1989), children as young as 16 months were elicited to imitate remembered event sequences. Sequences were simple action sequences such as cleaning a table with paper towel and spray cleaner, or making a picture on a chalkboard. Causal connections within the activities facilitated memory, even in the face of intervening interruptions with the sequences.

This excursion into structured event representations has a number of implications for cognition. Structured event schemata play a substantial role in cognition about events. They influence on-line processing of narrative texts as well as later memory for those texts. Memory for videotapes shows the influence of the same sorts of structures. Event schemata encode information about goals, and this structure in turn aligns with partonomic structure. Event schemata influence action as well as perception, by playing a role in planning and in how we think about our own past actions. While the precursors of event schemata seem to be present in toddlers, event representations increase in complexity into the teens. Thus, event schemata that relate goals and part structure can capture an impressive amount of information about peoples' conceptions of events.

Computational Models of Event Understanding

Much of the research we have reviewed was motivated in part by artificial intelligence models of story understanding and planning. These models have been widely analyzed (e.g., Schank, 1978; Schank & Abelson, 1977) and critiqued (e.g., Searle, 1980; Winograd & Flores, 1986). In this excursion, we will address their significance as heuristic devices for thinking about event structure perception.

Schank and Abelson's (1977) influential script theory of story understanding was implemented in a series of computer programs that were able to produce putatively human-like patterns in comprehension for texts. Two features of these models are especially relevant to event structure conception. First, the underlying formalism was based on Schank's (1975) conceptual dependency theory, which analyzed semantics in terms of a small number of primitive features. Notably, these basic components included 11 primitive action classes, called ACTs. These included physical actions (e.g., PROPEL, to impart of force to something), abstract actions (e.g., ATRANS, to abstractly transfer something, as in giving or selling), and mental actions (e.g., ATTEND, to direct attention to something). Second, the text comprehension models included structured representations of the temporal structure of particular event categories. These "scripts" constituted the strongest psychological claims of the models (Abelson, 1981). These features suggest that (a) decomposing events into primitive units and (b) maintaining structured representations of events are both important to the conceptual analysis of everyday activity.

SAM (for "script applier mechanism") was a set of programs that could parse text in terms of scripts for activities, using that structure to paraphrase, summarize, translate and answer questions about the story (Cullingford, 1981). FRUMP (for "fast reading understanding and memory program") used pared-down scripts to provide summaries of newspaper articles (Schank & Abelson, 1977). The program from this laboratory that most fully implemented the kind of schema theory described here was probably PAM (for "plan applier mechanism"). PAM explicitly represented goals and plans, inducing them from cues in a text, and interpreted statements in terms of those goals and plans (Wilensky, 1981).

Rumelhart's (1975; 1977) schema theory of story understanding was also implemented as a computer program that could parse and summarize folk tales. Understanding was taken to be the match of a sample of input text to a hierarchical propositional representation of a schema. Summarization was implemented by pruning the lower-level nodes of the hierarchy. Rumelhart showed that the automatically generated summaries were qualitatively similar to human-generated summaries, and also to human recall of stories from memory.

Event-based representations have also been applied to the automatic interpretation of simple motion sequences. The ABIGAIL system (Siskind, 1994) took sequences of line-drawings as its input and generates a propositional representation describing the activity. The system was based in part on an analysis of verbs that provides perceptual primitives for three classes of features of objects: support/contact relationships, motion categories, and part/sub-part relationships. The structure of the representational scheme explicitly reifies the temporal structure of the activity.

Newell and Simon's (1972) General Problem Solver, as described previously, was a computer model of planning as well as a psychological theory. More recently, Lansky (1994) has argued that actions and their relationships belong at the basis of AI planning systems. In previous systems such as GPS and its descendants, the basic representational framework refers to states of the world, and actions are represented indirectly, as transitions. In Lansky's COLLAGE planner, the atomic units in the representation are actions rather than states or objects. This amounts to an even stronger claim that the natural way to represent planning is in terms of events and their relations.

These results constitute a set of heuristically valuable existence proofs. They demonstrate that partonomically-organized event representations can support inference, summarization, question-answering, and planning in ways that are reminiscent of human behavior. Moreover, these models embody suggestions about how the mind might represent events, which have generated empirical research.

Constraints on a Theory of Event Structure Perception and Cognition

We began with a broad conception of event, a segment of time at a given location conceived as having a beginning and an end. We then examined the approaches of several disparate disciplines to this topic, notably, philosophy, psychology, computer science, and linguistics. These different approaches have drawn on different frameworks and observations, and, not surprisingly, have emphasized different features of events. At one extreme is the approach of the ecological psychologists, Barker and Wright (1954). They characterized the observable signs of episode shifts: change in sphere (verbal, social, intellectual), physical direction, object, tempo, or setting of behavior, or change in predominant part of body. These are all perceptible features of behavior in context. At the other extreme are the characterizations of those analyzing discourse which emphasize features of events not immediately observable, but perhaps inferable from observation, notably, causality and intentionality (e.g., Radvansky et al., 1998).

Despite the differences in approaches, there are striking convergences among the formulations based on perceptible features of behavior and those based on conceptual analyses of event structure. This is because people make use of multiple sources of information in perceiving and thinking about event structure, both bottom-up data-driven and top-down concept-driven sources. These include....

Covariation Between Multiple Sources of Information

It is clear that people make use of multiple sources of information in perceiving and thinking about event structure. These include partonomic relations, perceptual event boundaries, objective features of object and actor motion, perceptual causal properties, statistical patterns of occurrence, and goal relations. In the first section ("Lessons from Objects") we saw that people naturally talk about events in terms of part boundaries and part relationships. The second section ("Event Segmentation") showed how perceptually natural event boundaries can be identified in part based on objective physical features. Similarly, the causal properties of simple events can be perceptually identified and related to physical features. In the third section ("Characterizing Event Units") we described research indicating that infants and adults can learn event units based purely on the statistical pattern of their occurrence. In the fourth section ("Structured Representations of Events") we showed that people use information about goals and plans to reason about events.

Under normal circumstances all these sources of information are available—and covary. In the preceding sections we have identified a number of these correspondences directly. Partonomic relations correspond with perceptual structure: People spontaneously segment activity at different timescales in correspondence with a partonomic hierarchy (Zacks et al., in press). Partonomic relations also correspond with statistical patterns of occurrence, because sequences of activity recur in the environment (Avrahami & Kareev, 1994). Partonomic relations correspond with goal relationships because the acting out of a goal-subgoal decomposition generates a partonomic hierarchy. Perceptual event boundaries correspond with objective physical features of object and actor motion, because observers tend to segment activity at extrema in biological motion (Newtson et al., 1977). Finally, objective features of object motion correspond with perceptual causal properties (Michotte, 1946/1963).

From these direct relations we can derive a number of indirect relations. To name a few: Partonomic relations correspond with physical motion, because both correspond with perceptual segmentation. Perceptual segmentation corresponds with causal properties, because both correspond with physical motion. Physical motion corresponds with goal relations, because both correspond with partonomic relations. For example, think of a man packing boxes. Two successive parts of the event might be "placing contents into box" and "taping box closed." The boundary between those parts would also tend to be a point of distinctive physical motion (a change in posture, reaching, etc.). Observers would tend to segment the activity at this point in time. Finally, this point corresponds to the satisfaction of a goal (to fill the box). By tracing these indirect relationships it becomes clear that every source of information about event structure is mutually informative, to some degree, with every other source.

This dense pattern of covariation among information sources has strong implications for event perception and cognition. From a Bayesian perspective, it means that posterior probabilities of a feature in one domain (e.g., part relations) can be continuously updated based on the presence of features in several other domains (e.g., causality, goal relations, motion). From a connectionist perspective, it means that a model of these phenomena will likely include substantial connectivity between units representing each domain, to provide for interactive constraint satisfaction. In other words, the interdependence of information from different sources indicates that if different systems process each of these domains, event perception and cognition arise as an emergent phenomenon resulting from the interaction of all of the systems.

The requirement of real-time probability updating or constraint satisfaction rules out at least two classes of model of event perception: slot-filler models and some of the purely feed-forward models of ecological psychology. Neither can accommodate these mutual influences in real time.

Moreover, covariation across information systems argues against the view that event understanding can be explained solely in terms of one class of feature. Hypotheses that there is one fundamental domain to which the others reduce (e.g., intentionality, causal influence, or physical structure) are tempting for their parsimony. However, the fact that mutual information across domains appears to be pervasive militates against such reductions.

Representations That Span Timescales

From a range of sources comes evidence that people perceive and conceive of events in terms of representations that span time-scales. We described evidence from structured and unstructured descriptions of everyday activity (Barker & Wright, 1954; Bower et al., 1979; Slackman et al., 1986), perceptual segmentation (Zacks et al., in press), and memory (Abbott et al., 1985; Bower et al., 1979; Brewer & Dupree, 1983; Lichtenstein & Brewer, 1980; Rumelhart, 1977). These quite different approaches converge in suggesting that people are aware of relationships between events on small and large timescales. Moreover, they can use such relationships in perceiving, reasoning about, speaking about, and remembering events. This constitutes another form of mutual information: Characteristics of activity at a fine temporal grain constrain characteristics of activity at a coarse temporal grain, and vice versa.

What is important about cross-timescale relationships? Here we have emphasized hierarchical relations between parts and sub-parts. Partonomic hierarchies have attractive properties as tools for theory-building: simplicity, ease of description and visualization, and straightforward relationships to well-understood search algorithms. Moreover, several bodies of research show that analyzing events in terms of partonomic hierarchies gives reasonable fits to human performance. However, it seems unlikely strict hierarchies can give a complete account of event perception and cognition. It is all too easy to construct examples for which partonomic hierarchies are insufficient. To illustrate this, we return to the analogy between objects and events.

For objects, strictly hierarchical sub-part decompositions are often problematic. Constructing a partonomy for a river is a challenge because the boundaries of the object and its parts are not well-defined. Constructing a partonomy for a city is problematic because there are multiple reasonable spatial decompositions (by geographical boundaries, political subdivisions, sociological features). Thus, object part relations may be underdetermined or overdetermined—or both. The issues for events are at least as substantial.

Constructing a partonomy for a basketball game is a challenge because there are multiple actors with partially overlapping physical movements, goals, and causal interactions. Constructing a partonomy for preparing dinner is problematic because a person may interleave components of multiple sub-parts (e.g., interrupting cutting vegetables to stir the soup). Multiple actors, multiple goals, uncertainty in any source of event part information, and variations in observer expectation and experience can all induce complexities into a partonomic analysis of an event.

For both objects and events, the lesson to be drawn is <u>not</u> that part-sub-part relations are unimportant or that partonomies are psychologically implausible. Rather, it is that in real situations partonomy is a local and imperfect phenomenon. A reasonable hypothesis is that the same holds for psychological reflections of those situations. The fact that partonomic analyses provide good accounts of behavior does not establish that the mind or the brain implement strict hierarchical constraints. Conversely, the fact that some events are clearly not well-described by simple hierarchies does not imply that perceivers and conceivers fail to extract hierarchical structure where it is available.

Interactions Between Evolution and Development

Any comprehensive theory of event structure perception and conception will need to account for interactions between the evolutionary history of the human organism and the personal developmental history of individuals. The available evidence indicates that these interactions are pervasive and richly structured. Event understanding relies on the interaction of mechanisms that appear early and late in development, that are domain-general and domain-specific, and that experience-dependent and experience-independent.

The fact that aspects of event perception correspond to objective physical features of physical motion (Michotte, 1946/1963; Newtson et al., 1977), which have changed little if at all over geologic time, suggests that these features could be shaped largely by natural selection. The plausibility of this hypothesis is strengthened by the finding that infants can identify features of event structure early in development (Baldwin, Baird, Saylor, & Clark, submitted; Meltzoff, 1995; Woodward, 1998; Woodward & Sommerville, 2000; Wynn, 1996; Wynn & Chiang, 1998).

However, there are also indications that experience importantly shapes event perception and conception. To the extent that schemata for events condition understanding, the cognitive system is constantly "turn[ing] round upon its own 'schemata' and to construct them afresh" (Bartlett, 1932, p. 206). Effects of familiarity on event perception indicate that even basic perceptual processes are subject to shaping by experience (Graziano et al., 1988; Zacks et al., in press). Effects of familiarity and expertise in development show that experience exerts pervasive influence on event conception (Bauer & Mandler, 1989; Nelson & Gruendel, 1986).

Thus, while some aspects of event perception and conception show effects of shaping by evolution, others clearly show that these effects interact with effects of an individual's experience. A complete theory will need to explain how effects of natural selections and of learning interact to produce behavior in the mature organism (see also Baldwin & Baird, 1999).

Perspectives on Events

Events are the stuff of our lives. When we tell about our lives, we relate a series of events. What did we do today? Got up, took a run, showered, got dressed, grabbed some food, went to work. What is the stuff of events? We have considered many takes on events, from comic book artists to linguists, from recorders of minute aspects of behavior to designers of architectures of the mind, from babies to philosophers. All see events as occurring in time and space, as having beginnings and ends. Some add other attributes: physical features, causal attributes, statistical relationships, goals, plans, and intentions.

Time and space provide contrasting perspectives on events. A temporal perspective highlights the sequence of transitions, the dynamic changes from segment to segment, things in motion. A spatial perspective highlights the sequence of states, the static spatial configuration, things caught still. Capturing the temporal and the spatial at once seems elusive; like waves and particles, the dynamic and the static appear to complement each other.

References

Abbott, V., Black, J. H., & Smith, E. E. (1985). The representation of scripts in memory. Journal of Memory and Language, 24, 179-199.

Abelson, R. P. (1981). Psychological status of the script concept. American Psychologist, 36, 715-729.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. Psychological Science, 9, 321-324.

Avrahami, J., & Kareev, Y. (1994). The emergence of events. Cognition, 53, 239-261.

Baird, J. A., Saylor, M. M., & Baldwin, D. A. (1999, April). Action parsing in infancy and the origins of intentional understanding. Paper presented at the Society for Research in Child Development, Albuquerque, New Mexico.

Baldwin, D. A., & Baird, J. A. (1999). Action analysis: A gateway to intentional inference. In P. Rochat (Ed.), Early social cognition (pp. 215-240). Hillsdale, NJ: Lawrence Erlbaum Associates.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (submitted). Infants detect structure in human action: A first step toward understanding others' intentions? .

Barker, R. G. (1963). The stream of behavior as an empirical problem. In R. G. Barker (Ed.), The stream of behavior (pp. 1-22). New York: Appleton-Century-Crofts.

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

Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology. New York: The Macmillan Company.

Bauer, P. J., & Mandler, J. M. (1989). One thing follows another: Effects of temporal structure on 1- to 2-year-olds' recall of events. Developmental Psychology, 25, 197-206.

Bennett, J. (1996). What events are. In R. Casati & A. C. Varzi (Eds.), Events (pp. 137-151). Aldershot, England: Dartmouth.

Biederman, I. (1985). Human image understanding: Recent research and a theory. Computer Vision, graphics, and Image Processing, 32, 29-73.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. Psychological Review, 94, 115-117.

Boltz, M. (1992a). Temporal accent structure and the remembering of filmed narratives. Journal of Experimental Psychology: Human Perception & Performance, 18, 90-105.

Boltz, M. G. (1992b). The remembering of auditory event durations. Journal of Experimental Psychology-Learning Memory and Cognition, 18, 938-956.

Boltz, M. G. (1995). Effects of event structure on retrospective duration judgments. Perception & Psychophysics, 57, 1080-1096.

Bower, G. (1982). Plans and goals in understanding episodes. In A. Flammer & W. Kintsch (Eds.), Discourse Processing (pp. 2-15). Amsterdam: North-Holland Publishing Company.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. Cognitive Psychology, 11, 177-220.

Brent, M. R. (1999). An efficient, probabilistically sound algorithm for segmentation and word discovery. Machine Learning, 34, 71-105.

Brewer, W. F., & Dupree, D. A. (1983). Use of plan schemata in the recall and recognition of goal-directed actions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 117-129.

Bülthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are three-dimensional objects represented in the brain? Cerebral Cortex, 5, 247-260.

Byrne, R. W. (1999). Imitation without intentionality: Using string parsing to copy the organization of behaviour. Animal Cognition, 2, 63-72.

Casati, R., & Varzi, A. C. (1996). Events. Aldershot, England ; Brookfield, Vt.: Dartmouth.

Cohen, C. E., & Ebbesen, E. B. (1979). Observational goals and schema activation: a theoretical framework for behavior perception. Journal of Experimental Social Psychology, 15, 305-329.

Cohen, G. (2000). Hierarchical models in cognition: Do they have psychological reality? European Journal of Cognitive Psychology, 12, 1-36.

Cullingford, R. (1981). SAM. In R. C. Schank & C. K. Riesbeck (Eds.), Inside computer understanding: five programs plus miniatures (pp. 75-119). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances: an artificial- language study with implications for native-language acquisition. Journal of Experimental Psychology: General, 128, 165-185.

Danto, A. (1963). What we can do. Journal of Philosophy, 60, 435-445.

Davidson, D. (1970/1996a). Events as particulars. In R. Casati & A. C. Varzi (Eds.), Events (pp. 99-107). Aldershot, England: Dartmouth (reprinted from *Nous*, 4, pp 25-32).

Davidson, D. (1970/1996b). The individuation of events. In R. Casati & A. C. Varzi (Eds.), Events (pp. 286-284). Aldershot, England: Dartmouth (reprinted from in *Essays in Honor of Carl G. Hempel*, by N. Rescher (ed.), Dordrecht: D. Reiderl, pp. 216-234.).

Davidson, D. (1966/1996). The logical form of action sentences. In R. Casati & A. C. Varzi (Eds.), Events (pp. 3-18). Aldershot, England: Dartmouth (reprinted from *The Logic of Decision and Action*, by N. Rescher (ed.), Pittsburgh: University of Pittsburgh Press, pp. 81-95).

Deerwester, S., Dumais, S., Landauer, T., Furnas, G., & Beck, L. (1988). Improving Information-Retrieval With Latent Semantic Indexing. Proceedings of the Asis Annual Meeting, 25, 36-40.

Dennett, D. C. (1987). The intentional stance. Cambridge, Mass.: MIT Press.

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

Ebbesen, E. B. (1980). Cognitive processes in understanding ongoing behavior. In R. Hastie (Ed.), Person memory: the cognitive basis of social perception (pp. 179-225). Hillsdale, NJ: Lawrence Erlbaum Associates.

Foss, C. L., & Bower, G. H. (1986). Understanding actions in relation to goals. In N. E. Sharkey (Ed.), Advances in cognitive science (Vol. I, pp. 94-124). Chichester: Ellis Horwood, Ltd.

Franklin, N., & Bower, G. H. (1988). Retrieving actions from goal hierarchies. Bulletin of the Psychonomic Society, 26, 15-18.

Goldman, A. I. (1970). A theory of human action. Englewood Cliffs, N.J.,: Prentice-Hall.

Graziano, W. G., Moore, J. S., & Collins, J. E. (1988). Social cognition as segmentation of the stream of behavior. Developmental Psychology, 24, 568-573.

Hacker, P. M. S. (1982/1996). Events, ontology and grammar. In R. Casati & A. C. Varzi (Eds.), Events (pp. 79-88). Aldershot, England: Dartmouth (reprinted from *Philosophy*, 57, pp. 477-86).

Hanson, C., & Hanson, S. J. (1996). Development of schemata during event parsing: Neisser's perceptual cycle as a recurrent connectionist network. Journal of Cognitive Neuroscience, 8, 119-134.

Hanson, C., & Hirst, W. (1989). On the representation of events: A study of orientation, recall, and recognition. Journal of Experimental Psychology: General, 118, 136-147.

Hanson, C., & Hirst, W. (1991). Recognizing differences in recognition tasks: A reply to Lassiter and Slaw. Journal of Experimental Psychology: General, 120, 211-212.

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

Hemeren, P. E. (1996). Frequency, ordinal position and semantic distance as measures of cross-cultural stability and hierarchies for action verbs. Acta Psychologica, 91, 39-66.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. Cognition, 18, 65-96.

Horgan, T. (1978/1996). The case against events. In R. Casati & A. C. Varzi (Eds.), Events (pp. 243-262). Aldershot, England: Dartmouth (reprinted from *Philosophical Review*, 87, pp. 28-47).

Hudson, J. A. (1988). Children's memory for atypical actions in script-based stories: Evidence for a disruption effect. Journal of Experimental Child Psychology, 46, 159-173.

Johnson-Laird, P. N. (1983). Mental models : towards a cognitive science of language, inference, and consciousness. Cambridge, Mass.: Harvard University Press.

Kim, J. (1975/1996). Events as property exemplifications. In R. Casati & A. C. Varzi (Eds.), Events (pp. 117-136). Aldershot, England: Dartmouth (reprinted from *Action Theory*, by Myles Brand and Douglas Walton (eds.), Dordrecht: D. Reidel, pp. 159-77).

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. Chicago: University of Chicago Press.

Landauer, T. K. (1998). Learning and representing verbal meaning: The latent semantic analysis theory. Current Directions in Psychological Science, 7, 161-164.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, 104, 211-240.

Lansky, A. L. (1994). Action-based planning. In K. J. Hammond (Ed.), Proceedings of the Second Annual Conference on Artificial Intelligence Planning Systems (pp. 110-115). Chicago: AAAI Press.

Lassiter, G. D., & Slaw, R. D. (1991). The unitization and memory of events. Journal of Experimental Psychology: General, 120, 80-82.

Lassiter, G. D., Stone, J. I., & Rogers, S. L. (1988). Memorial consequences of variation in behavior perception. Journal of Experimental Social Psychology, 24, 222-239.

Lichtenstein, E. D., & Brewer, W. F. (1980). Memory for goal-directed events. Cognitive Psychology, 12, 412-445.

McCloud, S. (1993). Understanding comics: The invisible art. Northampton, MA: Kitchen Sink Press.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. Developmental Psychology, 31, 838-850.

Michotte, A. E. (1946/1963). The perception of causality (T. R. Miles

Elaine Miles, Trans.). New York: Basic Books.

Miller, G. A., & Johnson-Laird, P. N. (1976). Language and perception. Cambridge, MA: Harvard University Press.

Minsky, M. (1972). A framework for representing knowledge. In P. H. Winston (Ed.), The psychology of computer vision . New York: McGraw-Hill.

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

Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. Memory & Cognition, 18, 407-418.

Narayanan, S. (1997). Talking the talk is like walking the walk: A computational model of verbal aspect, 19th Annual Meeting of the Cognitive Science Society (pp. 548-553). Stanford, California: Ablex.

Nelson, K. (1986). Event knowledge and cognitive development. In K. Nelson (Ed.), Event knowledge: Structure and function in development (pp. 1-19). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nelson, K., & Gruendel, J. (1986). Children's scripts. In K. Nelson (Ed.), Event knowledge: Structure and function in development (pp. 21-46). Hillsdale, NJ: Lawrence Erlbaum Associates.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, N.J.,: Prentice-Hall.

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

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. Journal of Experimental Social Psychology, 12, 436-450.

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

Newtson, D., Hairfield, J., Bloomingdale, J., & Cutino, S. (1987). The structure of action and interaction. Special Issue: Cognition and action. Social Cognition, 5, 191-237.

Parish, D. H., Sperling, G., & Landy, M. S. (1990). Intelligent temporal subsampling of American Sign Language using event boundaries. Journal of Experimental Psychology: Human Perception & Performance, 16, 282-294.

Pustejovsky, J. (1991). The syntax of event structure. Special Issue: Lexical and conceptual semantics. Cognition, 41, 47-81.

Quine, W. V. (1985/1996). Events and reification. In R. Casati & A. C. Varzi (Eds.), Events (pp. 107-116). Aldershot, England: Dartmouth (reprinted from *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, by Ernest LePore and Brian P. McLaughlin (eds.), Oxford: Blackwell. pp. 162-71).

Radvansky, G. A., Zwaan, R. A., Federico, T., & Franklin, N. (1998). Retrieval from temporally organized situation models. Journal of Experimental Psychology: Learning, Memory, & Cognition, 24, 1224-1237.

Rifkin, A. (1985). Evidence for a basic level in event taxonomies. Memory & Cognition, 13, 538-556.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. Cognitive Brain Research, 3, 131-141.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), Cognition and categorization (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosch, E., & et al. (1976). Basic objects in natural categories. Cognitive Psychology, 8, 382-439.

Rui, Y., & Anandan, P. (2000, June 13-15). Segmenting visual actions based on spatio-temporal motion patterns. Paper presented at the IEEE

Computer Society Conference on Computer Vision and Pattern Recognition, Hilton Head, SC.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), Representation and understanding; studies in cognitive science. (pp. 211-236). New York: Academic Press.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. Laberge & S. J. Samuels (Eds.), Basic processes in reading: Perception and comprehension (pp. 265-303). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education (pp. 33-58). Hillsdale, N.J.: L. Erlbaum Associates.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge (pp. 97-135). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old infants. Science, 274, 1926-1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. Journal of Memory & Language, 35, 606-621.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & et al. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. Psychological Science, 8, 101-105.

Schank, R. C. (1975). Conceptual information processing. Amsterdam: North-Holland.

Schank, R. C. (1978). Computer understanding of natural language. Behavior Research Methods, Instruments & Computers, 10, 132-138.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals, and understanding: an inquiry into human knowledge structures. Hillsdale, N.J.: L. Erlbaum Associates.

Searle, J. R. (1980). Minds, brains, and programs. Behavioral & Brain Sciences, 3, 417-457.

Siskind, J. M. (1994). Grounding language in perception. Artificial Intelligence Review, 8, 371-391.

Slackman, E. A., Hudson, J. A., & Fivush, R. (1986). Actions, actors, links, and goals: The structure of children's event representations. In K. Nelson (Ed.), Event knowledge: Structure and function in development (pp. 47-69). Hillsdale, NJ: Lawrence Erlbaum Associates.

Talmy, L. (1975). Semantics and syntax of motion. In J. P. Kimball (Ed.), Syntax and Semantics (Vol. 4, pp. 181-238). New York: Academic Press, Inc.

Talmy, L. (1988). Force dynamics in language and cognition. Cognitive Science, 12, 49-100.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. Psychonomic Bulletin & Review, 2, 55-82.

Thibadeau, R. (1986). Artificial perception of actions. Cognitive Science, 10, 117-149.

Trabasso, T., & Stein, N. L. (1994). Using goal-plan knowledge to merge the past with the present and the future in narrating events on line. In M. Haith (Ed.), The Development of future-oriented processes (pp. 323-349). Chicago: University of Chicago Press.

Travis, L. L. (1997). Goal-based organization of event memory in toddlers. In P. W. van den Broek, P. J. Bauer, & T. Bovig (Eds.), Developmental spans in event comprehension and representation: Bridging fictional and actual events (pp. 111-138). Mahwah, NJ: Lawrence Erlbaum Associates.

Tversky, B. (1990). Where partonomies and taxonomies meet. In S. L. Tsohatzidis (Ed.), Meanings and prototypes: studies in linguistic categorization (pp. 334-344). London: Routledge.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. Journal of Experimental Psychology: General, 113, 169-193.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. Psychological Review, 94, 3-15.

van den Broek, P., Lorch, E. P., & Thurlow, R. (1996). Children's and adults' memory for television stories: The role of causal factors, story-grammar categories, and hierarchical level. Child Development, 67, 3010-3028.

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

Wegner, D. M., Vallacher, R. R., Kiersted, G. W., & Dizadji, D. (1986). Action identification in the emergence of social behavior. Social Cognition, 4, 18-38.

Wilder, D. A. (1978a). Effect or predictability on units of perception and attribution. Personality and Social Psychology Bulletin, 4, 281-284.

Wilder, D. A. (1978b). Predictability of behaviors, goals, and unit of perception. Personality and Social Psychology Bulletin, 4, 604-607.

Wilensky, R. (1981). PAM. In R. C. Schank & C. K. Riesbeck (Eds.), Inside computer understanding: Five programs plus miniatures (pp. 136-179). Hillsdale, NJ: Lawrence Erlbaum Associates.

Winograd, T., & Flores, F. (1986). Understanding computers and cognition. Norwood, NJ: Ablex Corporation.

Wolff, P., & Gentner, D. (1995). What language might tell us about the perception of cause. In J. D. Moore & J. F. Lehman (Eds.), Proceedings of the seventeenth annual conference of the Cognitive Science Society . Pittsburgh: Lawrence Erlbaum Associates.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. Cognition, 69, 1-34.

Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month infants interpret action in context. Psychological Science, 11, 73-77.

Wynn, K. (1996). Infants' individuation and enumeration of actions. Psychological Science, 7, 164-169.

Wynn, K., & Chiang, W. C. (1998). Limits to infants' knowledge of objects: The case of magical appearance. Psychological Science, 9, 448-455.

Zacks, J., Tversky, B., & Iyer, G. (in press). Perceiving, remembering, and communicating structure in events. Journal of Experimental Psychology: General.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. Psychological Science, 6, 292-297.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. Psychological Bulletin, 123, 162-185.

Author Note

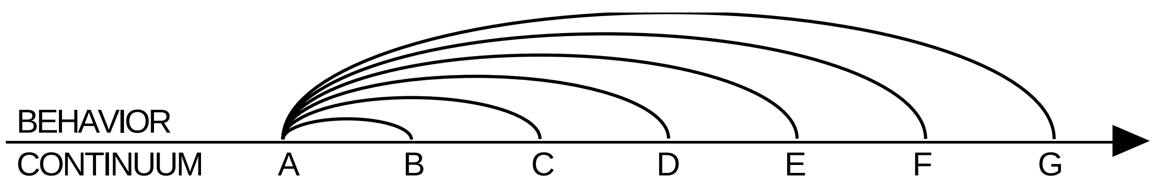
This work was supported by a National Science Foundation Graduate Fellowship and the Stanford University Humanities and Sciences Dissertation Fellowship. We would like to thank Gordon Bower, Herb Clark, John Gabrieli and Denis Pelli for thoughtful discussion.

Correspondence concerning this article should be addressed to: Jeffrey M. Zacks, Washington University, Department of Psychology, St. Louis, MO 63130-4899 (jzacks@artsci.wustl.edu).

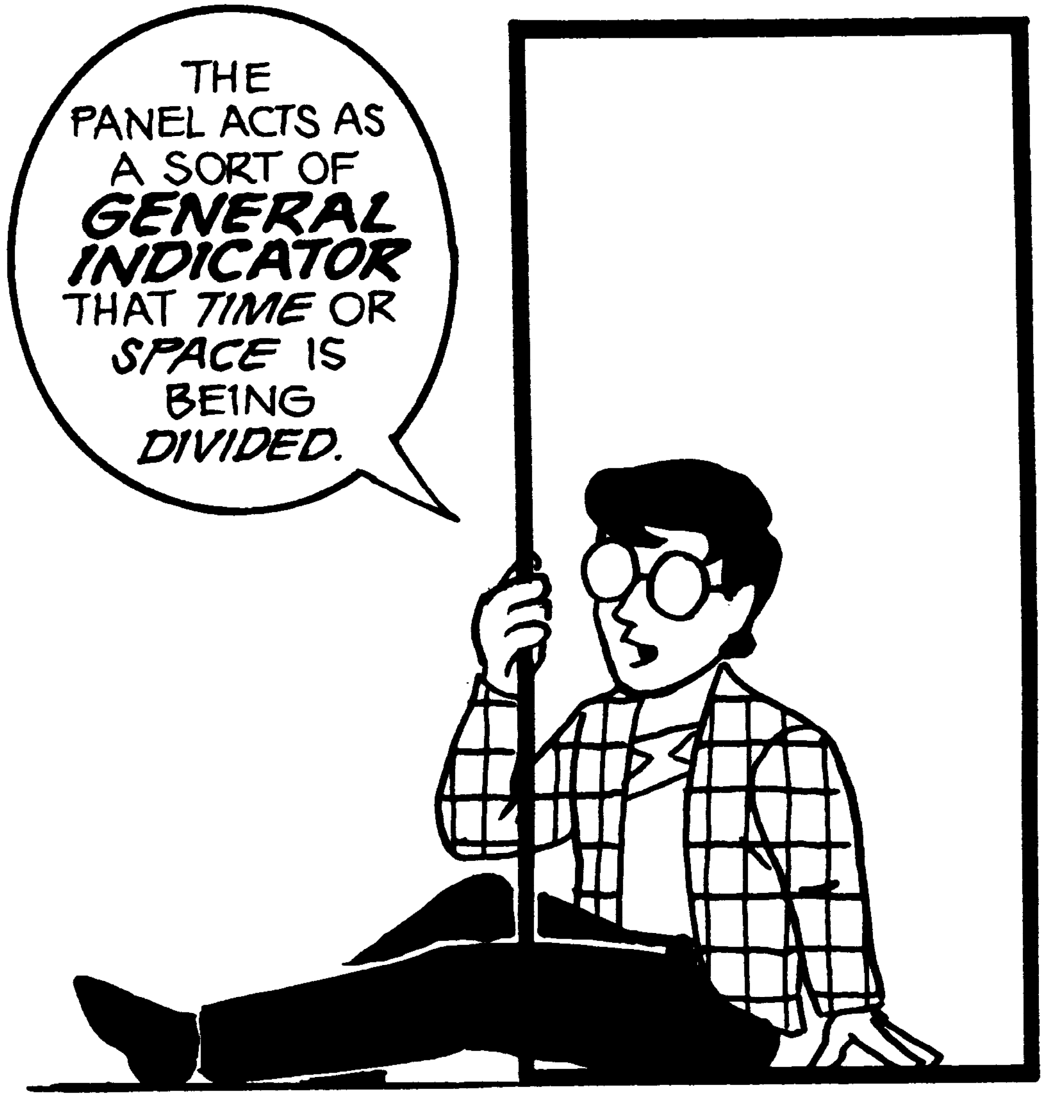
Figure 1: Behavior on different time scales (Barker & Wright, 1954).

Figure 2: Because comics allow their creator to represent time with space, they provide a mechanism to control how events are segmented (McCloud, 1993).

Figure 3: Schematic diagram of the launching effect. Initially, there is one object on the left of the screen and another in the middle. The first object moves toward the second (top panel) and stops at the point at which they are touching (middle panel). The second object begins moving at this moment or with a slight delay (bottom panel) and then stops. With appropriate configurations of space and velocity, this gives rise to a strong causal percept: the first object launches the second.



A TO B: STEPPING DOWN FROM THE CURB
A TO C: CROSSING STREET
A TO D: WALKING TO SCHOOL
A TO E: WORKING TO "PASS" FROM THE THIRD GRADE
A TO F: GETTING AN EDUCATION
A TO G: CLIMBING TO THE TOP IN LIFE


THE PANEL ACTS AS A SORT OF GENERAL INDICATOR THAT TIME OR SPACE IS BEING DIVIDED.