Is There an End in Sight?

Viewers' Sensitivity to Abstract Event Structure

Yue Ji[a] and Anna Papafragou[b]

a. Department of Linguistics and Cognitive Science, University of Delaware. 125 E. Main Street, Newark, DE, USA, 19716. E-mail: Jiyue@Udel.Edu

b. Department of Linguistics, University of Pennsylvania. 3401-C Walnut St., Philadelphia, PA, USA, 19104. E-mail: anna4@sas.upenn.edu

Correspondence concerning this article should be addressed to Yue Ji, Department of Linguistics and Cognitive Sciences, University of Delaware, DE, 19716. E-mail: Jiyue@Udel.Edu

Abstract

People segment the stream of experience into events, or temporal segments that have a beginning and an ending. But how are such event boundaries defined? Linguistic theories of event encoding draw a distinction between *bounded* events that include an inherent endpoint (“eat a pretzel”) and *unbounded* events that lack such an endpoint (“eat cheerios”). Even though the literature on event cognition has not focused on such abstract aspects of event structure, we hypothesize that sensitivity to boundedness could shape the way events are processed. In the present study, we show that viewers are sensitive to event boundedness in a category identification task and distinguish it from event completion; furthermore, viewers identify bounded events more easily than unbounded events. Sensitivity to boundedness emerges even when viewers are prevented from encoding the events linguistically and thus does not depend on the online use of linguistic distinctions. We conclude that event cognition relies on highly abstract properties of events and their boundaries, and sketch implications of these findings for the way events are described, processed, and used to interact with the world.

Keywords: event, telicity, aspect, boundedness, event cognition

Is There an End in Sight? Viewers' Sensitivity to Abstract Event Structure

1. Introduction

People segment continuous experience into events, themselves broadly defined as temporal segments that are perceived to have "a beginning and an end" (Zacks & Tversky, 2001). According to an influential proposal (Zacks & Swallow, 2007), people identify the boundaries of an event by tracking changes in perceptual features such as the direction, location, or speed of an action (cf. an arrow hitting a target), or conceptual features such as the goal-directedness or causal structure of an experience (cf. a person on a diet hitting a target). An event boundary is perceived when changes along a perceptual or conceptual dimension accumulate during a specific temporal window and create uncertainty about what may happen next: in such cases of maximal prediction error, people update their event model and perceive an event boundary.

Event boundaries have been shown to have a privileged status in memory and to provide anchors for later learning and describing (Swallow, Zacks, & Abrams, 2009). Event endpoints, in particular, have been argued to be critical for how events are represented. For instance, when comparing two events, the resultant state (e.g., whether a ball knocked over a tower or just a few blocks from the tower structure) has been shown to have more psychological weight than other event features (e.g., the moving direction of the ball) (He & Arunachalam, 2016; cf. also Sakarias & Flecken, 2018; Strickland & Keil, 2011). Relatedly, in the well-studied domain of motion events, the goal of motion (i.e., the endpoint, as in *Maria ran* *to the store*) is more accurately encoded in both language and memory as opposed to the source (as in *Maria ran* *away from the house*; Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007).

Despite the richness of the literature on event segmentation and its emphasis on the salience of endpoints, the *nature* of event endpoints has been less discussed. In most event segmentation studies, the stimuli are actions by an intentional actor (e.g., a person putting up a tent) and the endpoint is taken to be well-defined (e.g., the moment the tent is put up; cf. Zacks & Swallow, 2007). In studies of motion events, the endpoint appears similarly self-evident and is typically the moment that a moving entity reaches the goal. However, across a broader range of events, the notion of endpoint is not always straightforward. Consider the following situations described by the sentences in (1) and (2):

(1) The girl played the Moonlight Sonata.

(2) The girl played the piano.

There is subtle difference between the two situations. The situation in (1) comes to an end when the last note of the sonata was played. By contrast, it is hard to specify how or when the situation in (2) ends: the girl could stop playing at any point. The endpoint is inherent in the former event but arbitrary in the latter. Such contrasts have been discussed extensively in the linguistic literature on situation aspect (i.e., the linguistic encoding of the temporal contour of events, otherwise known as *telicity*). In this literature, the distinction between the two event descriptions above is captured by assuming that (1) encodes an experience as a "bounded" event but (2) encodes it as an "unbounded" event (Bach, 1986; Jackendoff, 1991; Harley, 2003).

Bounded events have internal structure leading to a "built-in terminal point" (Comrie, 1976), "climax" (Vendler, 1957), or "culmination" (Parsons, 1990). If we divide a bounded event into temporal slices with minimal duration, each slice represents a different stage of development towards this endpoint. By contrast, unbounded events are homogenous (Hinrichs, 1985) or cumulative (Krifka, 1989); such events have no inherent endpoint and may terminate at any arbitrary moment. An unbounded event can be divided into any number of temporal slices and each slice can still be regarded as an event of the same kind. In this sense, bounded events can be considered "atomic" or individuated (Bach, 1986), and unbounded events unspecified with respect to these dimensions (for brevity, we will sometimes refer to them as non-atomic and unindividuated). The boundedness of event descriptions can be diagnosed via specific linguistic tests. For instance, a bounded description such as *play the Moonlight Sonata* licenses a delimited temporal phrase such as *in an hour,* which denotes that the event described by the verb phrase has reached its endpoint within a certain amount of time. By contrast, an unbounded description such as *play the piano* is congruent with a durative temporal phrase such as *for an hour*, which specifies how long the event denoted by the verb phrase has lasted (Dowty, 1979; Smith, 1991; Vendler, 1957). Notice that language can encode the same experience as a "bounded" or an "unbounded" event (Wagner & Carey, 2003). For instance, *play the Moonlight Sonata* in (1) and *play the piano* in (2) could be used to describe the same perceptual stimulus. Linguistic boundedness thus reflects a perspective on the temporal profile of experience.[1]

According to the rich linguistic research on how event endpoints are encoded in language, boundedness is computed in part compositionally through an interaction between the verb and its argument (e.g., Filip, 2012; Krifka, 1998; see Borer, 2005; Jackendoff, 1996 for a more nuanced view). Therefore, one can distinguish between two sources of intuitions about whether an event is bounded or not. First, intuitions about boundedness may be due to the nature of the action encoded in the sentential verb. In particular, some actions lead to a change of state in the affected

[1] Boundedness has been defined by some as "expressing an action tending towards a goal" (Garey, 1975) but this view has faced criticism (e.g., Pustejovsky, 1991; see discussion in Borik, 2006). First, bounded expressions may denote events that lack an agent and thus are not goal-directed (e.g., *The lake froze last night*). Second, unbounded expressions may describe events that involve an intentional agent whose action is goal-directed without specifying the goal. For instance, in (2), the girl could play the piano to finish an assignment from her piano class, or just to relax. Thus boundedness and goal-directedness are independent notions even though they can be related. Similar issues apply to causality which is neither sufficient nor necessary for the presence of an inherent endpoint (Levin, 2000; Van Valin & LaPolla, 1997). For instance, causative verbs may denote unbounded events (e.g., *I walked the dog for half an hour*) and, conversely, non-causative verbs may denote bounded events (e.g., *I crossed the finish line in half an hour*).

object, such that the endpoint is the resultant state (bounded events); other actions do not affect the object in a perceptible way or the change lacks a well-defined resultant state (unbounded events). The contrast is shown in (3)-(4) below: (3) describes a bounded event – the cards are split when the girl finishes, whereas (4) describes an unbounded event – no predictable result follows from the girls’ shuffling. Although both events involve the same object, the difference in actions leads to the contrast in boundedness:

(3) The girl split the cards.

(4) The girl shuffled the cards.

Second, intuitions about boundedness may be due to the nature of the affected object encoded in the argument of the verb. When the argument is a quantified count noun, as in (5), it picks out a specific quantity and thus the event is bounded. Other kinds of object expressions such as bare plurals or mass nouns, as in (6), do not quantify specific amounts and therefore do not place any specific boundary on the event. Here there is a homomorphism between the affected object and the time course of the event (Dowty, 1991; Krifka, 1989), such that the changes in the object track or “measure out” the way the event develops (Tenny, 1987). In the example in (5), the event unfolds as the bubble emerges and ends at the moment when the bubble is fully formed; in (6), by contrast, there is no clear number of bubbles that needs to be produced and thus no point at which the event “ends”, even though the boy can stop at any time. In that sense, the two events inherit their (un)boundedness from the properties of the objects involved.

(5) The boy blew a bubble.

(6) The boy blew bubbles.

Within linguistic theory, boundedness is a foundational distinction in the event domain that represents a basic commitment to the kind of state of affairs the speaker has in mind and wants to talk about (much like the count/mass distinction is a foundational distinction in the nominal domain; Bach, 1986). Boundedness has been argued to be a potential semantic universal (von Fintel & Matthewson, 2008), even though it can be encoded in different ways across languages (Bar-el, Davis, & Matthewson, 2005; Botne, 2003). Most importantly for present purposes, boundedness in language has been widely assumed to be grounded in non-linguistic cognition: on this view, "[t]he notions 'bounded' and 'unbounded' belong to a finite set of primitives that characterizes parts of conceptual structure" (Filip, 1993, p.10; cf. also Folli & Harley, 2006; Malaia, 2014).

Existing evidence supports the presence of a connection between linguistic and cognitive representations of boundedness. In a recent study (Strickland et al., 2015), people who lacked any knowledge of sign languages tended to map signs with a visual boundary to verbs describing bounded events (e.g., *decide*) and signs lacking a visual boundary to verbs describing unbounded events (e.g., *think*). Because this pattern mirrored the iconic way in which sign languages tend to express boundedness (signs for bounded events have a salient visual boundary while signs for unbounded events do not), it suggested that the iconic mapping between the meaning (i.e., boundedness) and form (visual boundary) of a sign is universally accessible to both signers and non-signers. Furthermore, neurological research has revealed that similar brain areas are activated when signers of American Sign Language process signs of bounded or unbounded events and non-signers perform event segmentation tasks (Malaia, Ranaweera, Wilbur, & Talavage, 2012). Lastly, we know that linguistic (un)boundedness can provide a zoom lens for construing events: upon seeing a video of a girl eating an ice-cream cone in three bites, adults (and, to a lesser degree, 3-to-5-year olds) give different counts to the question "How many times was the ice cream cone eaten?" compared to "How many times did the girl eat?" (Wagner &

Carey, 2003; cf. Barner, Wagner & Snedeker, 2008; Wellwood, Hespos, & Rips, 2018). Even though these finding are suggestive, the literature on event cognition currently lacks direct tests of the role of boundedness in determining event categories, and little is known about whether viewers are sensitive to such abstract aspects of event representation.

Incorporating the formal property of boundedness into theories of event cognition would allow natural connections between cognition and language: being able to categorize an event as having or lacking an inherent boundary can offer the conceptual starting point for planning aspectually marked sentences such as those in (1)-(6) (Levelt, 1989; Konopka & Brown-Schmidt, 2018; Papafragou & Grigoroglou, 2019), as well as scaffold the way learners acquire the tools for encoding aspect in their native tongue (van Hout, 2007, 2016, 2018; Wagner, 2012). More broadly, evidence that viewers are sensitive to the boundedness distinction for classes of events would support the frequently cited but rarely tested idea that the linguistic representation of events is aligned with and builds on components of cognitive representations (cf. Pinker, 1989; Papafragou, 2015; Shipley & Zacks, 2008).

Furthermore, a linguistically inspired notion of boundedness could offer a more nuanced way of defining endpoints in event cognition, and ultimately of delimiting events themselves. Currently, the most prominent event segmentation accounts (e.g., Zacks & Swallow, 2007; Swallow et al., 2009) offer a mechanism for chunking dynamic input from experience into discrete event tokens but do not address the representational content of such tokens or the way tokens are related to one another to form event types. These accounts were not designed to capture the intuition that otherwise dissimilar events such as those described in (1), (3) and (5) (and, correspondingly, (2), (4) and (6)) share an underlying representational structure, including an abstract notion of when each event ends. Similarly, these accounts were not meant to accommodate the intuition that, even though both the events of playing the Moonlight Sonata and playing the piano can come to an end, the endpoint represents something different in each case (culmination vs. mere cessation). Positing a cognitive ability to distinguish bounded from unbounded events can organize these intuitions. We propose that this ability is foundational for the very units of thoughts about events, much like the cognitive ability to distinguish between objects and substances is foundational for the very units of thoughts about entities in the world.

Finally, the boundedness profile of an event could have cognitive consequences for how that event is mentally processed, remembered, and used to predict or plan future actions. Returning to our examples in (1) and (2), thinking that the girl played the Moonlight Sonata anchors the girl’s action in terms of distinct points of internal development (corresponding to the beginning, midpoint and endpoint of the piece) that can be used to support memory but thinking that the girl played the piano includes no such reference points. Similarly, prior to the event, thinking that the girl will play the Moonlight Sonata creates different expectations about how long the performance should last (about 15 minutes) compared to thinking that the girl will play the piano (where no specific expectations arise). Each of these thoughts can have different consequences about further planning and decision-making. Current theories of how events are organized and retrieved from memory do not incorporate such facts (Elman & McRae, 2019; Cooper, 2019; Radvansky & Zacks, 2014; Swallow et al., 2009).

Here, in a series of four experiments, we investigate whether boundedness is represented in event cognition to address these issues. Specifically, we ask whether viewers can group events into the bounded vs. unbounded category in a category identification task. Drawing on the linguistic literature above in which the category of bounded events is definitionally more specific (in that it denotes atomic entities) compared to unbounded events, we further ask whether there is an asymmetric relation between the two types of events in non-linguistic cognition. If so, the category of bounded events might be identified by observers more easily compared to that of unbounded events.

2. Experiment 1

Experiment 1 was a category identification task. Participants were exposed to pairs of bounded and unbounded events and had to extract a generalization about one member of these pairs.

2.1. Method

2.1.1. Participants

Forty adults (20 female, 20 male; M_{age} = 19.3, age range: 18.0 – 23.1) participated in the experiment. All were undergraduates at the University of Delaware and received course credit for participation. Data from 2 additional adults were collected but excluded because these adults were color-blind and failed to identify an important test feature (a red frame) consistently. Our sample size was based on similar work on event perception and event categorization (Strickland & Keil, 2011; Strickland & Scholl, 2015; Mereu, Zacks, Kurby, & Lleras, 2014).

2.1.2. Stimuli

Twenty pairs of videos were created, such that each pair showed a bounded and an unbounded event (see Table 1). Within each pair, the videos had the same duration (range: 4.5s-13s; M = 7.98s). All of the videos involved the same actor performing an action in a lab room. The videos began with the actor picking up an object or tool, and came to an end with the actor putting down the object or tool and removing her hand. To create the videos, we were inspired by the linguistic literature detailing the factors that can determine whether an event is bounded or unbounded (see Introduction) and used two sources to create the contrast in boundedness across related events – the nature of the action and the nature of the affected object. For half of the cases, paired bounded and unbounded events involved the same object but differed in terms of the nature of the action performed on the object (Figure 1a and 1b): the bounded event displayed an action that caused a clear and temporally demarcated change of state in the object (e.g., dress a teddy bear) while its unbounded counterpart did not involve such a change (e.g., pat a teddy bear). For the other half of the cases, the bounded and unbounded events involved the same action but differed in terms of the nature of the affected object (Figure 1c and 1d): the bounded event involved a single object (e.g., a pretzel) but its unbounded counterpart involved either an unspecified plurality of objects or a mass quantity (e.g., a large bag of cheerios), where there was no certain amount that need to be completed by the actor, thus the event endpoint was indeterminate.

To ensure that all video stimuli would illustrate the contrast in boundedness, a separate group of 18 English native speakers from the same population was asked to watch a subset of the clips and describe what happened in a full English sentence. The 20 pairs of events were split into two lists, such that each list included only one member of each pair with boundedness and the source of boundedness counterbalanced. The 18 participants were randomly assigned to one of the two lists. Their descriptions were coded for the verb phrase which was composed of the verb used to describe the action and the noun phrase used to describe the affected object. The verb phrases underwent linguistic tests for boundedness (e.g., co-occurrence with phrases such as *in an hour* vs. *for an hour*; see Dowty, 1979; Smith, 1991; Vendler, 1957, and Introduction). As expected, stimuli of bounded events elicited bounded descriptions that included change-of-state verbs (e.g., *dress a teddy bear*) or quantified count noun phrases (e.g., *eat a/the pretzel*) 98.2% of the time. Stimuli of unbounded events elicited unbounded verb phrases that included verbs of activity (e.g., *pat a teddy bear*) or unquantified noun phrases (bare plurals or mass nouns: e.g., *eat cheerios /cereal*) 92.8% of the time.

Figure 1. Pairs of video stimuli in Experiment 1: (a) dress a teddy bear (bounded) vs. (b) pat a teddy bear (unbounded); and (c) eat a pretzel (bounded) vs. (d) eat cheerios (unbounded).

We asked a separate group of 20 participants to rate the degree of intentionality for all videos on a scale from 1 (totally unintentional) to 7 (intentional). There was no significant difference between scores for bounded events (M = 5.67) and unbounded events (M = 5.62) (t (19) = 1.34, p = .195). We also assessed the degree of visual cohesion of the bounded vs.

unbounded class of events used in the training phase (see Table 1). We created a new stimulus set by putting together all possible pairwise combinations of the 8 videos of bounded events and intermixing them with all possible pairwise combinations of the 8 videos of unbounded events. We asked a different group of 20 people to rate the degree of visual similarity for each pairwise combination on a scale from 1 (least similar) to 7 (most similar). The average rating for pairs of bounded events (M = 2.38) did not differ from those of unbounded events (M = 2.52) (t (19) = -1.43, p = .169).

The video stimuli were arranged into three basic sets corresponding to the three phases of the experiment (see Table 1). For the initial training phase, we arranged 8 pairs of events (4 in which boundedness was due to the Action and 4 in which boundedness was due to the Affected Object) into a pseudorandomized presentation list. Within each list, a single video was played in the center of the screen and was followed immediately by the other video within the pair (the order of bounded-unbounded events within pairs was counterbalanced within the list). For the later testing phase, we arranged another 8 pairs of videos into 2 lists. Each list contained one video from each pair. We counterbalanced whether the event was bounded or unbounded and whether source of boundedness was the action or the affected object across lists. For the final (short) surprise testing phase, we used the last 4 pairs of videos, arranged into 2 lists. The same counterbalancing was used as in the (main) testing phase.

Table 1

Video stimuli used in Experiment 1. Each row depicts a pair of events. In the training phase, participants saw both events within a pair. In the testing phase, participants saw only one event from each pair.

Phase	Boundedness Source	No.	Bounded Events	Unbounded Events
Training	Nature of Action	1	fold up a handkerchief	wave a handkerchief
		2	put up one's hair	scratch one's hair
		3	stack a deck of cards	shuffle a deck of cards
		4	group pawns based on color	mix pawns of two colors
	Nature of Affected Object	5	draw a balloon	draw circles
		6	tie a knot	tie knots
		7	eat a pretzel	eat cheerios
		8	flip a postcard	flip pages
Testing	Nature of Action	9	dress a teddy bear/	pat a teddy bear
		10	roll up a towel/	twist a towel
		11	fill a glass with milk/	shake a bottle of milk
		12	scoop up yogurt/	stir yogurt
	Nature of Affected Object	13	peel a banana/	crack peanuts
		14	blow a balloon/	blow bubbles
		15	tear a paper towel/	tear paper towels
		16	paint a star/	paint stuff
Surprise Testing	Nature of Action	17	close a fan/	use a fan for oneself
		18	crack an egg/	beat an egg
	Nature of Affected Object	19	cut a ribbon in half/	cut ribbon from a roll
		20	stick a sticker/	stick stickers

2.1.3. Procedure

Participants were randomly assigned to one of two conditions. In the Bounded condition, the videos of bounded events shown in the training phase were surrounded by a red frame while their unbounded counterparts were surrounded by a black frame. In the Unbounded condition, the reverse assignment occurred.

In the training phase for both conditions, participants were asked to watch a few videos and to pay attention to those appearing within a red frame. Their task was to figure out what kind of videos were given the red frame and to decide whether a new video could have the red frame or not. In the testing phase, participants saw a new set of videos and for each one they were asked: "Could the video have a red frame or not?" (test question). Lastly, in the surprise testing phase, participants were unexpectedly asked: "Could the video have a black frame or not?" (surprise question). This question was included to probe whether participants had formed any hypotheses about the secondary event category present within the experiment, even though it was not the target of identification. Additionally, at a minimum, if participants were attentive to the red frame within the task, they should perform better at test compared to surprise questions.

At the end of the session, participants were asked to write down what kind of videos could have a red frame. This was used as an additional source of information about the category that participants had just formed.

2.2. Results

Results from Experiment 1 are shown in Figure 2. The data from this experiment (and all subsequent experiments) were analyzed using multi-level mixed modeling with crossed intercepts for Subjects and Items (Baayen, Davidson, & Bates, 2008; Barr, 2008). All models were fitted using the *glmer* function of the *glmertree* package in R. We first examined responses in test questions. The binary accuracy data were submitted to a logit model with Condition (Bounded vs. Unbounded) as the fixed factor. [2] As shown in Table 2, the model revealed an effect of Condition, with greater accuracy in the Bounded (M = 92.5%) than in the Unbounded

[2] We also examined whether the Source of Boundedness (Action vs. Affected Object) would influence accuracy. Neither this predictor nor its interaction with Condition (Bounded vs. Unbounded) significantly improved the model fit. Therefore, Boundedness Source was excluded from further analyses.

condition (M = 76.3%) (z = -3.49, p < .001). Performance was significantly different from chance level in both the Bounded (t (19) = 20.17, p < .001) and the Unbounded condition (t (19) = 6.49, p < .001). The responses in surprise questions were analyzed in the same way (Table 2). No significant difference between the two conditions was found (z = 0.68, p >.250). As in the test questions, participants performed at levels different from chance in both the Bounded (t (19) = 3.94, p = .001) and the Unbounded condition (t (19) = 4.29, p < .001).

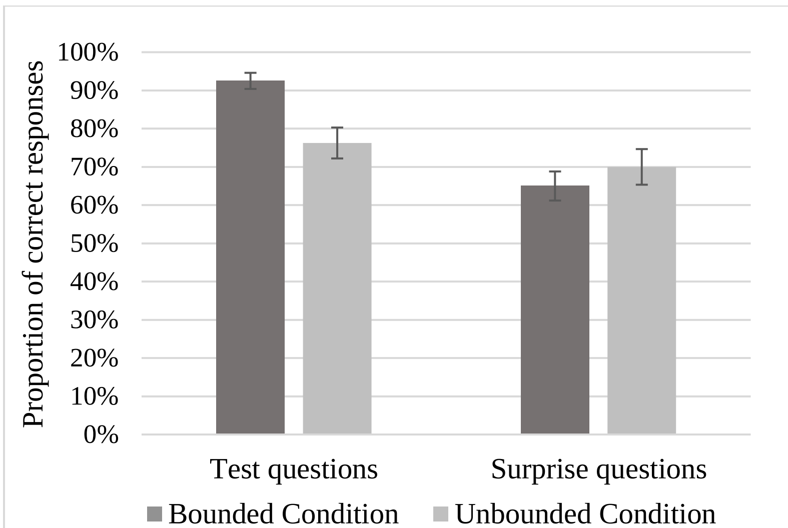

Figure 2. Proportion of correct responses in Experiment 1. Error bars represent ±SEM. Note that surprise questions probed a non-target category (unbounded events in the Bounded condition, and bounded events in the Unbounded condition).

Table 2

Fixed effect estimates for multi-level model of event category identification in Experiment 1.

Question Type	Effect	Estimate	SE	z value
Test question	(Intercept)	2.74	0.41	6.74***
	Condition	-1.43	0.41	-3.49***
Surprise question	(Intercept)	0.62	0.23	2.64**
	Condition	0.22	0.34	0.68

Note. Formula in R: AccTest ~ 1 + (1|ID) + (1|Item) + Condition; AccSurprise ~ 1 + (1|ID) + (1|Item) + Condition

* $p < .05$, ** $p < .01$, *** $p < .001$

To see if the presence of the red frame affected category identification, we conducted an additional analysis on all data examining the fixed effects of Question Type (Test vs. Surprise), Condition (Bounded vs. Unbounded), and the interaction between them (random slopes did not significantly improve model fit and were thus excluded from further analyses; see Table A.1 in Appendix A for this analysis). There was a main effect of Question Type (z = -4.57, $p < .001$), such that participants were more accurate in answering test questions compared to surprise questions, and of Condition ($z = -3.84$, $p < .001$), such that participants were better at the Bounded than the Unbounded condition. There was also an interaction between Question Type and Condition ($z = 3.24$, $p = .001$): in the Bounded condition, participants were better in test questions compared to surprise questions ($z = -4.97$, $p < .001$), but the difference was not significant in the Unbounded condition ($z = -.83$, $p > .250$). Thus the presence of a cue indicating the target category drove identification only for bounded events.

Responses to the final question about the meaning of the red frame confirmed that participants paid attention to the internal structure of events. Of a total of 40 answers, 55% made reference to the organization in the stimuli (with bounded events called “organized” or “structured”, and unbounded events “unorganized” or “lacking structure”); 30% referred to neatness (with bounded events called “neat” and “clean”, and unbounded events “messy” and “untidy”); and 15% referred to intention (bounded events were characterized as “intentional”, or “on purpose”, and unbounded events as “lacking an end or purpose”).

2.3. Discussion

After a brief exposure to a few contrastive examples of bounded and unbounded events, viewers were able to extract the (un)boundedness category and extend this category to new exemplars. Furthermore, as shown by the surprise trials, participants also extracted a generalization about the non-target category in both conditions. These two findings show that participants were sensitive to the dimension of boundedness when perceiving events, and grouped events into boundedness categories even when not explicitly instructed to do so by the task.

A particularly interesting finding was that participants were better at forming the category of bounded events compared to that of unbounded events. In further support of this asymmetry, in the Bounded condition, participants were more successful when asked about the target category (test questions) compared to a non-target category (surprise questions), but in the Unbounded condition there was no target category advantage. Given that events within each category were equated for intentionality and visual similarity, we hypothesize that the advantage of the bounded category relates to the nature of the category itself. We revisit this hypothesis in the following sections.

3. Experiment 2

In Experiment 1, participants were exposed to successively presented pairs of bounded and unbounded events in the training phase, a presentation choice that highlighted the contrast between the two event categories. In Experiment 2, we introduced a training phase with single instances of events and asked whether participants could identify boundedness categories without any support from contrast.

3.1. Method

3.1.1. Participants

A new group of forty undergraduates (22 female, 18 male; M_{age} = 19.4, age range: 18.0 - 24.0) at the University of Delaware completed the experiment. Data from 2 additional adults were collected but excluded because they failed to finish all the questions.

3.1.2. Stimuli

Video stimuli in the testing phase were identical to those used in Experiment 1. For the training phase, we split the 8 pairs of videos used in the training phase of Experiment 1 into 2 lists of 8 individual videos and ensured that no list contained both members of a single pair. Half of the events within each list were bounded and the other half unbounded (Source of Boundedness was also counterbalanced). To create a training set of equal length to Experiment 1 (16 total videos), we created a set of additional videos using the same actor as Experiment 1. Half depicted bounded events (stack 5 cups on the table, put some Q-tips together, erase a star, organize 3 pairs of socks by color) and the other half unbounded events (grind biscuits, roll a ball back and forth, pull a towel, sprinkle pepper). We then distributed these videos across the two training lists, such that each final list contained 8 bounded and 8 unbounded videos (unpaired) presented in a pseudorandomized order.

We conducted two norming studies on the 8 newly-introduced stimuli following our practice in Experiment 1. For the first of these, 18 new adults were asked to describe each of the newly-created videos in a full English sentence. Bounded events elicited change-of-state verbs or quantified noun phrases 97.2% of the time, while unbounded events elicited verbs of activity or unquantified noun phrases 88.9% of the time. For the second study, a separate group of 20 people did a visual similarity judgment task as in Experiment 1. There was no significant difference between the rating of members of the bounded (M = 2.58) and unbounded (M = 2.63) events class (t (19) = -.66, p > .250).

3.1.3. Procedure

Participants were randomly assigned to either the Bounded or the Unbounded condition. Within each condition, participants were split between the two training lists. Otherwise, the procedure was identical to that of Experiment 1.

3.2. Results

Results from Experiment 2 are shown in Figure 3. As shown in Table 3, accuracy of the test questions was significantly higher in the Bounded (M = 80.6%) than in the Unbounded condition (M = 53.8%) (z = -5.00, p < .001). By contrast, accuracy of the surprise questions did not differ between the Bounded (M = 62.5%) and the Unbounded condition (M = 56.3%) (z = -.075, p > .250). Unlike the Bounded condition (t (19) = 15.96, p < .001), performance on the test trials in the Unbounded condition was not significantly different from chance level (t (19) = .84, p >.250). Performance was not above chance level in either condition for surprise questions (Bounded: t (19) = 1.81, p = .086; Unbounded: t (19) = .96, p >.250).

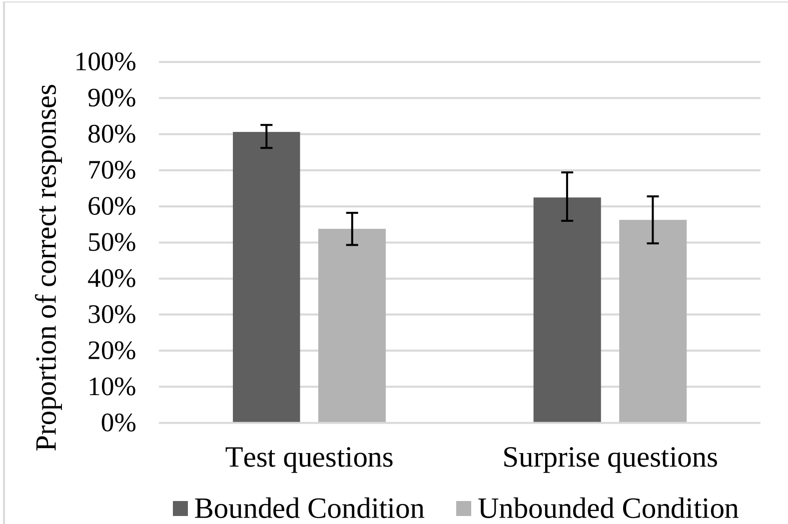

Figure 3. Proportion of correct responses in Experiment 2. Error bars represent ±SEM.

Table 3

Fixed effect estimates for multi-level model of event category identification in Experiment 2.

Question Type	Effect	Estimate	SE	z value
Test question	(Intercept)	1.43	0.20	7.13***
	Condition	-1.28	0.26	-5.00***
Surprise question	(Intercept)	0.81	0.30	2.71**
	Condition	-0.30	0.40	-0.75

Note. Formula in R: AccTest ~ 1 + (1|ID) + (1|Item) + Condition; AccSurprise ~ 1 + (1|ID) + (1|Item) + Condition

* $p < .05$, ** $p < .01$, *** $p < .001$

The binary accuracy data were submitted to a mixed logit model with Question Type (Test vs. Surprise), and Condition (Bounded vs. Unbounded) as predictors (see parameter estimates of fixed effects in Table A.2). There were main effects of Question Type ($z = -2.25$, $p = .025$; with performance being better in test than in surprise questions) and Condition ($z = -4.72$, *p* < .001; with accuracy being higher in the Bounded than the Unbounded condition), as well as a significant interaction between the two predictors (*z* = 2.42, *p* = .016). Specifically, participants were more accurate in test questions compared to surprise questions (*z* = -2.23 *p* = .026) in the Bounded condition but no such difference was found in the Unbounded condition (*z* = 1.13, *p* > .250).

Answers to the open question about the nature of the target category were more diffuse compared to previous experiments. Organization, neatness and intention (or lack thereof) accounted for 35%, 25% and 10% of responses respectively; however, 30% of the answers just listed individual events seen in the experiment. These participants had difficulty summarizing the common event structure underlying the examples in the training phase.

To compare performance between Experiment 1 and 2, the binary data from test questions were submitted to a mixed logit model with Condition (Bounded vs. Unbounded) and Experiment (1 vs. 2) as predictors. Table 4 reports parameter estimates of fixed effects. There was a fixed effect of Condition (*z* = -6.17, *p* < .001), with accuracy being higher in the Bounded compared to the Unbounded condition. There was also an effect of Experiment: performance was better in Experiment 1 where contrastive pairs were available in the training phase than Experiment 2 (*z* = -5.05, *p* < .001). Responses in surprise questions were analyzed in the same way. However, neither Condition (Bounded vs. Unbounded), nor Experiment (1 vs. 2) or their interaction could improve the model fit.

Table 4

Fixed effect estimates for multi-level model of event category identification in Experiments 1 and 2.

Effect	Estimate	SE	z value
(Intercept)	2.50	0.24	10.64***
Condition	-1.31	0.21	-6.17***
Experiment (1 vs. 2)	-1.05	0.21	5.05***

Note. Adding the interaction between Condition and Experiment to the model did not significantly improve the fit. Therefore, this interaction is not reported in the table.

Formula in R: Acc ~ 1 + (1|ID) + (1|Item) + Condition + Experiment

* $p < .05$, ** $p < .01$, *** $p < .001$

3.3. Discussion

In Experiment 2, the lack of contrastive pairs during training made the category identification task harder than in Experiment 1. Exposed to single examples of an event category, participants were able to categorize bounded events, but failed to identify the category of unbounded events. It seems that unbounded events could be identified as a group only through comparisons with their bounded counterparts. We return to the significance of this asymmetry in later sections.

Could sensitivity to a property other than boundedness account for the data in Experiments 1 and 2? One possibility is that viewers attended to intentionality or goal-directedness in the videos that might be thought to characterize bounded more than unbounded events. Recall, however, that the stimuli within the bounded and unbounded class were judged to be equally intentional in a prior norming study (see Experiment 1), and so this possibility is unlikely.

An alternative possibility is that viewers might have relied on physical information in the videos to classify events without extracting the abstract boundedness feature. We tried to remove this possibility, particularly in Experiment 2, by having bounded and unbounded events in the training phase involve several different actions and objects (and ensuring that events within a class would be equally visually similar; see norming data). One might point out that at least some of the unbounded event stimuli (e.g., draw circles, tie knots) suggest repetition while their bounded counterparts do not (e.g., draw a balloon, tie a knot). However, in our stimuli, overall, around one third of the bounded events involved repetition; for instance, in the video of stacking 5 cups on the table, the action of putting one cup on top of another was repeated. Most importantly, answers to the debriefing question across both Experiment 1 and 2 mentioned physical properties of the displays extremely rarely – for instance, fewer than 5% of responses referred to repetitiveness. As discussed already, the majority of responses made reference to the organization of the stimuli, with bounded events described as "organized" or "structured", and unbounded events as "unorganized" or "random". We take this as evidence that participants formed a specific type of abstract generalization about the target event category.

4. Experiment 3

A different alternative explanation for the patterns of performance in Experiments 1 and 2 is that participants tracked simply whether the actor in the videos completed the depicted action, especially for bounded events. Completion (realizing an endpoint) is different from boundedness (having an inherent endpoint, whether realized or not; Dahl, 1981). Debriefing responses already mitigate against this possibility, since they refer to organization or structure as opposed to completion. To exclude this possibility more definitively, especially given the subtlety of the difference, in the training phase of Experiment 3, participants watched the same videos as in Experiment 1; however, in the testing phase, half of the videos depicted the entire event, as in previous experiments (full videos), and the other half only the very beginning of the event (truncated videos). The two types of videos differed in terms of whether the events were actually completed but not in terms of whether the events were bounded (i.e., whether a potential end was in sight): even for the truncated stimuli, we anticipated participants to project the event and its predicted endpoint (e.g. eating a pretzel), despite the sparse evidence (e.g., seeing the actor take a first bite).

If participants in the Bounded condition based their conjectures on event completion after watching full videos in the training phase, they should only accept full videos as members of the learned category in the testing phase. However, if participants based their conjectures on whether the event had a potential, but not necessarily realized, endpoint, they should generalize the learned category to both full and truncated videos.

4.1. Method

4.1.1. Participants

A new group of forty undergraduates (21 female, 19 male; M_{age} = 19.6, age range: 18.4 – 22.5) at the University of Delaware participated in the experiment.

4.1.2. Stimuli

In the training phase, videos were the same as in Experiment 1. In the testing phase, half of the videos were truncated and the other half were full. The truncated videos were created by editing the 8 pairs of videos used in the testing phase of Experiment 1 such that only the first 25% of each video remained (see Figure 4 for examples). As a result, truncated bounded events were incomplete while truncated unbounded events were unterminated. The full videos consisted of the 8 unpaired videos (4 bounded, 4 unbounded) that had been created for the purposes of Experiment 2.

The truncated videos were arranged into 2 lists in the same way as in Experiment 1. Then the 8 full videos were added to each list and intermixed with the truncated videos. Each of the resulting testing lists was thus composed of 16 videos. Whether the event was bounded or unbounded and whether the video was truncated or full was counterbalanced.

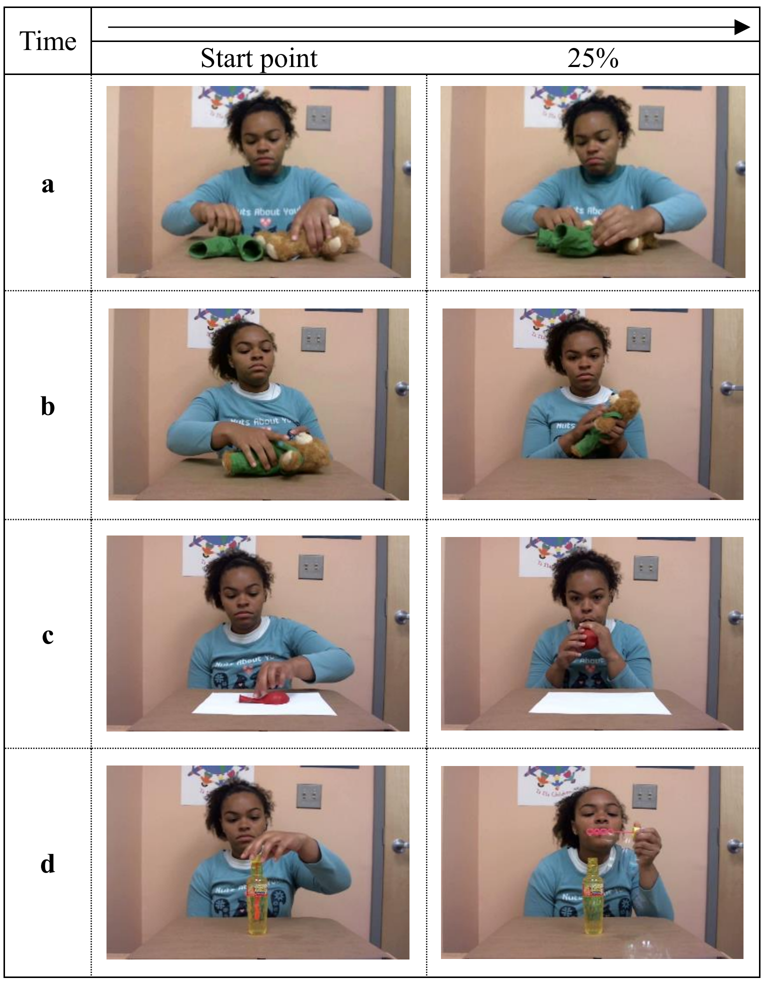

Figure 4. Truncated videos in Experiment 4: (a) dress a teddy bear vs. (b) pat a teddy bear, (c) blow a balloon vs. (d) blow bubbles.

4.1.3. Procedure

The procedure was identical to that of Experiment 1 except that the surprise testing phase was removed.

4.2. Results

Results from Experiment 3 are shown in Figure 5. We examined the fixed effects of Condition (Bounded vs. Unbounded) and Video Type (Truncated vs. Full) on accuracy. As shown in Table 5, a main effect of Condition was detected (z = -2.49, p = .013): accuracy was higher in the Bounded (M = 85.0%) than in the Unbounded condition (M = 73.8). There was no effect of Video type (z = -0.87, p > .250), and no interaction between Video Type and Condition (z = 0.43, p > .250). Thus participants treated full and truncated videos similarly. In both conditions, responses to both full and truncated videos differed from chance level (all *ps* <.001).

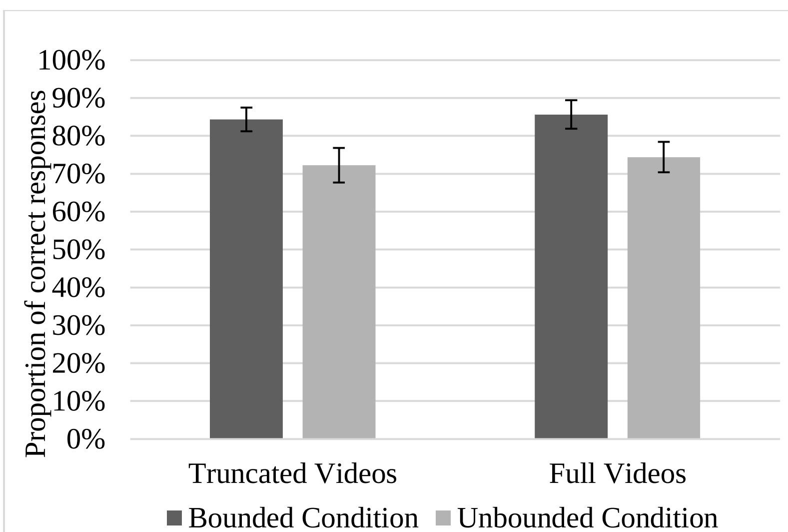

Figure 5. Proportion of correct responses in Experiment 3. Error bars represent ±SEM.

Table 5.

Fixed effect estimates for multi-level model of event category identification in Experiment 3.

Effect	Estimate	SE	z value
(Intercept)	2.18	0.36	6.00***
Condition	-0.93	0.37	-2.49*
Video Type	-0.39	0.45	-0.87
Video Type*Condition	0.18	0.42	0.43

Note. Formula in R: Acc ~ 1 + (1 | ID) + (1 | Item) + Condition + Video Type + Condition : Video Type

* $p < .05$, ** $p < .01$, *** $p < .001$

In support of this conclusion, answers to the open question showed that participants mainly associated the target category with presence or absence of organization, neatness and intention (50%, 35% and 10% of responses respectively). Completion appeared in only 2 answers (5% of the time).

4.3. Discussion

In Experiment 3, participants were able to classify an event into the bounded vs. unbounded category even when they had watched only the beginning of the event in the testing phase. This suggests that participants did not simply track whether an event was completed or not but instead formed conjectures about the inherent temporal profile of events.

5. Experiment 4

A further alternative explanation of the results of Experiments 1-3 is that participants might have subliminally encoded what they saw in linguistic terms and then identified the target event category based on these prior linguistic descriptions. If so, the results would reflect a linguistic distinction (the encoding of boundedness in English verbs and verb phrases), not a dimension of non-linguistic event cognition. In Experiment 4 we introduced a secondary counting task to disrupt linguistic encoding in the training phase and assessed whether this affected viewers' computation of boundedness categories.

5.1. Method

5.1.1. Participants

A new group of forty undergraduates (22 female, 18 male; M_{age} = 18.9, age range: 18.0 – 21.5) at the University of Delaware participated in the experiment. Data from 4 additional adults were collected but excluded because they failed the dual task.

5.1.2. Stimuli

The stimuli and procedure were identical to those in Experiment 1 with one exception. In the training phase, a 3-digit number appeared in the upper right corner of the screen at the onset of the first, the fifth, the ninth and the thirteenth video and was displayed for 3s. Participants were asked to count forward by 2 from the number they saw until they encountered a new number, at which point they had to switch to the new number and start counting forward by 2. This secondary task of counting numbers served to suppress any linguistic encoding of the events throughout the training phase.

5.2. Results

Results from Experiment 4 are shown in Figure 6. Separate accuracy data for test and surprise questions were submitted to a logit model with Condition (Bounded vs. Unbounded) as the fixed factor (see Table 6). As the table shows, accuracy in test questions was significantly higher in the Bounded (M = 82.5%) than in the Unbounded condition (M = 68.1%) (z = -3.06, p = .002). By contrast, accuracy in surprise questions saw no significant difference between the two conditions (Bounded: M = 50.0%, Unbounded: M = 47.5%; z = -.60, p > .250). Answers to the test questions differed from chance level in both the Bounded condition (t (19) = 15.42, p < .001) and the Unbounded condition (t (19) = -5.66, p < .001), but answers to the surprise questions were not significantly different from chance level in either condition (Bounded: t (19) = 0.00, p > .250; Unbounded: t (19) = -.42, p > .250).

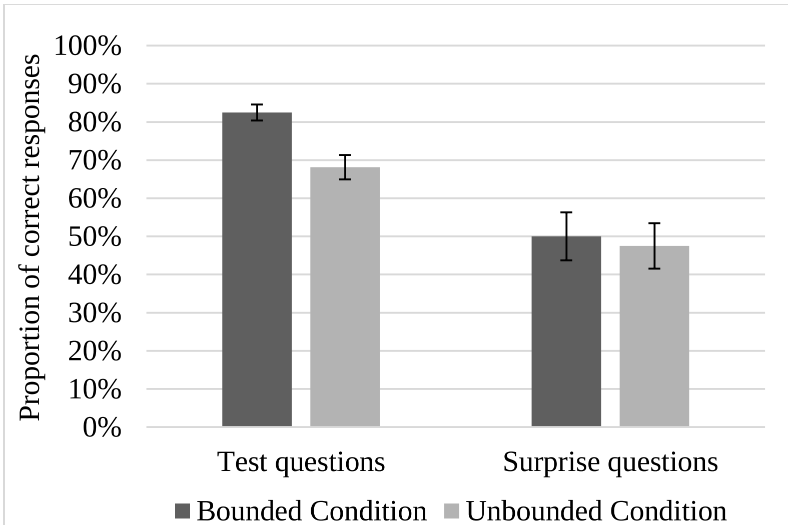

Figure 6. Proportion of correct responses in Experiment 4. Error bars represent ±SEM.

Table 6

Fixed effect estimates for multi-level model of event category identification in Experiment 4.

Question Type	Effect	Estimate	SE	z value
Test question	(Intercept)	1.55	0.21	7.45***
	Condition	-0.82	0.27	-3.06**
Surprise question	(Intercept)	0.01	0.24	0.00
	Condition	-0.30	0.35	-.60

Note. Formula in R: AccTest ~ 1 + (1|ID) + (1|Item) + Condition; AccSurprise ~ 1 + (1|ID) + (1|Item) + Condition

* p < .05, ** p < .01, *** p < .001

All data were submitted to a mixed logit model with Question Type (Test vs. Surprise), and Condition (Bounded vs. Unbounded) as predictors (see Table A.3 in Appendix A for this analysis). There were main effects of Question Type ($z = -5.08$, $p < .001$; with better performance in test questions compared to surprise questions) and of Condition ($z = -3.06$, $p = .002$; with greater accuracy in the Bounded than the Unbounded condition), but no significant interaction between them ($z = 1.49$, $p = .136$).

Of the 40 answers to the open question about the meaning of the red frame, organization, neatness and intention (or lack thereof) appeared 40%, 30% and 15% of the time respectively. In 15% of answers participants just gave examples of events seen in the experiment, presumably because the secondary task made performing the main task more difficult.

This difficulty was confirmed in an analysis comparing performance in Experiments 1 and 4 by examining the fixed effects of Question Type (Test vs. Surprise), Condition (Bounded vs. Unbounded), and Experiment (1 vs. 4), as well as their interactions (see Table 7). Unsurprisingly, there was a main effect of Question Type ($z = -6.97$, $p < .001$) and Condition ($z = -4.85$, $p < .001$), as well as an interaction between them ($z = 3.29$, $p = .001$). There was also a main effect of Experiment, with performance being worse in the dual task of Experiment 4 compared to Experiment 1 ($z = -4.51$, $p < .001$).

Table 7

Fixed effect estimates for multi-level model of event category identification in Experiment 1 and 4.

Effect	Estimate	SE	z value
(Intercept)	2.35	0.20	11.82***
Question Type	-1.68	0.24	-6.97***
Condition	-1.03	0.21	-4.85**
Experiment	-0.71	0.15	-4.51***
Question Type* Condition	1.03	0.31	3.29**

Note. The interaction reported in the Table significantly improved the model fit compared to the corresponding model without the interaction term. Other interactions do not appear because adding them to the models did not significantly improve the fit.

Formula in R: Acc ~ 1 + (1|ID) + (1|Item) + Question Type + Condition+ Experiment + Question Type : Condition

* $p < .05$, ** $p < .01$, *** $p < .001$

5.3. Discussion

Experiment 4 showed that viewers were able to form boundedness-driven generalizations without resorting to language. As in previous experiments, performance was better for bounded compared to unbounded events. Performance was worse in this dual task compared to Experiment 1 (and, unlike Experiment 1, performance in the surprise questions was no different from chance – presumably because the dual task increased cognitive effort and prevented participants from extracting task-irrelevant generalizations). We conclude that boundedness is encoded in the non-linguistic representation of events.

6. General Discussion

Our findings show that viewers are sensitive to the internal temporal profile of events – specifically, they draw a distinction between bounded events that possess a finely differentiated internal structure leading toward an inherent endpoint and unbounded events that are characterized by a largely undifferentiated structure with no discernible endpoint (Experiments 1-2). Viewers treat boundedness in event cognition as different from simple event completion (Experiment 3). Furthermore, viewers identify bounded and unbounded events even when linguistic encoding is suppressed (Experiment 4). Throughout our experiments, the category of bounded events was identified with greater ease compared to that of unbounded events. We conclude that event cognition incorporates abstract properties of events and their boundaries.

In the next sections, we elaborate on the implications of the sensitivity to boundedness for cognitive and linguistic theories (Section 6.1) and explore the theoretical importance of the difference between boundedness categories (Section 6.2).

6.1. Boundedness in language and event cognition

Within the linguistics literature, it is often assumed that the notions "bounded" and "unbounded" (originally developed to capture linguistic generalizations in the aspectual domain) extend beyond the realm of language – specifically, they "belong to a finite set of primitives that characterizes parts of conceptual structure" (Filip, 1993, p.10; cf. also Folli & Harley, 2006; Malaia, 2014; and Introduction). Our findings offer the first direct piece of support for the psychological reality of boundedness distinctions in non-linguistic cognition. These data throw light onto the conceptual foundations of aspect, and support the presence of homologies between event representations in language and cognition more generally (Folli & Harley, 2006; Harfi, Papafragou, & Trueswell, 2013; Lakusta & Landau, 2005; Malaia, 2014; Strickland et al., 2015; Tversky, Zacks, Morrison, & Hard, 2011).

Within the domain of event cognition, our results bear on theories of event segmentation, since they suggest that the very definition of event needs to accommodate a highly schematic construal of event boundaries, especially endpoints (cf. Radvansky & Zacks, 2014; Zacks & Tversky, 2001). In the minds of current participants, otherwise dissimilar events within the bounded (or unbounded) class share internal structure, including a notion of when each event ends. Moreover, people can make assumptions about the presence of an event boundary even when the boundary cannot be observed in the visual stream (see Experiment 3). This approach is compatible with the basic insight from event segmentation theories, namely that event boundaries coincide with significant changes in event features (Swallow et al., 2009; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). However, the present approach distinguishes different types of boundaries: in bounded events, the endpoints are the moments when maximum changes occur in the affected objects (e.g., the moment a teddy bear gets dressed); in unbounded events, the endpoints do not involve salient changes in the objects but other discontinuities in the input (e.g., the moment a teddy bear stops being patted). Furthermore, on the present approach, the property of (un)boundedness goes beyond the factors that have been taken to characterize event breakpoints within event segmentation models: it is neither reducible to lower-level perceptual features of the stimuli nor is it identical to higher-level conceptual features such as intentionality (see discussion of Experiment 2). Rather, as a property of a *class* of events, it is closer to an architectural feature of event structure, the very temporal backbone of events themselves (reflected also in participants' overt conjectures about the "organization", "neatness", or lack thereof, within the target categories).

The present approach has the potential to inform broader theories of event organization. A first major class of such theories seeks to model general knowledge underlying event construals and event memory, including, e.g., the contribution of event participants such as agents, patients and instruments, the interactions of such participants over space and time, and the role of higher-order factors such as causality and intentionality, especially for animate participants, in structuring event knowledge (Elman & McRae, 2019; Cooper, 2019; Jackendoff, 2007; Radvansky & Zacks, 2014; cf. also earlier work on scripts, frames or schemas; Bartlett, 1932; Schank & Abelson, 1977; Rumelhart, 1980). Even though these theories capture many aspects of event knowledge, they are not designed to address the granularity of representations that support event encoding in all its complexity, especially from a linguistic perspective. The present work strongly suggests that part of the representation of events involves sensitivity to abstract properties such as (un)boundedness that can form the impetus for linguistic-semantic event encoding.

A second, more recent kind of model proposes that events are represented as a series of intersecting representations of the objects in them (Altmann & Ekves, 2019). This line of research has examined the brain activity of participants as they read event descriptions including either a substantial change of state in an object (e.g., an onion is chopped) or a minimal change (e.g., an onion is weighed) (Hindy, Altmann, Kalenik, & Thompson-Schill, 2012; Solomon, Hindy, Altmann, Thompson-Schill, 2015). The results revealed that people are sensitive to different degrees of change that an object has undergone in the described event, such that event comprehension involves representations of an object's initial and final states as well as the trajectory between the two states. Our study shows that event observers are sensitive to not only the degrees but also the *nature* of the change of state in objects, i.e., whether the change has a specified resultant state. However, in our data, the nature of the affected object did not behave differently from the nature of the action in defining the type of change in the event. It would be important to compare these two factors more systematically in future studies of how (un)boundedness is conceptualized.

6.2. An asymmetry between bounded and unbounded events

A particularly striking aspect of our results was the fact that, across experiments, the category of bounded events was identified with greater ease compared to that of unbounded events. This asymmetry occurred even after we ensured that exemplars for each class used in the training phase were equated in terms of intentionality and visual similarity to other members of the class. As far as we know, this asymmetry is not explicitly anticipated by any current linguistic theory of aspect, and is also novel from the perspective of theories of event cognition. In what follows, we show that the asymmetry can be used to make new contributions to both types of theory.

We propose that the presence of internal structure that culminates in defined endpoints makes bounded events easier to individuate, track, compare to each other and generalize over compared to unbounded events – an important difference that explains their steady advantage in our categorization task. Specifically, during the categorization task, participants had to inspect a class of event tokens, extract similarities across the tokens and identify the class that all tokens instantiated. "Atomic", i.e., bounded, events (see Introduction) were better individuals and therefore provided a more stable unit of comparison across instances: the defined endpoints could provide enough information for participants to detect similarities across members of the class. In contrast, "non-atomic", i.e., unbounded, events were characterized by *lack* of defined endpoints. Precisely because participants in the Unbounded condition had to figure out what was *missing* in the class of events they were exposed to, their performance fell to chance levels when there was no contrast between bounded-unbounded event pairs in the training phase (as in Experiment 2).

This perspective is reminiscent of the classic linguistic assumption that boundedness in the domain of events has a close quantificational counterpart in the domain of objects. Specifically, a number of commentators have argued that the notion of atomicity characterizes the linguistic semantics of both events and objects (Bach, 1986; Jackendoff, 1991; see also Wagner & Carey, 2003; Wellwood et al., 2018). According to this idea, bounded events (expressed, e.g., by telic phrases such as *dress a teddy bear*) resemble objects (expressed, e.g., by count nouns such as *a sandcastle*) because they are "atoms" that can be individuated and compared to each other. Similarly, unbounded events (expressed, e.g., by atelic phrases such as *pat a teddy bear*) resemble substances (expressed, e.g., by mass nouns such as *sand*) in that they are unspecified for atomic features.

Evidence for the parallel between the cognitive representations of events and objects has so far mostly come from studies of how linguistic form connects to conceptualization (e.g., Prasada, Ferenz, & Haskell, 2002; Barner et al., 2008; Wellwood et al., 2018). In a recent study (Wellwood et al., 2018), people strongly preferred to label both images with "natural" spatial breaks and animations with "natural" temporal breaks using plural count nouns (e.g., *some gleebs*) and telic verb phrases (e.g., *do some gleebs*) but labelled unnaturally-divided – and hence probably non-atomic – images and animations mostly with mass labels (e.g., *some gleeb/ do some gleebing*). In addition, a handful of studies employing cognitive measures have revealed suggestive commonalities in object and (bounded) event construal for both infants (e.g., Sharon & Wynn, 1998; cf. Wood & Spelke, 2005) and adults (Maguire, Brumberg, Ennis, & Shipley, 2011). Within this context, our own data are consistent with the perspective that abstract underlying construals of entities as atomic or non-atomic determine, in part, the dimensions along which people compare them. Most importantly, our data offer the first piece of evidence from a cognitive task that abstract considerations of atomicity extend beyond the domain of spatially extended entities (objects/substances) to the domain of temporally extended entities (bounded/unbounded events).[3]

6.3. Final thoughts

To conclude, across four experiments, we have found that people can extract abstract information about the internal temporal profile of dynamically unfolding events and form event classes depending on the presence or absence of an inherent event boundary. This type of information is necessary for describing events in all natural languages and can provide the basis for further understanding and processing the dynamic nature of human experience.

[3] In our study, participants were presented with stimuli that were readily perceived in either bounded or unbounded terms. However, as mentioned already, the same visual input can often be construed as either a bounded or an unbounded event, depending on one's perspective (cf. *The girl is playing the piano* vs. *The girl is playing the Moonlight Sonata*). In future work, it is important to ask how the processing perspective of the viewer organizes any given temporally unfolding stream of sensory information in terms of the bounded/unbounded distinction.

Acknowledgements

The authors thank Raevyn Johnson for helping construct the videos, Audrey Sokol and Rebekah Mullin for helping with data collection, and Matthew Soule for transcribing description data. A short report of results from Experiments 1 and 2 appeared in Ji and Papafragou (2017).

Funding

This material is based upon work supported by a Doctoral Fellowship from the Office of Graduate and Professional Education of the University of Delaware (Y.J.) and National Science Foundation grant no. 1632849 (A.P.).

Declarations of interest: none.

Data: The data for the present study are available from osf.io/sbjqm Ji, Y., & Papafragou, A. (2019). Viewers' sensitivity to abstract event structure.

Appendix A

Table A.1

Fixed effect estimates for multi-level model of event category identification in Experiment 1

Effect	Estimate	SE	z value
(Intercept)	2.55	0.32	7.95***
Question Type	-1.92	0.42	-4.57***
Condition	-1.36	0.35	-3.84***
Question Type*Condition	1.59	0.49	3.24**

Note. Formula in R: Acc ~ 1 + (1|ID) + (1|Item) + Question Type + Condition + Question Type : Condition

* $p < .05$, ** $p < .01$, *** $p < .001$

Table A.2.

Fixed effect estimates for multi-level model of event category identification in Experiment 2

Effect	Estimate	SE	z value
(Intercept)	1.45	0.21	6.78***
Question Type	-0.71	0.31	-2.25*
Condition	-1.30	0.28	-4.72***
Question Type*Condition	1.02	0.42	2.42*

Note. Formula in R: Acc ~ 1 + (1|ID) + (1|Item) + Question Type + Condition + Question Type : Condition

* $p < .05$, ** $p < .01$, *** $p < .001$

Table A.3

Fixed effect estimates for multi-level model of event category identification in Experiment 4

Effect	Estimate	SE	z value
(Intercept)	1.55	0.21	7.45***
Question Type	-1.55	0.31	-5.08***
Condition	-0.82	0.27	-3.06**
Question Type* Condition	0.62	0.42	1.49

Note. Formula in R: Acc ~ 1 + (1|ID) + (1|Item) + Question Type + Condition + Question Type : Condition

* $p < .05$, ** $p < .01$, *** $p < .001$

References

Altmann, G. T. M., & Ekves, Z. (2019). Events as intersecting object histories: A new theory of event representation. *Psychological Review,* Advance online publication. doi: 10.1037/rev0000154

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi: 10.1016/j.jml.2007.12.005.

Bach, E. (1986). The algebra of events. *Linguistics and Philosophy, 9*, 5-16. doi: 10.1002/9780470758335.

Bar-El, L., Davis, H., & Matthewson, L. (2005). On non-culminating accomplishments. In L. Bateman, & C. Ussery (Eds.), *Proceedings of the 35th Annual Meeting of the North Eastern Linguistics Society*, Volume 1 (pp. 87-102). Amherst, MA: GLSA.

Barner, D., Wagner, L., & Snedeker, J. (2008). Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns. *Cognition, 106*, 805–832. doi: 10.1016/j.cognition.2007.05.001

Barr, D. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457–474. doi: 10.1016/j.jml.2007.09.002.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. United Kingdom: Cambridge University Press.

Borik, O. (2006). *Aspect and reference time*. Oxford: Oxford University Press.

Borer, H. (2005). *In name only: Structuring sense.* Vol. 1. *The normal course of events: Structuring sense*. Vol. 2. Oxford: Oxford University Press.

Botne, R. (2003). To die across languages: Towards a typology of achievement verbs. *Linguistic Typology, 7*, 233-278. doi: 10.1515/lity.2003.016.

Comrie, B. (1976). *Aspect*: *An introduction to the study of verb aspect and related problems.* Cambridge: CUP.

Cooper, R. P. (2019). Action production and event perception as routine sequential behaviors. *Topics in Cognitive Science*. doi: 10.1111/tops.12462. [Epub ahead of print]

Dahl, Ö. (1981). On the definition of the telic-atelic (bounded-unbounded) distinction. In P. J. Tedeschi & A. Zaenen (Eds.), *Syntax and semantics (Vol. 14)*: *Tense and aspect* (pp. 79-90). New York: Academic Press.

Dowty, D. (1979). *Word meaning and Montague grammar*. Dordrecht, The Netherlands: Kluwer.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*, 547-619. doi: 10.2307/415037

Elman, J., & McRae, K. (2019). A model of event knowledge. *Psychological Review*, *126*, 252-291. doi: 10.1037/rev0000133

Filip, H. (1993). *Aspect, situation types and nominal reference* (Unpublished doctoral dissertation). University of California at Berkeley, Berkeley, CA.

Filip, H. (2012). Lexical aspect. In R. I. Binnich (Ed.), *The Oxford handbook of tense and aspect* (pp. 721-751). Oxford: Oxford University Press.

Folli, R., & Harley, H., (2006). What language says about the psychology of events. *Trends in Cognitive Science*, *10*, 91-92. doi: 10.1016/j.tics.2006.01.002

Garey, H. B. (1957). Verbal aspects in French. *Language, 33,* 99-110.

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology, 142*, 880-905. doi: 10.1037/a0030045

Harley, H. (2003). How do verbs get their names? Denominal verbs, manner incorporation and the ontology of verb roots in English. In N. Erteschik-Shir, & T. Rapoport (Eds.), *The syntax of aspect: Deriving thematic and aspectual interpretation* (pp. 62-64). Oxford: Oxford University Press.

He, X., & Arunachalam, S. (2016, March). *How event endstates are conceptualized*. Poster presented at the 29th Annual CUNY Conference on Human Sentence Processing, Gainesville, FL. Abstract retrieved from http://tedlab.mit.edu/~mekline/ELC2016/25_ELC2016-He&Arunachalam.pdf

Hindy, N. C., Altmann, G. T. M., Kalenik, E., & Thompson-Schill, S. L. (2012). The effect of object state-changes on event processing: Do objects compete with themselves? *The Journal of Neuroscience, 32*, 5795–5803. doi:10.1523/JNEUROSCI.6294-11.2012

Hinrichs, E. (1985). *A Compositional Semantics for Aktionsarten and NP Reference in English* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.

Jackendoff, R. (1991). Parts and boundaries. *Cognition, 41*, 9-45. doi: 10.1016/0010-0277(91)90031-X

Jackendoff, R. (1996). The proper treatment of measuring out, telicity, and perhaps event quantification in English. *Natural Language and Linguistic Theory, 14*, 305-354.

Jachendoff, R. (2007). Shaking hands and making coffee: The structure of complex actions. In *Language, consciousness, culture: Essays on mental structure* (pp. 111-144). Cambridge, MA: MIT Press.

Ji, Y. & Papafragou, A. (2017). Viewers' sensitivity to abstract event structure. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 594-599). Austin, TX: Cognitive Science Society.

Konopka, A. E., & Brown-Schmidt, S. (2014). Message encoding. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 3-20). Oxford: Oxford University Press.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expression* (pp.75-115). Dordrecht: Foris Publications.

Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and grammar*. Dordrecht: Kluwer.

Lakusta, L., & Landau, B. (2005). Starting at the end: the importance of goals in spatial language. *Cognition*, *96*, 1-33. doi: 10.1016/j.cognition.2004.03.009

Lakusta, L., & Landau, B. (2012). Language and memory for motion events: Origins of the asymmetry between source and goal. *Cognitive Science*, *36*, 517-544. doi: 10.1111/j.1551-6709.2011.01220.x

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Levin, B. (2000). Aspect, lexical semantic representation, and argument expression. In A. K. Simpson (Ed.), *Proceedings of the 26th Annual Meeting of the Berkeley Linguistic Society* (pp. 413-429). Berkeley, CA: Berkeley Linguistics Society.

Maguire, M. J., Brumberg, J., Ennis, M., & Shipley, T. F. (2011). Similarities in object and event segmentation: A geometric approach to event path segmentation. *Spatial Cognition and Computation, 11*, 254-279. doi: 10.1080/13875868.2011.566955

Malaia, E. (2014). It still isn't over: event boundaries in language and perception. *Language and Linguistics Compass, 8,* 89-98. doi: 10.1111/lnc3.12071

Malaia, E. Renaweera, R., Wilbur, R., & Talavage, T. (2012). Event segmentation in a visual language: Neural bases of processing American Sign Language predicates. *Neuroimage, 19,* 4094-4101. doi:10.1016/j.neuroimage.2011.10.034.

Mereu, S., Zacks, J. M., Kurby, C. A., & Lleras, A. (2014). The role of prediction in perception: evidence from interrupted visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1372-1389. doi: 10.1037/a0036646

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, *34,* 1064-1092. doi: 10.1111/j.1551-6709.2010.01107.x

Papafragou, A. (2015). The representation of events in language and cognition. In E. Margolis & Laurence, S. (Eds.), *The conceptual mind: new directions in the study of concepts* (pp. 327-346). Cambridge, MA: MIT Press.

Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization in language production: Evidence from event encoding. *Language, Cognition and Neuroscience,* online pub. DOI: 10.1080/23273798.2019.1589540

Parsons, T. (1990). *Events in the semantics of English*: *A study in subatomic semantics*. Cambridge, MA: MIT Press.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

Prasada, S., Ferenz, K., & Haskell, T. (2002). Conceiving of entities as objects and as stuff. *Cognition, 83*, 141-165. doi: 10.1016/S0010-0277(01)00173-1

Pustejovsky, J. (1991). The syntax of event structure. *Cognition*, *41*, 47-81. doi: 10.1016/0010-0277(91)90032-Y

Radvansky, G., & Zacks, J. (2014). *Event cognition*. Oxford: Oxford University Press.

Regier, T., & Zheng, M. (2007). Attention to endpoints: A cross-linguistic constraint on spatial meaning. *Cognitive Science, 31,* 705-719. doi: 10.1080/15326900701399954

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.

Sakarias, M. & Flecken, M. (2018). Keeping the result in sight and mind: General cognitive principles and language-specific influences in the perception and memory of resultative events. *Cognitive Science, 43,* 1-30. doi: 10.1111/cogs.12708

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ:Erlbaum.

Sharon, T., & Wynn, K., (1998). Individuation of actions from continuous motion. *Psychological Science, 9*, 357-362. doi: 10.1111/1467-9280.00068

Shipley, T., & Zacks, J. (2008). *Understanding events: From perception to conception.* Oxford: Oxford University Press.

Smith, C. S. (1991). *The parameter of aspect*. Dordrecht, The Netherlands: Kluwer.

Solomon, S. H., Hindy, N. C., Altmann, G. T., & Thompson-Schill, S. L. (2015). Competition between mutually exclusive object states in event comprehension. *Journal of Cognitive Neuroscience, 27,* 2324–2338. doi: 10.1162/jocn_a_00866

Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M. & Pfau, R. (2015). Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences of the United States of America, 112,* 5968-73. doi: 10.1073/pnas.1423080112

Strickland, B., & Keil, F. (2011). Event completion: Event based inferences distorted memory in a matter of seconds. *Cognition, 121*, 409-415. doi: 10.1016/j.cognition.2011.04.007

Strickland, B., & Scholl, B. (2015). Visual perception involves event-type representations: The case of containment versus occlusion. *Journal of Experimental Psychology: General, 144*, 570-580. doi: 10.1037/a0037750

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

Tenny, C. (1987). *Grammaticalizing aspect and affectedness* (Unpublished doctoral dissertation). MIT, Cambridge, MA.

Tversky, B., Zacks, J., Morrison, J., & Hard, B. (2011). Talking about events. In J. Bohnemeyer, & E. Perdersen, (Eds), *Event representation in language and cognition* (pp. 216-227). Cambridge, UK: Cambridge University Press.

Van Hout, A. (2007). Acquiring telicity cross-linguistically: On the acquisition of telicity entailments associated with transitivity. In P. Brown, & M. Bowerman (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability* (pp. 255-278). Hillsdale: Routledge.

Van Hout, A. (2016). Lexical and grammatical aspect. In J. Lidz, W. Synder, & J. Pater (Eds.), *The Oxford handbook of developmental linguistics*. Oxford: Oxford University Press.

Van Hout, A. (2018). On the acquisition of event culmination. In K. Syrett, & S. Arunachalam (Eds.), *Semantics in Language Acquisition,* pp. 95-121. Amsterdam: John Benjamins Publishers.

Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning, and function.* Cambridge: Cambridge University Press.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review, 66*, 143-160. doi: 10.2307/2182371

Von Fintel, K., & Matthewson, L. (2008). Universals in semantics. *The Linguistic Review, 25*, 139-201. doi: 10.1515/TLIR.2008.004

Wagner, L. (2012). First Language Acquisition. In Binnick, R. (Ed.), *The Oxford Handbook of Tense and Aspect,* 458 - 480. Oxford: Oxford University Press.

Wagner, L., & Carey, S. (2003). Individuation of objects and events: A developmental study. *Cognition, 90,* 163–191. doi: 10.1016/S0010-0277(03)00143-4

Wellwood, A., Hespos, S. J., & Rips, L. (2018). The object : substance :: event : process analogy. In T. Lombrozo, J. Knobe, & S. Nichols (Eds.), *Oxford Studies in Experimental Philosophy*, Volume 2 (pp. 183-212). Oxford: Oxford University Press.

Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: numerical discrimination and its signature limits. *Developmental Science, 8,* 173-181. doi: 10.1111/j.1467-7687.2005.00404.x

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind/brain perspective. *Psychological Bulletin*, *133*, 273–293. doi:10.1037/0033-2909.133.2.273

Zacks, J., & Swallow, M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16*, 80-84. doi: 10.1111/j.1467-8721.2007.00480.x

Zacks, J., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3-21. doi: 10.1037/0033-2909.127.1.3