Midpoints, endpoints and the cognitive structure of events

Yue Ji & Anna Papafragou

LANGUAGE, COGNITION AND NEUROSCIENCE
2020, VOL. 35, NO. 10, 1465–1479
https://doi.org/10.1080/23273798.2020.1797839

Routledge
Taylor & Francis Group

REGULAR ARTICLE

Midpoints, endpoints and the cognitive structure of events

Yue Ji [a] and Anna Papafragou[b]

[a]Department of Linguistics and Cognitive Science, University of Delaware, Newark, USA.; [b]Department of Linguistics, University of Pennsylvania, Philadelphia, USA.

ABSTRACT
Events unfold over time, i.e. they have a beginning and endpoint. Previous studies have illustrated the importance of endpoints for event perception and memory. However, this work has only discussed events with a self-evident endpoint, and the internal temporal structure of events has not received much attention. In this study, we hypothesise that event cognition computes boundedness, an abstract feature of the internal temporal structure of events. We further hypothesise that sensitivity to boundedness affects how individual temporal slices of events (such as event midpoints or endpoints) are processed and integrated into a coherent event representation. The results of three experiments confirm these hypotheses. In Experiment 1, viewers identified the class of bounded (non-homogeneous, culminating) and unbounded (homogeneous, non-culminating) events in a categorisation task. In Experiments 2 and 3, viewers reacted differently to temporal disruptions in bounded versus unbounded events. We conclude that boundedness shapes how events are temporally processed.

ARTICLE HISTORY
Received 11 January 2020
Accepted 10 July 2020

KEYWORDS
Event; event cognition; aspect; boundedness

Introduction

The cognitive representation of events

Events unfold over time and can be defined as temporal segments with "a beginning and an end" (Zacks & Tversky, 2001). For instance, an event of climbing a mountain begins with people at the mountain foot and ends at the moment people reach the mountain top. Much work on event segmentation has explored how viewers determine the beginning and end of an event. Event Segmentation Theory (EST; Zacks et al., 2007) proposes that people form predictions about what is happening next based on the current event model; when important situation features change, the predictions are no longer accurate. As a result, people update the current event model and perceive the moment of accumulating change as an event boundary. The change can be perceptual, such as new spatial location, direction or speed of a motion (e.g. a train arriving at the station; Magliano et al., 2001; Newtson et al., 1977; Zacks et al., 2007), or conceptual, such as the achievement of a goal (e.g. a researcher arriving at a conclusion; Baldwin et al., 2001; Newtson, 1973; Vallacher & Wegner, 1987; Wilder, 1978; Zacks & Swallow, 2007; see also Papenmeier et al., 2019, for an alternative proposal to EST on which event boundaries may be identified through backward inferences, rather than predictive perception). A complementary framework known as the Event-Indexing model (Zwaan et al., 1995) posits that people construct and update a coherent event representation on the basis of several dimensions (including temporal, spatial, protagonist-related, causal and intentional aspects of the event).

Event boundaries have a privileged status in event comprehension and memory, and provide anchors for later learning and describing of events (e.g. Boltz, 1995; Huff et al., 2012; Newtson & Engquist, 1976; for an overview, see Radvansky & Zacks, 2017). For instance, objects present at an event boundary are retrieved more easily than objects present at any moment between event boundaries (Swallow et al., 2009). Furthermore, event endpoints appear to be particularly important compared to event beginnings. For instance, many studies on motion events have revealed a source-goal asymmetry: the goal of motion (i.e. the endpoint, as in *Felix walked to the bakery*) is more accurately encoded in both language and memory as opposed to the source (as in *Felix walked out of the room*; Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007; Wagner, 2009). Event endings also have an outsize influence on working memory: when observers track and estimate the number of overlapping motions, asynchronous endings but not asynchronous beginnings have an effect on estimation (Ongchoco & Scholl, 2019). Furthermore, the absence of an endpoint can trigger rapid reactions from event viewers. When the endpoint is not in

CONTACT Yue Ji Jiyue@Udel.Edu

sight, even infants as young as 12 months old can actively infer it from observed actions (Csibra et al., 2003). Relatedly, when the endpoint of a causally active event within an event chain remains unseen, adults form inferences about this event within seconds. For instance, after watching videos of someone launching an object (e.g. kicking a soccer ball) followed by the object's directed motion (e.g. the ball flying into a goal), participants mistakenly reported that they saw the moment of contact at launching, even when it was actually omitted from the display (Strickland & Keil, 2011). Furthermore, these inferences about event endings do not rely on prior experience or knowledge (e.g. in the soccer ball example, both novices and soccer referees inferred the moment of contact in similar ways; Brockhoff et al., 2016; see also Papenmeier et al., 2019).

Previous studies on event representation leave open two issues. A first issue is that, apart from beginnings and endpoints, other potentially salient temporal landmarks such as midpoints have not received much attention. To our knowledge, only one recent study has compared event midpoints to both event starting points and endpoints (Gold et al., 2017). In this study, participants watched movies of everyday activities (e.g. setting up for a party) which were composed of a series of meaningful events (e.g. taking plates and napkins out of a bag, laying the table, etc.). Some of the movies were edited by placing cues (a bell sound along with an arrow pointing to the affected object) either at event midpoints or at event boundaries. It was found that cues at event midpoints improved subsequent memory of movies, although the cues were less effective than cues at event boundaries. These findings suggest that event midpoints, although informative to some extent, are less salient compared to event starting points and endpoints.

The lack of attention to potential temporal landmarks such as midpoints within an event reflects a second issue: despite the fact that most prior work has only discussed events with a self-evident endpoint (such as the destination of a motion event), event boundaries and other temporal landmarks such as midpoints often need to be defined abstractly. More generally, a proper theoretical understanding of how midpoints, endpoints and other temporal landmarks contribute to event representation needs to take into account the specific temporal profile of individual events. However, until recently, not much research had looked into the abstract internal structure of events (but see Huff & Papenmeier, 2017).

In this paper, we propose to address both of these issues by arguing that the abstract property of boundedness structures the psychological representation of events and affects how individual temporal slices of events (such as event midpoints or endpoints) are processed and integrated into a coherent event representation.

Boundedness and event structure

Inspired by the extensive linguistic literature on situation aspect (i.e. the linguistic encoding of the temporal contour of events; see Bach, 1986; Jackendoff, 1991; Vendler, 1957), we distinguish between two types of events with different internal structures and different ways in which they come to an end. *Bounded* events such as the situation described by (1) below are developments leading to a "built-in terminal point" (Comrie, 1976), "climax" (Vendler, 1957) or "culmination" (Parsons, 1990). Bounded events include an inherent endpoint, "which is there from the outset and culminates if not interrupted" (Mittwoch, 2013). In (1), the endpoint is the moment when the teddy bear got dressed. *Unbounded* events such as the situation described by (2) have a homogenous structure and may terminate at any arbitrary moment. In (2), the endpoint could be any moment when the girl stopped patting.

(1) A girl dressed a teddy bear.
(2) A girl patted a teddy bear.

Two properties, homogeneity (Hinrichs, 1985) and cumulativity (Krifka, 1989, 1998; Taylor, 1977), have been used to characterise the differences in the temporal structure between bounded and unbounded events. Bounded events are non-homogeneous: if we divide a bounded event into temporal slices with minimal duration, each slice represents a different stage of development. In (1), each temporal slice of dressing a teddy bear could be a distinct step (e.g. putting the pants on one leg at a time, fastening a button, etc.). Bounded events are also non-cumulative: events of the same kind cannot be simply added up. For instance, if there is an event of a girl dressing a teddy bear between 9 am and 10 am, and another event of teddy-bear dressing by the same girl between 10 am and 11 am, the sum of these two events can no longer be described by (1). Instead, it results in a new event that can be described as "A girl dressed two teddy bears." By contrast, unbounded events are homogeneous: we can divide an unbounded event into any number of temporal slices and each slice can still be regarded as an event of the same kind. In (2), each moment of the girl's action is still an event of patting a teddy bear. Unbounded events are also cumulative, i.e. events of the same kind can be added together. If an event of a girl patting a teddy bear occurs from 9 am to 10 am, and another event of her patting from 10 am to 11 am, the sum of these two events leads to the same event of patting a teddy bear.

The difference in homogeneity and cumulativity makes the bounded-unbounded distinction in the event domain reminiscent of the object-substance distinction in the object domain. Bounded events resemble atomic, structured objects such as a sandcastle. Similar to objects which specify number as the default dimension of measurement (Barner & Snedeker, 2005), bounded events are quantified by counting how many endpoints have been achieved. Unbounded events resemble non-atomic substances such as sand. Substances are not specified in terms of dimensions of measurement (e.g. sand may be measured by volume, weight, etc.). Similarly, unbounded events are quantified according to arbitrary spatio-temporal criteria (e.g. an event of patting a teddy bear may be measured by its duration, as well as frequency; cf. Bach, 1986 for linguistic analyses).

Several studies have offered evidence that boundedness can be extracted from communicative signals (e.g. Malaia et al., 2012; Strickland et al., 2015), and have used it to explain mappings between the cognitive and linguistic representation of events (Barner et al., 2008; Van Hout, 2007, 2016, 2018; Wagner, 2012; Wagner & Carey, 2003; Wellwood et al., 2018). For present purposes, a study of particular relevance found that viewers can form categories corresponding to bounded and unbounded events after exposure to video clips instantiating the two categories (Ji & Papafragou, 2020, in press). Furthermore, viewers are better at identifying the category of bounded events (ibid.), presumably because of the presence of defined endpoints that make such events "atomic" and easier to compare and generalise over. Boundedness distinctions in this paradigm have been shown to differ from completion meanings (e.g. whether the endpoint is realised), and persist even when participants perform the categorisation task while engaged in a secondary, linguistic task (Ji & Papafragou, 2020, in press). The fact that viewers extract the feature of boundedness even when overt linguistic encoding is blocked reveals that event perception is sensitive to abstract boundedness considerations independent of whether events are encoded linguistically online. This result strongly raises the expectation that the psychological representation of event boundedness has broader, testable implications about how people track the internal temporal structure of events, including temporal landmarks such as midpoints and event boundaries. In the present study, we tested this expectation.

In Experiment 1, we began by confirming that viewers are sensitive to differences in (un)boundedness when representing events in a paradigm that builds on prior work (Ji & Papafragou, 2020, in press). In Experiments 2 and 3, we pursued the hypothesis that sensitivity to (un)boundedness should affect how viewers process and weigh temporal slices of different events. We focussed on two specific predictions that flow from this hypothesis. First, since bounded events have a finely differentiated internal structure that is defined on the basis of the availability of an inherent endpoint, we expected that endpoints should be particularly salient over other event slices such as midpoints for such events. By contrast, since unbounded events have a largely undifferentiated, homogeneous internal structure, endpoints should be treated largely similarly to midpoints or other points within the event's temporal profile. To test this prediction, in Experiment 2 we introduced interruptions at different time points in the temporal profile of bounded vs. unbounded events and measured sensitivity to these interruptions as viewers learned a novel way of categorising the stimuli.

Second, the finely differentiated temporal structure of bounded events should be more vulnerable to disruptions of temporal sequence compared to the homogeneous temporal structure of unbounded events. To test this prediction, in Experiment 3, we introduced global violations of the temporal contours of events and compared the effects of such violations on the apprehension of bounded vs. unbounded events.

Experiment 1

Experiment 1 was a category identification task (adapted from Ji & Papafragou, 2020, in press). Participants were exposed to pairs of bounded and unbounded events (separated by a few other videos) and had to extract a generalisation about one member of these pairs. The purpose of the study was to ascertain the psychological reality of (un)boundedness distinctions before we turned to the implications of this distinction for the way events are cognitively processed.

Method

Participants

Forty undergraduate students (18 female, 22 male; M_{age} = 19.6, age range: 18.1-22.2) at the University of Delaware participated. Data from an additional adult were collected but excluded because this participant failed to understand the task and did not finish all the questions. Our sample size in this and subsequent experiments was based on similar work on event perception and event categorisation (Mereu et al., 2014; Strickland & Keil, 2011; Strickland & Scholl, 2015).

Stimuli

Sixteen videos of bounded events and sixteen videos of closely related unbounded events were created (see Table 1). Related bounded-unbounded videos (i.e. each row in Table 1) had the same duration (duration: 4-12s, M = 6.9s) and involved the same actor wearing a blue shirt. All of the videos began with the actor picking up an object or tool to perform an action, and came to an end with the actor putting down the object or tool (see Figure 1). 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 Filip, 2004; Tenny, 1987; cf. Ji & Papafragou, 2020, in press). We 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 videos, the bounded-unbounded events involved the same object but differed in terms of the nature of the action performed on the object: the bounded event displayed an action that caused a clear and temporally demarcated change of state in the object (e.g. stack a deck of cards) while its unbounded counterpart did not involve such a change (e.g. shuffle a deck of cards; see Figure 1a-b). For the other half of the videos, the bounded and unbounded events involved the same action but differed in terms of the nature of the affected object: the bounded event involved a single object (e.g. blow a balloon) but its unbounded counterpart involved either an unspecified plurality of objects or a mass quantity (e.g. blow bubbles; see Figure 1c-d).

We conducted four types of norming studies for these stimuli. First, to ensure that all video stimuli would instantiate the contrast in boundedness, a separate group of 18 adults from the same population was asked to watch a subset of the clips and describe what happened in a full English sentence. For this task, the events in Table 1 were split into two lists, such that each list

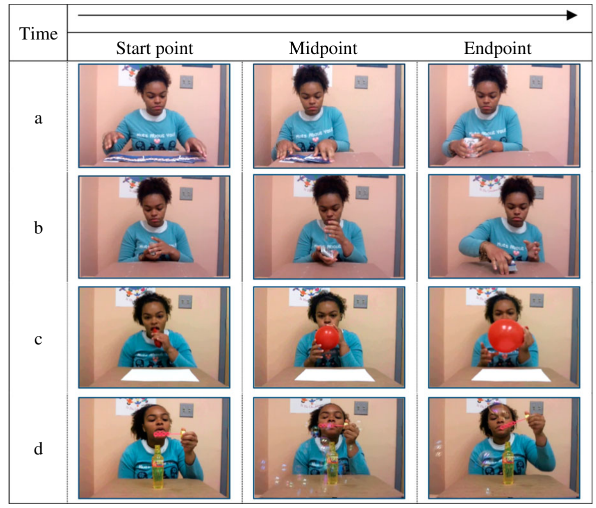

Figure 1. Examples of closely related bounded and unbounded events in Experiment 1: (a) stack a deck of cards (bounded) vs. (b) shuffle a deck of cards (unbounded); and (c) blow a balloon (bounded) vs. (d) blow bubbles (unbounded)

included only one video in each row of Table 1 and an equal number of bounded and unbounded events. 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(s). 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) by 4 English native speakers. Their judgment agreed 94.6% of the time and discrepancies were resolved through discussion. As expected, stimuli of bounded events elicited bounded descriptions that included change-of-state verbs or particle verbs (e.g. *stack a deck of cards*) or quantified count noun phrases (e.g. *blow a balloon*) 96.7% of the time. Stimuli of

Table 1. Event stimuli in Experiment 1. Related videos appear in the same row. In the training phase, participants saw both events within a pair. In the testing phase, participants saw only one event from each pair.

Phase	No.	Bounded Events	No.	Unbounded Events	Boundedness Source
Training	1	fold up a handkerchief	17	wave a handkerchief	Nature of Action
	2	put up one's hair	18	scratch one's hair	
	3	stack a deck of cards	19	shuffle a deck of cards	
	4	group pawns based on colour	20	mix pawns of two colours	
	5	draw a balloon	21	draw circles	Nature of Affected Object
	6	tie a knot	22	tie knots	
	7	eat a pretzel	23	eat cheerios	
	8	flip a postcard	24	flip pages	
Testing	9	dress a teddy bear	25	pat a teddy bear	Nature of Action
	10	roll up a towel	26	twist a towel	
	11	fill a glass with milk	27	shake a bottle of milk	
	12	scoop up yogurt	28	stir yogurt	
	13	peel a banana	29	crack peanuts	Nature of Affected Object
	14	blow a balloon	30	blow bubbles	
	15	tear a paper towel	31	tear slices off paper towels	
	16	paint a star	32	paint stuff	

unbounded events elicited unbounded verb phrases that included verbs of activity (e.g. *shuffle a deck of cards*) or unquantified noun phrases (bare plurals or mass nouns: e.g. *blow bubbles*) 91.2% of the time.

Second, to make sure that people would view the video stimuli as either bounded or unbounded events, we asked a new group of 40 participants to answer the following question after watching each video clip: “Does it make sense to think of the action in the video as something with a beginning, midpoint and specific endpoint?”. As in the linguistic norming study, the events in Table 1 were split into two lists, such that each list included only one member of each row. The 40 participants were randomly assigned to one of the two lists. As expected, the bounded stimuli elicited *Yes* responses 87.2% of the time while the unbounded stimuli did so only 20.3% of the time ($t(1, 39) = 20.05$, $p < .001$).

Third, to control for the role of agent’s intention on event perception, 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 = 5.61) and unbounded events (M = 5.48) (t $(19) = 1.14$, $p = .210$). Finally, we 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 $(19) = -1.43$, $p = .169$).

For the main experiment, the video stimuli were arranged into two basic sets corresponding to the two 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 such that any 2 videos within a pair were separated by at least 5 other videos (so that the contrast between the bounded and unbounded versions would not be highlighted). Within the list, a single video was played in the centre of the screen and was followed immediately by the next video. The order of bounded-unbounded events within pairs was counterbalanced within the list.

For the 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. Participants were randomly assigned to one of the 2 lists.

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?”

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.

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 modelling with crossed intercepts for Subjects and Items (Baayen et al., 2008; Barr, 2008). All models were fitted with the *glmertree* or *lme4* package in R. The binary accuracy data were submitted to a logit model with Condition

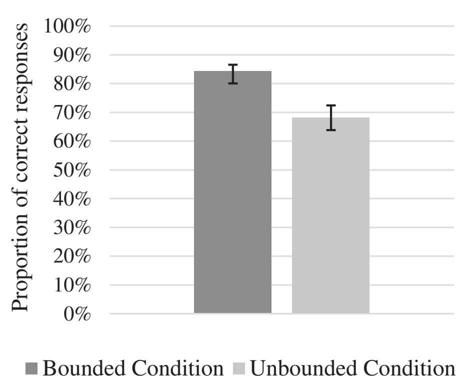

Figure 2. Proportion of correct responses in Experiment 1. Error bars represent ± SEM.

(Bounded vs. Unbounded) as the fixed factor using centred contrast coding (−0.5, 0.5). The same coding strategy was followed in all subsequent analyses. 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 the final model. As shown in Table 2, Condition had a significant effect such that participants gave more correct responses in the Bounded condition (M = 84.4%) than in the Unbounded condition (M = 68.1%) ($\beta = -0.99$, $z = -3.35$, $p < .001$). Participants performed at levels different from chance for both Conditions (both *ps* < .05).

Organisation, neatness and intention (or lack thereof) appeared in 40%, 20% and 15% respectively of the answers to the open question about the meaning of the red frame. Additionally, completion was used in 15% of the answers (exclusively for bounded events) and repetition was used in 10% of the answers (exclusively for unbounded events).

Discussion

Results from Experiment 1 replicate prior findings (Ji & Papafragou, 2020, in press) with a slightly modified design. Participants succeeded in identifying the categories of both boundedness and unboundedness. Furthermore, there was an advantage for forming the category of bounded compared to unbounded events. Participants' overt conjectures rule out alternative bases for their responses: they show, for instance, that responses were not driven primarily by whether the events were completed, intentional, or could be described in certain ways (see also Ji & Papafragou, 2020, in press). We conclude that participants represent event classes such as those of (un)bounded events in non-linguistic cognition.

Experiment 2

In Experiment 2, we asked whether the sensitivity to bounded and unbounded structure shown in Experiment 1 would affect how viewers process and weigh temporal slices of events. We adopted a variant of the "picky puppet task" (Waxman & Gelman, 1986): we presented people with videos containing an interruption that blocked either the endpoint or the midpoint of an event and told them that a picky girl liked only one of these kinds of videos (see also Papenmeier et al., 2019). We later asked people to guess whether the girl would like new videos containing mid- or end-interruptions. We hypothesised that people's own biases when processing events of different types would interact with the ability to learn someone else's (ostensibly arbitrary) preference towards disrupted event stimuli. Specifically, we expected that, for bounded events where an internal structure leads to a climax, people would treat the preference for interruptions during temporal endpoints vs. midpoints differently (and would have greater difficulty learning that someone else prefers endpoint interruptions). For unbounded events that have a largely undifferentiated internal structure, however, the endpoint-midpoint difference should diminish or disappear altogether.

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

Effect	Estimate	SE	z value
(Intercept)	1.82	0.32	5.66***
Condition	−0.99	0.30	−3.35***

Note. Formula in R: AccTest ~ 1 + (1|ID) + (1|Item) + Condition
* $p < .05$, ** $p < .01$, *** $p < .001$

Method

Participants

One hundred and sixty adults (female 99, male 61; M_{age} = 20.1, age range: 18.0-22.1) participated in the experiment. All were undergraduates at the University of Delaware and received course credit for participation. Data from an additional group of 8 adults were collected but excluded due to experimenter error.

Stimuli

The same stimuli as Experiment 1 were used. To increase the visual variety of the stimuli, we created an additional version of the videos that was identical to the first except that the actor wore a yellow shirt.

All of the videos were then edited in Corel VideoStudio X9 to introduce an interruption taking up one-fifth of the total video duration (range: 0.8-2.4s). During the interruption, a blurry picture appeared on the screen. To create the blurry picture, we selected the midpoint frame (for interruptions in the middle) or endpoint frame (for interruptions at the end) in the original video and then applied an Iris Blur Effect in Adobe Photoshop CS 6 (see the examples in Figure 3). Each video was edited twice, once to create a mid-interruption and once to create an end-interruption. The mid-interruption was centred on the video midpoint, while the end-interruption blocked the last 20% of the video.

Procedure

Participants were invited to watch some videos. The experimenter told participants that the girl in the

Figure 3. Examples of edited videos in Experiment 2 (a second version showed the actor in a yellow shirt): (a) stack a deck of cards (bounded) – mid-interruption; (b) stack a deck of cards (bounded) – end-interruption; (c) shuffle a deck of cards (unbounded) – mid-interruption; (d) shuffle a deck of cards (unbounded) – end-interruption.

videos liked performing, but was very picky about her videos: she liked some videos but not others. The task was to figure out what kind of videos the picky girl liked. Participants were randomly assigned to one of two conditions depending on the event type (Bounded or Unbounded) that they were exposed to throughout the experiment. In the training phase, participants watched a total of 16 videos. These were comprised of 8 events (No. 1–8 in Table 1 for bounded events, or 17–24 for unbounded events, presented in a random order), each with two versions shown in succession in the centre of the screen. The two versions differed in terms of both the actor's shirt colour (blue vs. yellow) and in terms of the placement of the interruption (mid-interruption vs. end-interruption, see Figure 4 for examples). Within this phase, half of the time mid-interruptions occurred when the actor wore a blue shirt and the other half they occurred when the actor wore a yellow shirt. Even though our hypothesis targeted detection of a mid- vs. end-interruption, we added the change of shirt colour to ensure that participants would treat the two (highly similar) versions of each event as different tokens. The order of mid-interruptions and end-interruptions, as well as shirt colours, was counterbalanced.

After each version the experimenter said either, "The girl likes the video", or "The girl doesn't like the video". Within the Bounded and Unbounded group, participants were randomly assigned to one of two Preference conditions. In the "Likes mid-interruption" condition, they were always told that the picky girl liked the video

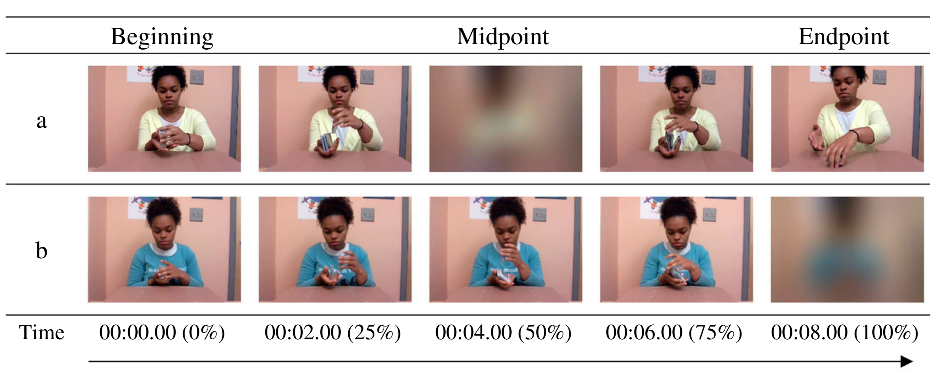

Figure 4. Example of a training trial for the unbounded event in Figure 3 (shuffle a deck of cards) that includes the two versions of the event: (a) mid-interruption (actor in yellow shirt), (b) end-interruption (actor in blue shirt).

after mid-interruptions but did not like the video after end-interruptions. In the "Likes end-interruption" condition, the girl's preference was reversed. Participants were expected to detect the girl's preference based on the placement of interruptions.

In the testing phase, participants watched a total of 8 videos corresponding to 8 new events (No. 9–16 in Table 1 for bounded events, or 25–32 for unbounded events, presented in random order). Half of these events were presented in their mid-interruption version and the other half in their end-interruption version. Each event type (bounded vs. unbounded) had 4 lists; participants were randomly assigned to one of the 4 lists. Each list included one interruption version (mid- or end-) of each event; the actor's shirt colour was counterbalanced for that event across lists. Types of interruptions and shirt colour changes were evenly split within each list. After watching each video, participants were asked: "Will the girl like this video or not?" They were requested to give a *Yes/No* response on an answer sheet.

At the end of the session, participants were asked to write down what kind of videos the girl liked. The answers were used as an additional source of information about whether participants related the girl's preference to a pattern of interruptions.

Results

Correct responses indicated successful identification of the type of interruptions the girl liked. For instance, in "Likes mid-interruption" condition, the correct answer was either a *Yes* response for a video with a mid-interruption or a *No* response for a video with an end-interruption. If participants based their conjecture on the actor's shirt colour (blue or yellow), they would give a correct response around 50% of the time since shirt colour was counterbalanced. As shown in Figure 5, the proportion of correct responses significantly differed from 50% in both Preference conditions and across

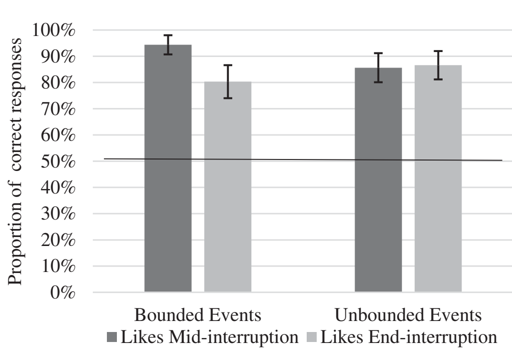

Figure 5. Proportion of correct responses in Experiment 2. Error bars represent ± SEM.

both Event types (all *ps* < .001). This suggests that participants overall were not misled by the change in shirt colour but instead tracked the placement of interruptions as expected.

The binary accuracy data were submitted to a mixed logit model examining the fixed effects of Preference (Likes mid-interruption vs. Likes end-interruption), Event Type (Bounded vs. Unbounded) and their interaction. We also examined whether shirt colour (Blue vs. Yellow), contrast (Action vs. Affect Object), placement of the interruption (Mid vs. End), as well as List would influence accuracy. Adding these factors to the model did not reliably improve the fit. Therefore, these factors were excluded from further analysis. Random intercepts were provided for each Subject and each Item (Baayen et al., 2008; Barr, 2008). As shown in Table 3, the model revealed an effect of Preference ($\beta = -1.43$, $z = -4.77$, $p < .001$), i.e. participants were more accurate when the picky girl liked mid-interruptions (M = 94.4%) compared to end-interruptions (M = 83.4%). There was also a significant effect of Event Type ($\beta = -1.05$, $z = -3.31$, $p < .001$), i.e. participants were more accurate when watching videos of bounded events (M = 87.3%) compared to unbounded events (M = 86.1%). Finally, a significant interaction between Preference and Event Type was detected ($\beta = 1.52$, $z = 3.86$, $p < .001$). When participants watched videos of bounded events, they were better at identifying a preference for mid-interruptions (M = 94.4%) compared to a preference for end-interruptions (M = 80.3%) ($\beta = -1.47$, $z = -4.28$, $p < .001$). By contrast, when participants watched videos of unbounded events, they were equally good at identifying that the girl liked mid-interruptions (M = 85.6%) or end-interruptions (M = 86.6%) ($\beta = 0.08$, $z = 0.34$, $p > .250$).

Recall that, at the end of the session, participants were asked to write down what kind of videos the girl liked. These responses are summarised in Table 4. As expected, the majority of participants mentioned the placement or effects of interruptions (e.g. "the girl likes when the video pauses in the middle and then resumes" in "Likes mid-interruption" condition; "the girl likes the videos where everything is masked at the end", "the girl likes the videos where she wasn't done with the task yet" in

Table 3. Fixed effect estimates for multi-level model of identifying the preference for interruptions in Experiment 2.

Effect	Estimate	SE	z value
(Intercept)	2.91	0.26	11.13***
Preference (Likes Mid- vs. End-interruption)	−1.43	0.30	−4.77***
Event Type (Bounded vs. Unbounded)	−1.05	0.31	−3.31***
Preference * Event Type	1.52	0.39	3.86***

Note. Formula in R: Acc ~ 1 + (1|Subject) + (1|Item) + Preference + Event Type + Preference: Event Type
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4. Counts (N) and percentages (%) of the two types of responses to the open question in Experiment 2.

Event Type	Preference	Interruption		Shirt colour	
		N	%	N	%
Bounded Bounded	Likes Mid-interruption	36/40	90%	4/40	10%
	Likes End-interruption	29/40	72.5%	11/40	27.5%
Unbounded Unbounded	Likes Mid-interruption	37/40	92.5%	3/40	7.5%
	Likes End-interruption	36/40	90%	4/40	10%

"Likes end-interruption" condition). A small number of participants responded on the basis of shirt colour (e.g. "The girl seems to alternate on the basis of the colour of the shirt"). As shown in Table 4, these conjectures were most frequent among participants exposed to bounded events in the "Likes end-interruption" condition. This fact confirms that it was harder to identify and accept a preference for missing endpoints, especially for bounded events.

Discussion

In Experiment 2, we tested the hypothesis that the abstract internal profile of events (or boundedness) that was found to be psychologically real in Experiment 1 should have implications for how individual time points are processed within events. This hypothesis predicts that midpoints and endpoints should behave asymmetrically in the representation of bounded events (since such events have non-homogeneous internal development and their endpoint coincides with the event's culmination) but should have comparable roles in the representation of unbounded events (since such events have a homogeneous structure and no culmination). In support of this prediction, our data show that participants had more difficulty tracking that someone might like end- compared to mid-interruptions of bounded events but there was no such difference for unbounded events. This result extends, explains and places boundaries on a robust finding from prior studies on event cognition, according to which event boundaries – especially event endpoints – are salient within the representation of an event (see, e.g. Lakusta & Landau, 2005, 2012; Papafragou, 2010; Strickland & Keil, 2011; Zacks & Swallow, 2007): here we find that the relative salience of endpoints in event cognition is tied to the internal temporal architecture of events.

Experiment 3

In Experiment 3, we tested a second prediction of the hypothesis that an abstract event property such as boundedness affects how the event timeline is processed. According to this prediction, a global disruption of the temporal development of an event should be more likely to affect event apprehension for bounded events (where there is non-homogeneous temporal structure leading to culmination) compared to unbounded events (where temporal structure is homogeneous). To test this prediction, we examined viewers' ratings of event videos after we either switched the order of the first and the second half of a video or showed the video in reverse. We know that such global disruptions affect event apprehension (e.g. Strickland & Keil, 2011) but are not fatal. For instance, in a recent study on event segmentation (Levine et al., 2017), viewers identified event boundaries (e.g. the point at which a skater finished spinning during a skating performance) in similar ways regardless of whether the stimuli were played forward or backward (cf. also Hard et al., 2006; Hard et al., 2011). Assuming that viewers would be able to extract pertinent event information from such deficient stimuli, we wanted to see whether their ratings would differ for bounded vs. unbounded events. We introduced two disruptions to the stimulus flow by either playing each video backwards or switching the order of the two video halves (presumably a milder change). We expected both of these disruptions (and especially video reversals) to be more likely to be judged as disruptive to event processing for bounded compared to unbounded events.

Method

Participants

Eighty eight adults (female 47, male 41; $M_{age} = 20.3$, age range: 18.3-22.9) participated in the experiment. All were undergraduates at the University of Delaware and received course credit for participation. Data from two additional adults were collected but excluded because they misunderstood the task.

Stimuli

Four videos of bounded events and four videos of unbounded events were created for the practice phase of this experiment (duration: 5-11s, M = 7.5s; see Table 5). The 32 unedited videos from Experiment 1 were used in the testing phase (see Table 1). To ensure the representativeness of the newly created stimuli, we did the same norming studies as in

Table 5. Video stimuli for the practice phase of Experiment 3.

Boundedness	No.	Events
Bounded	1	stack the 5 cups on the table
	2	put 20 Q-tips together
	3	erase a star
	4	organize 3 pairs of socks by colour
Unbounded	5	roll a ball back and forth
	6	grind biscuits
	7	sprinkle pepper
	8	pull a towel

Experiment 1. For the first of these, 18 new adults were asked to describe each of the 8 new videos in a full English sentence. The verb phrases in the descriptions were tested for boundedness. As expected, 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 an intentionality judgment task as in Experiment 1. There was no significant difference between the rating of members of the bounded ($M = 5.79$) and unbounded events class ($M = 5.60$) ($t(19) = 1.46$, $p = .190$).

The 40 videos were edited in Corel VideoStudio X9 to create two new versions. In the "switched halves" version, the first half and the second half of each video were switched (see Figure 6a and c). In the "reverse" version, each video was reversed (see Figure 6b and d).

These new video stimuli were arranged into 4 lists. Two lists had 20 bounded events and two lists had 20 unbounded events. Each list contained 10 videos with switched halves and 10 videos that were reversed. The "switched halves" and the "reverse" version were counterbalanced across lists. Participants were randomly assigned to one of the 4 lists.

Procedure

Participants were randomly assigned to one of two conditions depending on the Event type (Bounded or Unbounded) that they were exposed to throughout the experiment. The experiment started with a practice phase of 4 trials where participants watched two videos with switched halves and two videos in reverse. They were told that similar videos would be presented in the testing phase. The practice phase was aimed to familiarise participants with edited videos. After practice, participants were asked to watch one video at a time and decide how confusing the video looked on a scale from 1 (not confusing) to 7 (extremely confusing). On an answer sheet, they circled a number on the scale.

Results

Results from Experiment 3 are shown in Figure 7. Participants' rating of the videos were analyzed using multilevel mixed modelling with Event Type (Bounded vs. Unbounded) and Video Type (Switched halves and Reverse) as the fixed factors and crossed random intercepts for Subjects and Items. As in previous experiments, we also examined the effect of Contrast (Action vs. Affected Object). However, neither this factor alone nor its interaction with other factors improved model fit. Thus, it was not included in the final model. As shown

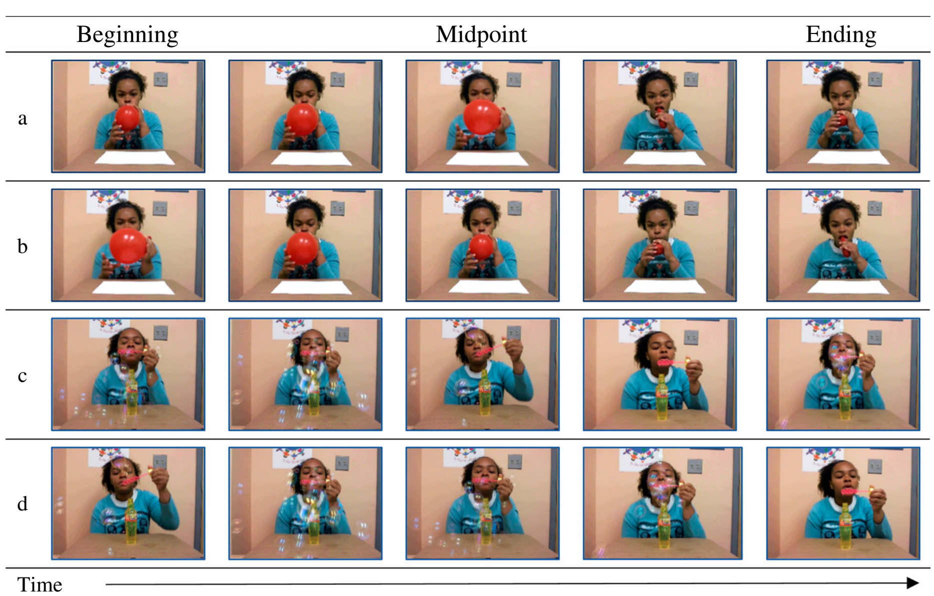

Figure 6. Examples of edited videos in Experiment 3: (a) blow a balloon (bounded) – switched halves; (b) blow a balloon (bounded) – reverse; (c) blow bubbles (unbounded) – switched halves; (d) blow bubbles (unbounded) – reverse.

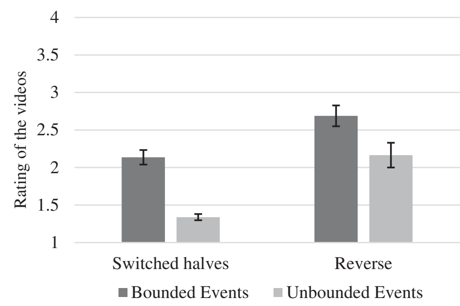

Figure 7. Average rating of the videos in Experiment 3 (on a 7-point scale). High ratings indicate highly confusing videos. Error bars represent ± SEM.

in Table 6, there was a fixed effect of Event Type ($\beta = -0.66$, $t = -5.42$, $p <. 001$), such that participants rated videos of bounded events as more confusing than videos of unbounded events, and of Video Type ($\beta = 0.69$, $t = 11.21$, $p < .001$), such that reverse videos were rated as more confusing than videos with switched halves. There was also a significant interaction between Event Type and Video Type ($\beta = 0.28$, $t = 2.24$, $p = .026$), i.e. the difference between the two conditions was greater in the "switched halves" videos than in the reverse videos.

Discussion

In this experiment, we disrupted the global temporal sequence within an event and asked whether these disruptions would be assessed differently for bounded and unbounded events. Even though these disruptions were not particularly troublesome overall (mean ratings stayed below the midpoint of a 7-point scale), they resulted in more confusion when they occurred within bounded compared to unbounded events; furthermore, the more extreme the disruption (as in video reversals), the more likely it was to affect bounded events disproportionately. These results provide empirical evidence for the conclusion that event cognition tracks boundedness (i.e. the abstract internal temporal profile of events) and uses such information to form a coherent event representation.

Table 6. Fixed effect estimates for multi-level model of video rating in Experiment 3.

Effect	Estimate	SE	t value
(intercept)	2.08	0.10	21.07***
Event Type (Bounded vs. Unbounded)	−0.66	0.12	−5.42***
Video Type (Switched halves vs. Reverse)	0.69	0.06	11.21***
Event Type*Video Type	0.28	0.12	2.24*

Note. Formula in R: Rating ~ 1 + Event Type + Video Type + Event Type: Video Type + (1 | ID) + (1 | Item)
* $p < .05$, ** $p < .01$, *** $p < .001$

General discussion

Midpoints, endpoints and the cognitive structure of events

Prior literature on event cognition has defined events as temporal segments with beginnings and endpoints (e.g. Zacks et al., 2007) and has revealed asymmetries between event processing at event boundaries compared to other temporal slices (such as midpoints, e.g. Gold et al., 2017; Newtson & Engquist, 1976; Swallow et al., 2009). Nevertheless, notions such as event midpoints and endpoints have remained hard to define, and have until recently not been integrated into a broader theory of events and event types.

The present study began by offering empirical support for the hypothesis that people are sensitive to differences in the internal temporal profile of events, i.e. they distinguish between bounded (non-homogeneous) and unbounded (homogeneous) events when they assemble representations of an event (Experiment 1). This hypothesis was consistent with linguistic treatments of aspect or telicity (e.g. Bach, 1986; Krifka, 1989, 1998) and prior experiments from non-linguistic categorisation tasks (Ji & Papafragou, 2020, in press; cf. Strickland et al., 2015). In two further experiments we provided support for two predictions of this hypothesis about the way people process information along the timeline of dynamically unfolding stimuli. In Experiment 2, we found that people represent temporal slices such as midpoints and endpoints differently depending on the perceived nature of the event (bounded vs. unbounded): interruptions that blocked either event midpoints or endpoints were treated identically for unbounded events but asymmetrically for bounded ones. In Experiment 3, we found that violations of temporal sequence within an event are tolerated to a different degree for bounded compared to unbounded events: disruptions of the temporal sequence within bounded events were judged as more confusing compared to similar disruptions within unbounded events.

Together, these results bear on several strands of research on event cognition. At the broadest level, they confirm the conclusion that people attend to the property of boundedness (i.e. the internal temporal contour of events) when assembling representations of dynamically unfolding experience. Thus the present data contribute to growing evidence that the units of event cognition are themselves quite abstract (Ji & Papafragou, 2020, in press; Strickland et al., 2015; see also Hard et al., 2011; Hard et al., 2019; Loucks et al., 2017). Furthermore, by showing that the notion of boundedness that was inspired by treatments of linguistic meaning extends to properties of events in conceptual structure, our data indirectly support the presence of parallels between event language and event perception (see also Brockhoff et al., 2016; Folli & Harley, 2006; Hafri et al., 2013; Lakusta & Landau, 2005; Malaia, 2014; Papafragou, 2015; Strickland et al., 2015; Tversky et al., 2011).

More relevantly for present purposes, our findings have implications about how event cognition filters the temporal dynamics of experience in the process of building coherent event representations of different types. Most obviously, our data provide new evidence for the importance of event endpoints, extending and enriching previous literature that compared endpoints with starting points in motion events (Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007; Wagner, 2009), and other work comparing event boundaries to intermediate points in an event timeline (Gold et al., 2017). Our results connect the salience of endpoints over other event slices to a broader framework concerning the fundamentals of event structure. Within this framework, endpoints carry different weight depending on the type of event that is being perceived: for bounded events that have internal structure, endpoints are privileged over other temporal slices such as midpoints but not so for unbounded events that lack internal structure. Furthermore, notions such as endpoints and midpoints are embedded into a broader event architecture that tracks properties such as homogeneity and can be disrupted in different ways depending on the profile of the perceived event.

Furthermore, our data bear on Event Segmentation theory (e.g. Zacks et al., 2007; Zacks & Swallow, 2007) that offers a mechanism for chunking dynamic input from experience into discrete event tokens. Recall that, on this theory, people predict boundaries of events by tracking changes in the input and place event boundaries at moments of maximum, highly unpredictable change. The present data are consistent with the major tenets of this theory, since the differences between bounded and unbounded events can be related to the predictability of the internal temporal development of an event. For unbounded events with a homogeneous structure (e.g. shuffling a deck of cards), internal development can be predicted from the very beginning. By contrast, for bounded events with a non-homogeneous structure (e.g. stacking a deck of cards), it is harder to make predictions about upcoming development because such events are composed of several steps. Put differently, the endpoints in bounded events, when maximum changes occur (e.g. the cards become a stack), are highly salient to observers and can be actively used for both identifying the event category (Experiment 1) and reasoning about it (Experiments 2 and 3). By contrast, the endpoints of unbounded events – since they do not involve major changes - are not as salient and do not drive cognitive processing effectively. This overall line of reasoning is consistent with findings from a recent eye tracking study showing that, at the video offset, people pay more attention to the action and the object in resultative events that involve a high degree of change in the affected object compared to events with a less salient change of state (Sakarias & Flecken, 2019).

Our results go beyond Event Segmentation theory in one important respect, however. In that model, the notion of boundary does not specify whether the event has reached its natural endpoint or has simply stopped. In the present account, the notion of boundedness allows us to specify boundaries in a more nuanced way: even though both bounded and unbounded events can and do come to an end, endpoint represents something different in each case (reaching an inherent endpoint, i.e. culmination, as opposed to mere cessation respectively; see Figure 1). Thus the linguistically based notion of boundedness offers a subtler way of defining events and event boundaries in event cognition.

Finally, our results bear on a recent proposal according to which events are represented as a series of intersecting representations of the objects in them (Intersecting Objects History theory; Altmann & Ekves, 2019). This proposal is supported by neural evidence that participants track multiple object changes when reading event descriptions, presumably connecting such changes to how an event unfolds (Hindy et al., 2012; Solomon et al., 2015). The idea that object state changes play a critical role for event cognition is consistent with our findings which show that observers are sensitive to the difference between bounded events that involve pronounced object changes and unbounded events that lack a specific or observable change in the affected object. However, unlike this theory, our own approach does not take object changes (and corresponding events) to occur independently of an observer but rather views boundedness as a property of the cognitive representation of events (in that sense, it leaves open the possibility that the very same object state in the world might be construed as either a bounded or an unbounded event; see next section).

Extensions and future directions

The present results can be extended in a number of directions. First, the methods used to test the processing of event stimuli in the present experiments involved judgment tasks without time constraints. These methods allow us to understand human cognitive processing but could be extended to include additional, time-sensitive measures of event interpretability (e.g. measures tapping how the boundedness profile of an event is calculated as the mind processes a dynamically unfolding perceptual input stream; cf. also Huff & Papenmeier, 2017).

Second, our video stimuli, created against the clean background of a lab room, highlighted the actor, the action and the affected object, controlled for how intentional the events looked, and were constructed so as to be readily perceived in either bounded or unbounded terms. In reality, however, events are encoded in richer, more complex contexts which can influence how viewers represent the event structure. Furthermore, in many cases, the same situation can be encoded as either bounded or unbounded depending on the perspective of the viewer. For instance, an instance of an unbounded event used in our experiments, shuffling poker cards, may be conceptualised as a bounded event in a setting where a couple of friends are about to play a poker game (the endpoint of the event would be defined by the moment that the cards are ready to use, i.e. by context; Depraetere, 2007; Filip, 2001; Hard et al., 2006; Zacks, 2004; Zacks & Tversky, 2001). In other cases, whether an event is construed as bounded or not might depend on the zoom lens on the situation (cracking a peanut is bounded but cracking peanuts is not). Future research needs to address how perceptual information and higher-level contextual knowledge are integrated in determining the boundedness profile of an event.

Final thoughts

In sum, we have presented evidence suggesting that viewers represent the abstract internal profile of events (or boundedness) and process individual event time points in accordance with underlying assumptions about event structure. These results contribute to our understanding of how humans perceive and interpret dynamically unfolding events and bear on ways in which event representations can be further used for describing, remembering and processing human experience.

Acknowledgements

The authors thank Raevyn Johnson for helping construct the videos, and Audrey Sokol, Rebekah Mullin and Bryce Lamers for helping with various aspects of data collection. A portion of the data in Experiment 2 has been reported in Ji and Papafragou (2018).

Disclosure statement

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

Funding

This material is based upon work supported by the National Science Foundation under grant # 1632849.

Data availability statement

The data for the present study are available from osf.io/p89nw Ji and Papafragou (2020). Midpoints, endpoints and the cognitive structure of events.

ORCID

Yue Ji http://orcid.org/0000-0002-9978-4006

References

Altmann, G. T. M., & Ekves, Z. (2019). Events as intersecting object histories: A new theory of event representation. *Psychological Review*, *126*(6), 817–840. https://doi.org/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*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bach, E. (1986). The algebra of events. *Linguistics and Philosophy*, *9*(1), 5–16. https://doi.org/10.1007/BF00627432.

Baldwin, D., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, *72*(3), 708–717. https://doi.org/10.1111/1467-8624.00310

Barner, D., & Snedeker, J. (2005). Quantity judgements and individuation: Evidence that mass nouns count. *Cognition*, *97*(1), 41–66. https://doi.org/10.1016/j.cognition.2004.06.009

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*(2), 805–832. https://doi.org/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*(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Boltz, M. G. (1995). Effects of event structure on retrospective duration judgments. *Perception and Psychophysics*, *57*(7), 1080–1096. https://doi.org/10.3758/BF03205466

Brockhoff, A., Huff, M., Maurer, A., & Papenmeier, F. (2016). Seeing the unseen? Illusory causal filling in FIFA referees, players, and novices. *Cognitive Research: Principles and Implications*, *1*(1), 1–12. https://doi.org/10.1186/s41235-016-0008-5

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

Csibra, G., Biro, S., Koos, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111–133. https://doi.org/10.1207/s15516709cog2701_4

Depraetere, I. (2007). (A)telicity and intentionality. *Linguistics*, *45*(2), 243–269. https://doi.org/10.1515/LING.2007.008

Dowty, D. (1979). *Word meaning and Montague grammar*. Kluwer.

Filip, H. (2001). Nominal and verbal semantic structure: Analogies and interactions. *Language Sciences*, *23*(4), 453–501. https://doi.org/10.1016/S0388-0001(00)00033-4

Filip, H. (2004). The telicity parameter revisited. In R. Young (Ed.), *Proceedings of the 14th Semantics and linguistic theory Conference* (pp. 92–109). LSA.

Folli, R., & Harley, H. (2006). What language says about the psychology of events. *Trends in Cognitive Science*, *10*(3), 91–92. https://doi.org/10.1016/j.tics.2006.01.002

Gold, D., Zacks, J., & Flores, S. (2017). Effects of cues to event segmentation on subsequent memory. *Cognitive Research: Principles and Implications*, *2*(1), 1–15. https://doi.org/10.1186/s41235-016-0043-2

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: General*, *142*(3), 880–905. https://doi.org/10.1037/a0030045

Hard, B. M., Meyer, M., & Baldwin, D. (2019). Attention reorganizes as structure is detected in dynamic action. *Memory & Cognition*, *47*(1), 17–32. https://doi.org/10.3758/s13421-018-0847-z

Hard, B. M., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of Experimental Psychology: General*, *140*(4), 586–604. https://doi.org/10.1037/a0024310

Hard, B. M., Tversky, B., & Lang, D. S. (2006). Making sense of abstract events: Building event schemas. *Memory & Cognition*, *34*(6), 1221–1235. https://doi.org/10.3758/BF03193267

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*(17), 5795–5803. https://doi.org/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.

Huff, M., & Papenmeier, F. (2017). Event perception: From event boundaries to ongoing events. *Journal of Applied Research in Memory and Cognition*, *6*(2), 129–132. https://doi.org/10.1016/j.jarmac.2017.01.003

Huff, M., Papenmeier, F., & Zacks, J. M. (2012). Visual target detection is impaired at event boundaries. *Visual Cognition*, *20*(7), 848–864. https://doi.org/10.1080/13506285.2012.705359

Jackendoff, R. (1991). Parts and boundaries. *Cognition*, *41*(1–3), 9–45. https://doi.org/10.1016/0010-0277(91)90031-X

Ji, Y., & Papafragou, A. (2018). Midpoints and endpoints in event perception. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Meeting of the cognitive Science Society* (pp. 1874–1879). Cognitive Science Society.

Ji, Y., & Papafragou, A. (2020). Is there an end in sight? Viewers' sensitivity to abstract event structure. *Cognition*, *197*, 104197. https://doi.org/10.1016/j.cognition.2020.104197

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, Groningen-Amsterdam studies in semantics 11* (pp. 75–115). Foris Publications.

Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and grammar* (pp. 197–235). Kluwer.

Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, *96*(1), 1–33. https://doi.org/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*(3), 517–544. https://doi.org/10.1111/j.1551-6709.2011.01220.x

Levine, D., Hirsh-Pasek, K., Pace, A., & Golinkoff, R. (2017). A goal-bias in action: The boundaries adults perceive in events align with sites of actor intent. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(6), 916–927. https://doi.org/10.1037/xlm0000364

Loucks, J., Mutschler, C., & Meltzoff, A. N. (2017). Children's representation and imitation of events: How goal organization influences 3-year-old children's memory for action sequences. *Cognitive Science*, *41*(7), 1904–1933. https://doi.org/10.1111/cogs.12446

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology*, *15*(5), 533–545. https://doi.org/10.1002/acp.724

Malaia, E. (2014). It still isn't over: Event boundaries in language and perception. *Language and Linguistics Compass*, *8*(3), 89–98. https://doi.org/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*, *59*(4), 4094–4101. https://doi.org/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*(4), 1372–1389. https://doi.org/10.1037/a0036646

Mittwoch, A. (2013). On the criteria for distinguishing accomplishments from activities, and two types of aspectual misfits. In B. Arsenijević, B. Gehrke, & R. Marín (Eds.), *Studies in the composition and decomposition of event predicates* (pp. 27–48). Springer.

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

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, *12*(5), 436–450. https://doi.org/10.1016/0022-1031(76)90076-7

Newtson, D., Engquist, G., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology*, *35*(12), 847–862. https://doi.org/10.1037/0022-3514.35.12.847

Ongchoco, J. D. K., & Scholl, B. J. (2019). Did that just happen? Event segmentation influences enumeration and working memory for simple overlapping visual events. *Cognition*, *187*, 188–197. https://doi.org/10.1016/j.cognition.2019.01.002

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, *34*(6), 1064–1092. https://doi.org/10.1111/j.1551-6709.2010.01107.x

Papafragou, A. (2015). The representation of events in language and cognition. In E. Margolis, & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 327–345). MIT Press.

Papenmeier, F., Brockhoff, A., & Huff, M. (2019). Filling the gap despite full attention: The role of fast backward inferences for event completion. *Cognitive Research: Principles and Implications*, *4*(1), 1–17. https://doi.org/10.1186/s41235-018-0151-2

Parsons, T. (1990). *Events in the semantics of English: A study in subatomic semantics*. MIT Press.
Radvansky, G., & Zacks, J. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, *17*, 133–140. https://doi.org/10.1016/j.cobeha.2017.08.006
Regier, T., & Zheng, M. (2007). Attention to endpoints: A cross-linguistic constraint on spatial meaning. *Cognitive Science*, *31*(4), 705–719. https://doi.org/10.1080/15326900701399954
Sakarias, M., & Flecken, M. (2019). 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), e12708–30. https://doi.org/10.1111/cogs.12708
Smith, C. S. (1991). *The parameter of aspect*. 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*(12), 2324–2338. https://doi.org/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*(19), 5968–5973. https://doi.org/10.1073/pnas.1423080112
Strickland, B., & Keil, F. (2011). Event completion: Event based inferences distorts memory in a matter of seconds. *Cognition*, *121*(3), 409–415. https://doi.org/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*(3), 570–580. https://doi.org/10.1037/a0037750
Swallow, K., Zacks, J., & Abrams, R. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236–257. https://doi.org/10.1037/a0015631
Taylor, B. (1977). Tense and continuity. *Linguistics and Philosophy*, *1*, 199–220. https://doi.org/10.1007/BF00351103
Tenny, C. (1987). *Grammaticalizing aspect and affectedness* [Unpublished doctoral dissertation]. MIT.
Tversky, B., Zacks, J., Morrison, J., & Hard, B. (2011). Talking about events. In J. Bohnemeyer, & E. Pederson (Eds.), *Event representation in language and cognition* (pp. 216–227). Cambridge University Press.
Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*(1), 3–15. https://doi.org/10.1037/0033-295X.94.1.3
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). Routledge.
Van Hout, A. (2016). Lexical and grammatical aspect. In J. Lidz, W. Synder, & J. Pater (Eds.), *The Oxford handbook of developmental linguistics* (pp. 587–610). Oxford University Press.
Van Hout, A. (2018). On the acquisition of event culmination. In K. Syrett, & S. Arunachalam (Eds.), *Semantics in language acquisition* (pp. 96–121). John Benjamins Publishers.
Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, *66*(2), 143–160. https://doi.org/10.2307/2182371
Wagner, L. (2009). Manners and goals in pre-linguistic thought: The origins of aspectual construal. In J. Chandlee, M. Franchini, S. Lord, & G. Rheiner (Eds.), *Proceedings of the 33th Annual Boston University Conference on language development* (pp. 599–610). Cascadilla Press.
Wagner, L. (2012). First language Acquisition. In R. Binnick (Ed.), *The Oxford handbook of tense and aspect* (pp. 458–480). Oxford University Press.
Wagner, L., & Carey, S. (2003). Individuation of objects and events: A developmental study. *Cognition*, *90*(2), 163–191. https://doi.org/10.1016/S0010-0277(03)00143-4
Waxman, S., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development*, *1*(2), 139–156. https://doi.org/10.1016/S0885-2014(86)80016-8
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 University Press.
Wilder, D. A. (1978). Predictability of behaviors, goals, and unit of perception. *Personality and Social Psychology Bulletin*, *4*(4), 604–607. https://doi.org/10.1177/014616727800400422
Zacks, J. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*(6), 979–1008. https://doi.org/10.1207/s15516709cog2806_5
Zacks, J., Speer, N., Swallow, K., Braver, T., & Reynolds, J. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273
Zacks, J., & Swallow, M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16*(2), 80–84. https://doi.org/10.1111/j.1467-8721.2007.00480.x
Zacks, J., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. https://doi.org/10.1037/0033-2909.127.1.3
Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x