THE INTERNAL TEMPORAL STRUCTURE OF EVENTS

IN LANGUAGE AND COGNITION

by

Yue Ji

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics

Summer 2020

THE INTERNAL TEMPORAL STRUCTURE OF EVENTS

IN LANGUAGE AND COGNITION

by

Yue Ji

Approved: _____

Benjamin Bruening, Ph.D.
Chair of the Department of Linguistics and Cognitive Science

Approved: _____

John Pelesko, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____

Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education
and Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Anna Papafragou, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Satoshi Tomioka, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Zhenghan Qi, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Helene Intraub, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

I am grateful for so many people who encouraged, inspired, helped and supported me as I completed this dissertation during the COVID-19 outbreak.

I would like to first thank my advisor Anna Papafragou for her guidance and support over the past five years. I can't thank her enough for patiently teaching me how to think, write and communicate like a scientist. I owe nearly everything that I am as a researcher to her. She couldn't have been a better role model.

I thank my committee members Zhenghan Qi, Satoshi Tomioka and Helene Intraub for their insightful questions and helpful comments on this work. I also thank Li Sheng, the advisor for my second qualifying paper, for her valuable research and career guidance.

I extend my gratitude to my undergraduate and master's advisor, Xiaolu Yang. I was initially drawn to the world of language science, particularly to the amazing field of child language development, by the enthusiasm of her teaching and research. She has taught me so much about pursuing an academic career, as well as enjoying life outside of the lab. I am grateful for her continuing encouragement and support.

Many thanks go to the members of the Language and Cognition Lab for their creative feedback, critical discussions and invaluable support. Special thanks to Myrto Grigoroglou, who gave me so many good suggestions to get started with experiments,

and helped me go through my first mixed-effect model for data analysis, and to Alyssa Kampa, without whom it would be much harder to survive the first year of the Ph.D. program.

Thanks to the National Science Foundation (#1632849) for funding these studies. Special thanks to the Office of Graduate and Professional Education for the Doctoral Fellowship in 2018-2019 that allowed me to focus on my research, and the Department of Linguistics and Cognitive Science for the summer funding that supported data collection.

Many thanks to the Early Learning Center and the Laboratory Preschool at the University of Delaware, and the Delaware Children's Museum for helping me recruit child participants in my studies. Thanks also to the hundreds of UD undergraduate students who participated in my studies.

Thanks to Daneil Bresnahan, Bryce Lamers, Matthew Soule, Audrey Sokol, Rebekah Mullin, Sierrah Harries, Victor Ike, Jamie Epstein for their assistance with data collection and transcription in these studies. Many thanks to Raevyn Johnson, the actor in the experiment video stimuli, for her patience and time.

Finally, I owe a great deal of thanks to my parents, who have always supported me and rarely questioned about my boring lifestyle. Without their understanding and positive attitude, I would not be able to do any of this.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

People segment their continuous stream of experience into events, or temporal segments that have a beginning and an endpoint. But how are such event boundaries defined? Linguistic theories of event encoding draw a distinction between *bounded* events that are non-homogeneous, structured temporal developments leading to an inherent endpoint (e.g., “dress a teddy bear”) and *unbounded* events that are homogeneous temporal entities with no inherent endpoint (e.g., “pat a teddy bear”). 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 dissertation, we fill this gap. We hypothesize that event cognition tracks boundedness, and the sensitivity to boundedness affects how individual temporal slices of events (such as midpoints or endpoints) are processed and integrated into a coherent event representation.

In Chapter 2, 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. The 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. In

Chapter 3, we find that viewers react differently to interruptions at the end compared to interruptions in the middle of an event – but only when perceiving events with a bounded structure; moreover, viewers are more disturbed by temporal disruptions of a bounded event compared to an unbounded event. Chapter 4 tests our hypotheses from a developmental perspective and shows continuity in mechanisms of event cognition. We demonstrate that 4-to-5-year-old children compute boundedness during event categorization and treat it as different from event completion. Furthermore, 4-to-5-year-olds evaluate interruptions at event endpoints vs. midpoints differently when exposed to bounded but not unbounded events.

We conclude that event viewers are sensitive to foundational and abstract properties of event temporal structure. This sensitivity can emerge from an early age, support the acquisition of linguistic-aspectual distinctions, and scaffold the way children conceptualize and process their dynamic experience.

Chapter 1

INTRODUCTION

1.1 Event and event temporal structure in cognition and language

The world provides us with a fluid stream of dynamic experience but we encode this experience as concrete units, i.e., events. Segmenting events from continuous input lies at the heart of event cognition. It is an automatic and foundational process that provides the basis for understanding, remembering and talking about the unfolding experience (Kurby & Zacks, 2008; Richmond & Zacks, 2017). The early studies of Newtson and colleagues (e.g., Newtson, 1973; Newtson & Engquist, 1976) have shown that people watching a series of activities have similar intuitions about when a meaningful unit (i.e., an event) ends and another begins. Recent event segmentation work further demonstrates that adults and even young infants are able to decide event boundaries by tracking salient changes in important event features (e.g., a new place as in mom walking into the kitchen; Baldwin, Baird, Saylor, & Clark, 2001; Saylor, Baldwin, Baird, & LaBounty, 2007; Zacks, 2001; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). This work provides a spatiotemporal framework that defines an event as "a segment of time at a given location that is conceived by an observer to have a beginning and an end" (Zacks & Tversky, 2001).

Although research on event segmentation has revealed that event boundaries are moments of accumulating change, in many cases (e.g., mom walking around the house), it is not clear what counts as an event beginning or ending (e.g., Hard, Recchia, & Tversky, 2011; Swallow, Zacks, & Abrams, 2009; for a recent review, see Radvansky & Zacks, 2017). The existing spatiotemporal framework does not capture *how* a change along a specific dimension occurs over time. What is missing is an account of the event temporal structure that can distinguish, for example, an event of walking into the kitchen from an event of walking around the house. More generally, the definition of events across time and space does not directly address how people identify, compare between and categorize events.

This dissertation aims to fill the gap. We are inspired by how language encodes event temporal structure as well as the nature of event boundaries. Linguistic descriptions individuate an event by "carving out" a "time slice" from what is "out there in the world" (Filip, 1993; Parsons, 1990; Partee, 1999; Van Hout, 2016). Moreover, event descriptions, particularly the verb phrases, distinguish different "time schemata", i.e., the ways events unfold over time (Vendler, 1957). A basic division lies between events that are temporally bounded (e.g., walking into the kitchen, fixing a car) from those temporally unbounded (e.g., walking around the house, driving a car; Filip, 2012). The two types of events have different internal temporal structures and ways they come to an end.

Would the bounded-unbounded distinction in language be part of conceptual representations of events? Would observers be sensitive to the different temporal

structures of bounded and unbounded events in cognition? We test these possibilities through a series of non-linguistic tasks. Our investigation starts with adults and then extends to 4-to-5-year-old children. The latter age group demonstrates some awareness of the linguistic bounded-unbounded distinction but has not fully acquired how boundedness is expressed in their language (e.g., Wagner, 2006; Ogiela, 2007; Penner, Schulz & Wymann, 2003; van Hout, 1998; for an overview, see van Hout, 2018). We ask whether these young learners are sensitive to the temporal profile of events in their mind, despite of their still developing linguistic knowledge.

1.2 Outline of the dissertation

This dissertation consists of three sets of experiments. The first set investigates the cognitive representation of bounded and unbounded event categories in adults (Chapter 2). The second set looks into viewers' sensitivity to the internal temporal structure of bounded and unbounded events in adults (Chapter 3). The last set examines young children's awareness of boundedness in event temporal structure (Chapter 4). Each of these chapters was developed as a journal publication and thus has an independent structure. A modified version of Chapter 2 appeared as Ji, Y. & Papafragou, A. (in press). Is there an End in Sight? Viewers' sensitivity to abstract event structure. *Cognition*. A modified version of Chapter 3 was submitted to *Language, Cognition and Neuroscience,* entitled "Midpoints, endpoints and the cognitive structure of events" and is under revision. A modified version of Chapter 4 was submitted to *Developmental Psychology*, entitled "Children track the internal

temporal structure of events" and is under review. There is some overlap in the introductory section of each chapter. Lastly, in Chapter 5 we draw conclusions based on our findings from each study.

Chapter 2

IS THERE AN END IN SIGHT? VIEWERS' SENSITIVITY TO ABSTRACT EVENT STRUCTURE

2.1 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; Newtson, Engquist, & Bois, 1977; Magliano, Miller, & Zwaan, 2001; Zacks, Speer, Swallow, Braver, & Reynolds, 2007), or conceptual features such as the goal-directedness or causal structure of an experience (cf. a person on a diet hitting a target; Zacks & Swallow, 2007). 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 (Boltz, 1995; Newston & Engquist, 1976; Huff, Papenmeier, & Zacks, 2012; 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, 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 <u>to the store</u>*) is more accurately encoded in both language and memory as opposed to the source (as in *Maria ran <u>away from the house</u>*; 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 sight, even infants as young as 12 months old can actively infer it from observed actions (Csibra, Biro, Koos, & Gergely, 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 (i.e., the endpoint of the kicking event), even when it was actually omitted from the display (Strickland & Keil, 2011).

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 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.

[1] Boundedness has been defined by some as "expressing an action tending towards a goal" (Garey, 1957) 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*).

(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 & Snedeker, 2005; 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 five 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.

2.2 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.

2.2.1 Method

2.2.1.1 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).

2.2.1.2 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.

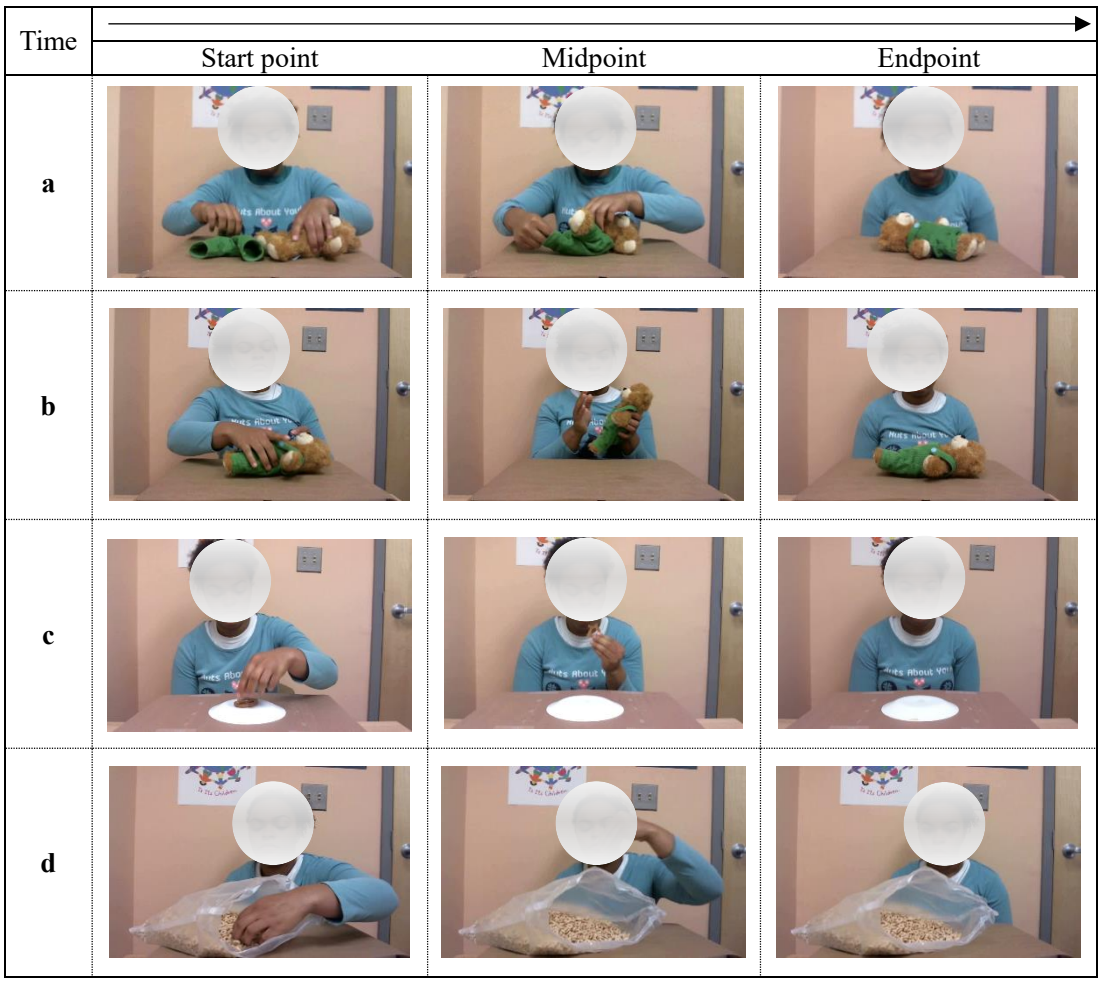

Figure 1 Pairs of video stimuli in Experiment 1: (a) dress a teddy bear (bounded) vs. (b) pat a teddy bear (unbounded); and (c) eat a pretzel (bounded) vs. (d) eat cheerios (unbounded).

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	No.	Unbounded Events
Training	Nature of Action	1	fold up a handkerchief	21	wave a handkerchief
		2	put up one's hair	22	scratch one's hair
		3	stack a deck of cards	23	shuffle a deck of cards
		4	group pawns based on color	24	mix pawns of two colors
	Nature of Affected Object	5	draw a balloon	25	draw circles
		6	tie a knot	26	tie knots
		7	eat a pretzel	27	eat cheerios
		8	flip a postcard	28	flip pages
Testing	Nature of Action	9	dress a teddy bear/	29	pat a teddy bear
		10	roll up a towel/	30	twist a towel
		11	fill a glass with milk/	31	shake a bottle of milk
		12	scoop up yogurt/	32	stir yogurt
	Nature of Affected Object	13	peel a banana/	33	crack peanuts
		14	blow a balloon/	34	blow bubbles
		15	tear a paper towel/	35	tear paper towels
		16	paint a star/	36	paint stuff
Surprise Testing	Nature of Action	17	close a fan/	37	use a fan for oneself
		18	crack an egg/	38	beat an egg
	Nature of Affected Object	19	cut a ribbon in half/	39	cut ribbon from a roll
		20	stick a sticker/	40	stick stickers

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.

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.

2.2.1.3 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.

2.2.2 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. 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 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).

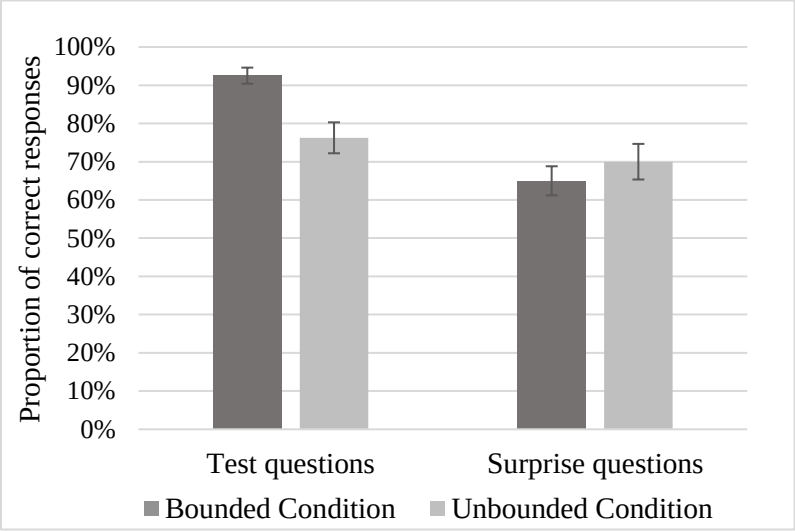

Figure 2 Proportion of correct responses in Experiment 1. Error bars represent ±SEM. Note that surprise questions probed a non-target category (unbounded events in the Bounded condition, and bounded events in the Unbounded condition).

Table 2 *Fixed effect estimates for multi-level model of event category identification in the testing and surprise testing phase of 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). As shown in Table 3, 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.

Table 3 *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
* $p < .05$, ** $p < .01$, *** $p < .001$

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").

2.2.3 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.

2.3 Experiment 2

In Experiment 1, participants might have benefited from the successive, contrastive presentation of paired videos in the training phase. Experiment 2 asked whether boundedness categories could be extracted from a less supportive context compared to the training phase of Experiment 1.

2.3.1 Method

2.3.1.1 Participants

A new group of forty undergraduates (18 female, 22 male; M_{age} = 19.6, age range: 18.1-22.2) at the University of Delaware was recruited. Data from an additional

adult were collected but excluded because this participant failed to understand the task and did not finish all the questions.

2.3.1.2 Stimuli and procedure

The stimuli and procedure were identical to those in Experiment 1 with one exception. In the training phase, the sequence of the 16 videos was pseudo-randomized 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 category was weakened.

2.3.2 Results

Results from Experiment 2 are shown in Figure 3. 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 4). For test questions, 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%) (z = -3.35, p < .001). By contrast, for surprise questions, no significant difference between the Bounded (M = 60.0%) and the Unbounded condition (M = 68.8%) was found (z = 0.82, p > .250). Participants performed at levels different from chance for each Question Type by Condition combination (all ps < .05) with the exception of surprise questions in the Bounded condition, i.e., those that targeted unbounded events (t (19) = 1.71, p = .104).

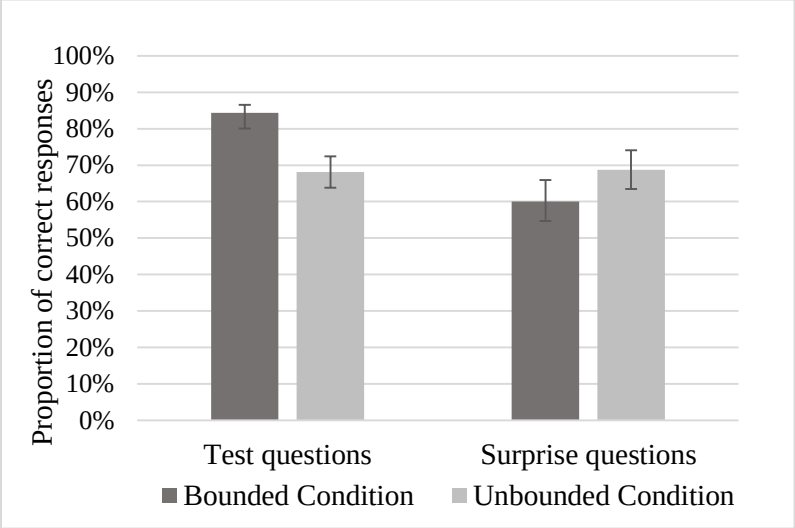

Figure 3 Proportion of correct responses in Experiment 2. Error bars represent ±SEM.

Table 4 *Fixed effect estimates for multi-level model of event category identification in the testing and surprise testing phase of Experiment 2.*

Question Type	Effect	Estimate	SE	z value
Test question	(Intercept)	1.82	0.32	5.66***
	Condition	-0.99	0.30	-3.35***
Surprise question	(Intercept)	0.41	0.23	1.78
	Condition	0.27	0.33	0.82

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 parameter estimates of fixed effects in Table 5). There were main effects of Question Type ($z = -3.31$, $p < .001$; with performance being better in test compared to surprise questions),

and of Condition (z = -3.42, p < .001; with accuracy being higher in the Bounded than the Unbounded condition). There was also an interaction between the two predictors (z = 2.84, p = .005). Specifically, participants in the Bounded condition were more accurate in test questions compared to surprise questions (z = -2.96, p = .003) but no such difference was found in the Unbounded condition (z = -.27, p > .250).

Organization, 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).

Table 5 *Fixed effect estimates for multi-level model of event category identification in Experiment 2.*

Effect	Estimate	SE	z value
(Intercept)	1.76	0.27	6.57***
Question Type	-1.34	0.41	-3.31***
Condition	-0.96	0.28	-3.42***
Question Type*Condition	1.23	0.43	2.84**

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

2.3.3 Discussion

Results from Experiment 2 are reminiscent of Experiment 1: participants succeeded in identifying the target event category in both conditions. Furthermore, there was an advantage for forming the category of bounded compared to unbounded

events. The category of bounded events could still be extracted at levels different from chance even when not specifically targeted in the task (cf. surprise questions), unlike the category of unbounded events.

2.4 Experiment 3

In Experiments 1 and 2, participants were exposed to (either successively or discontinuously 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 3, we introduced a training phase with single instances of events and asked whether participants could identify boundedness categories without any support from contrast.

2.4.1 Method

2.4.1.1 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.

2.4.1.2 Stimuli

Video stimuli in the testing phase were identical to those used in Experiments 1 and 2. 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 the previous experiments (16 total videos), we created a set of additional videos using the same actor. 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).

2.4.1.3 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 Experiments 1 and 2.

2.4.2 Results

Results from Experiment 3 are shown in Figure 4. As shown in Table 6, 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).

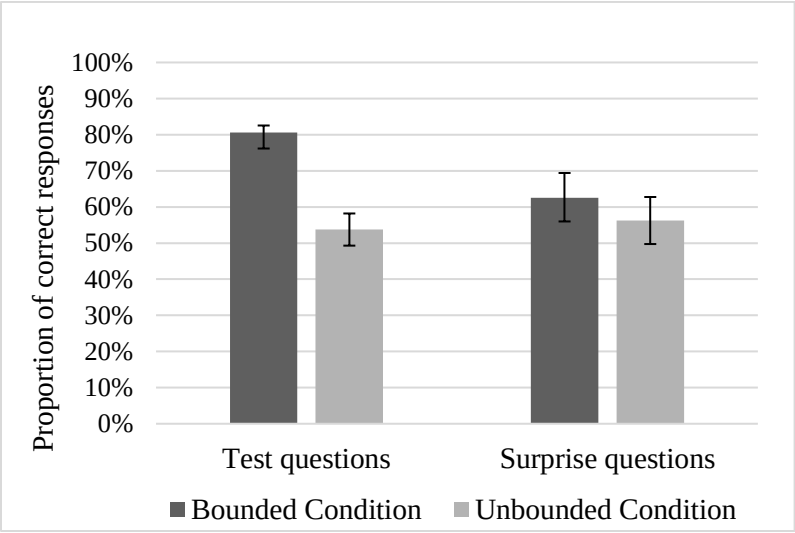

Figure 4 Proportion of correct responses in Experiment 3. Error bars represent ±SEM.

Table 6 *Fixed effect estimates for multi-level model of event category identification in the testing and surprise testing phase of Experiment 3.*

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 7). 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 ($z = 1.13$, $p > .250$).

Table 7 *Fixed effect estimates for multi-level model of event category identification in Experiment 3.*

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

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.

2.4.3 Comparison across Experiments 1-3

To compare performance across experiments, the binary data from test questions were submitted to a mixed logit model with Condition (Bounded vs. Unbounded) and Experiment (1 vs. 2 vs. 3) as predictors. The fixed effect of

Experiment was analyzed with planned comparisons using simple contrast coding (c_1: -.33, -.33, .66, c_2: .50, -.50, 0), which first compared Experiments 1 and 2 vs. Experiment 3 to identify the role of a training phase with and without contrastive pairs, and further compared Experiment 1 vs. Experiment 2 to identify the role of strong vs. weak contrast within the training phase. Table 8 reports parameter estimates of fixed effects. There was a fixed effect of Condition (z = -6.98, 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 Experiments 1 and 2 (where contrastive pairs were available in the training phase) compared to Experiment 3 (z = -4.59, p < .001); furthermore, Experiment 1 that included a strong contrast between bounded and unbounded events led to better results than Experiment 2 where the contrast was weak (z = 2.62, p = .008). Responses in surprise questions were analyzed in the same way. However, neither Condition (Bounded vs. Unbounded), nor Experiment (1 vs. 2 vs. 3) or their interaction could improve the model fit.

Table 8 *Fixed effect estimates for multi-level model of event category identification in Experiments 1-3.*

Effect	Estimate	SE	z value
(Intercept)	1.92	0.19	10.32***
Condition	-1.20	0.17	-6.98***
Experiment (3 vs. 1&2)	-0.78	0.17	-4.59***
Experiment (1 vs. 2)	0.56	0.21	2.62**

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

2.4.4 Discussion

In Experiment 3, the lack of contrastive pairs during training made the category identification task harder than in previous experiments. 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 1-3? 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 3, 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 the three experiments 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.

2.5 Experiment 4

A different alternative explanation for the patterns of performance in Experiments 1-3 s 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, Mittwoch, 2013). 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 4, 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.

2.5.1 Method

2.5.1.1 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.

2.5.1.2 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 5 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 3.

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.

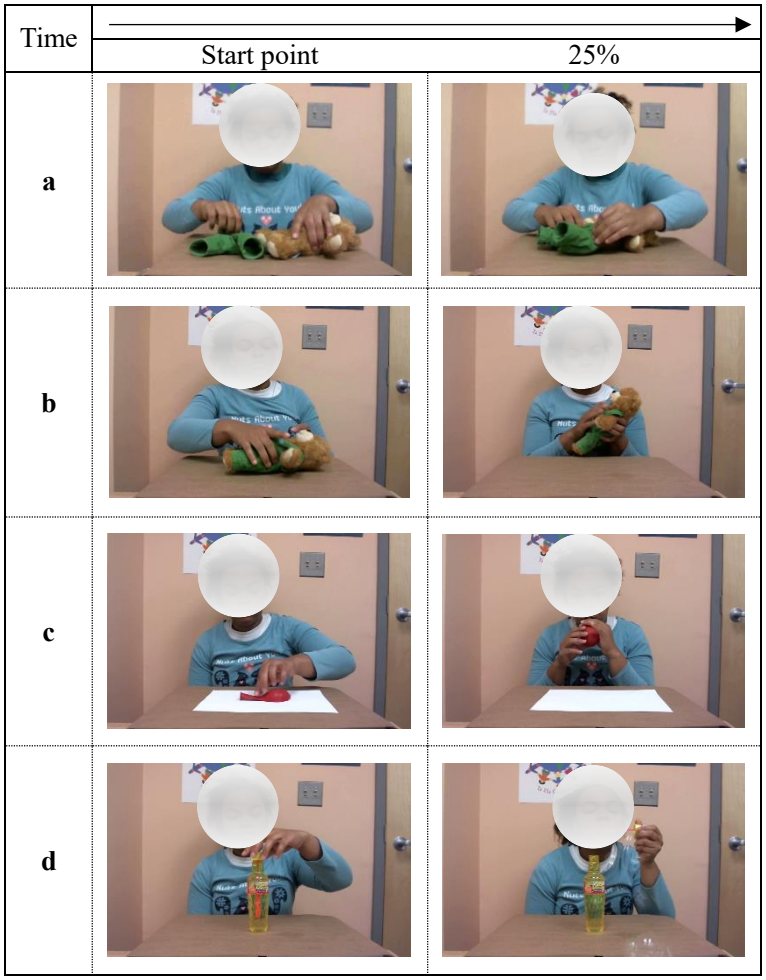

Figure 5 Truncated videos in Experiment 4: (a) dress a teddy bear vs. (b) pat a teddy bear, (c) blow a balloon vs. (d) blow bubbles.

2.5.1.3 Procedure

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

2.5.1.4 Results

Results from Experiment 4 are shown in Figure 6. We examined the fixed effects of Condition (Bounded vs. Unbounded) and Video Type (Truncated vs. Full) on accuracy. As shown in Table 9, 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).

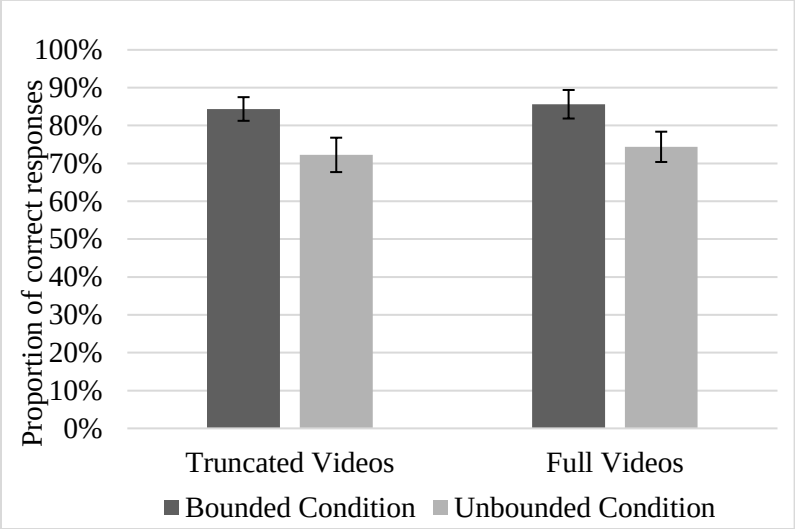

Figure 6 Proportion of correct responses in Experiment 4. Error bars represent ±SEM.

Table 9 *Fixed effect estimates for multi-level model of event category identification in Experiment 4.*

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).

2.5.2 Discussion

In Experiment 4, 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.

2.6 Experiment 5

A further alternative explanation of the results of Experiments 1-4 is that participants might have silently 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 5 we introduced a secondary counting task to disrupt linguistic encoding in the training phase and assessed whether this affected viewers' computation of boundedness categories.

2.6.1 Method

2.6.1.1 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.

2.6.1.2 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.

2.6.2 Results

Results from Experiment 5 are shown in Figure 7. 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 10). 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).

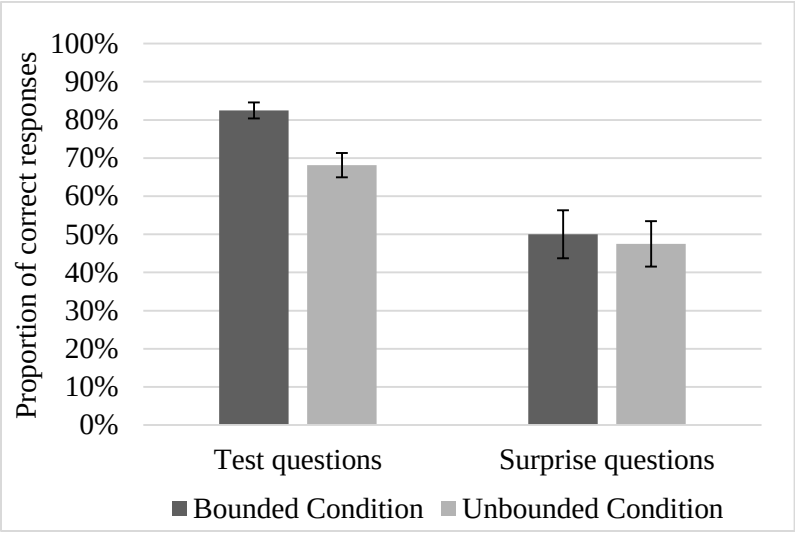

Figure 7 Proportion of correct responses in Experiment 5. Error bars represent ±SEM.

Table 10 *Fixed effect estimates for multi-level model of event category identification in the testing and surprise testing phase of Experiment 5.*

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. As shown in Table 11, 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 ($z = 1.49$, $p = .136$).

Table 11 *Fixed effect estimates for multi-level model of event category identification in Experiment 5.*

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

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 5 by examining the fixed effects of Question Type (Test vs. Surprise), Condition (Bounded vs. Unbounded), and Experiment (1 vs. 5), as well as their interactions (see Table 12). 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 5 compared to Experiment 1 ($z = -4.51$, $p < .001$).

Table 12 *Fixed effect estimates for multi-level model of event category identification in Experiment 1 and 5.*

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

2.6.3 Discussion

Experiment 5 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.

2.7 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 1-3). Viewers treat boundedness in event cognition as different from simple event completion (Experiment 4). Furthermore, viewers identify bounded and unbounded events even when linguistic encoding is suppressed (Experiment 5). 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 and explore the theoretical importance of the difference between boundedness categories.

2.7.1 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; Hafri,

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 4). 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 3). 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.

2.7.2 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 3).

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).[2]

2.7.3 Final thoughts

To conclude, across five experiments, we have found that people can extract abstract information about the temporal structure 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.

[2] In this 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.

Chapter 3

MIDPOINTS, ENDPOINTS AND THE COGNITIVE STRUCTURE OF EVENTS

3.1 Introduction

In Chapter 2, we have shown that event cognition tracks boundedness, an abstract feature of the internal temporal structure of events. This result strongly raises the possibility that the psychological representation of event boundedness has broader, testable implications about how people process individual temporal slices of events (such as event midpoints or endpoints) and form a coherent event representation. Here in Chapter 3 we test this possibility.

As noted in Chapter 2, research on event segmentation has revealed the privileged status of event boundaries compared to non-boundary moments. This work on event representation leaves 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, Zacks, & Flores, 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.

In this chapter, we propose to address both of these issues by examining how abstract properties of event structure affect how people track the internal temporal structure of events, including temporal landmarks such as midpoints and event boundaries. 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 bounded and unbounded events with different internal structures and different ways in which they come to an end. We focus on the property of homogeneity, or cumulativity (Hinrichs, 1985; Krifka, 1989,

1998; Taylor, 1977) that has been used to characterize the differences in the temporal structure between the two types of events. Consider the following bounded-unbounded event pair used in the first set of experiments.

(1) A girl dressed a teddy bear.

(2) A girl patted a teddy bear.

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.

In two experiments, we pursue the hypothesis that the property of homogeneity should affect how viewers process and weigh temporal slices of different events. We focus 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 expect 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 6 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 categorizing 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 7, 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.

3.2 Experiment 6

In this experiment, we asked whether the differences in bounded and unbounded structure 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. We later asked people to guess whether the girl would like new videos containing mid- or end-interruptions. We hypothesized 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.

3.2.1 Method

3.2.1.1 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.

3.2.1.2 Stimuli

The same event stimuli (showing16 pairs of bounded and unbounded events) as in the training and testing phase of 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 8 and Figure 9 below). Each video was edited twice, once to create a mid-interruption and once to create an end-interruption. The mid-interruption was centered on the video midpoint, while the end-interruption blocked the last 20% of the video.

3.2.2 Procedure

Participants were invited to watch a couple of videos. The experimenter told participants that the girl in the 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 Event type 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 21-28 for unbounded events, presented in a random order), each with two versions shown in succession in the center of the screen. The two versions differed in terms of both the actor's shirt color (blue vs. yellow) and in terms of the placement of the interruption (mid-interruption vs. end-interruption, see Figure 8 and Figure 9 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 color 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 colors, 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 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.

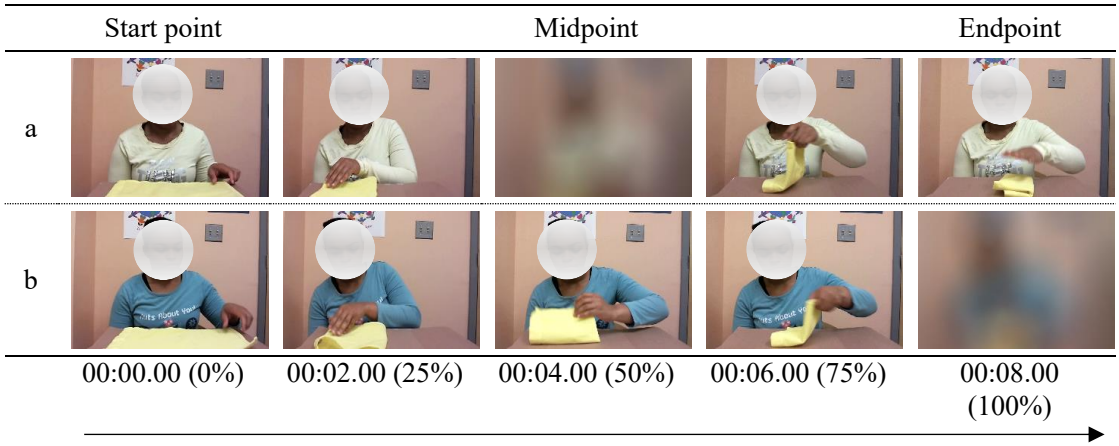

Figure 8 Examples of a training trial for a bounded event (fold up a handkerchief) that includes the two versions of the event: (a) mid-interruption (actor in yellow shirt), (b) end-interruption (actor in blue shirt). The still-frame images were extracted from the video stimuli showing what happened at the beginning, when 25% of the time passed, at the midpoint, when 75% of the time passed and at the end respectively.

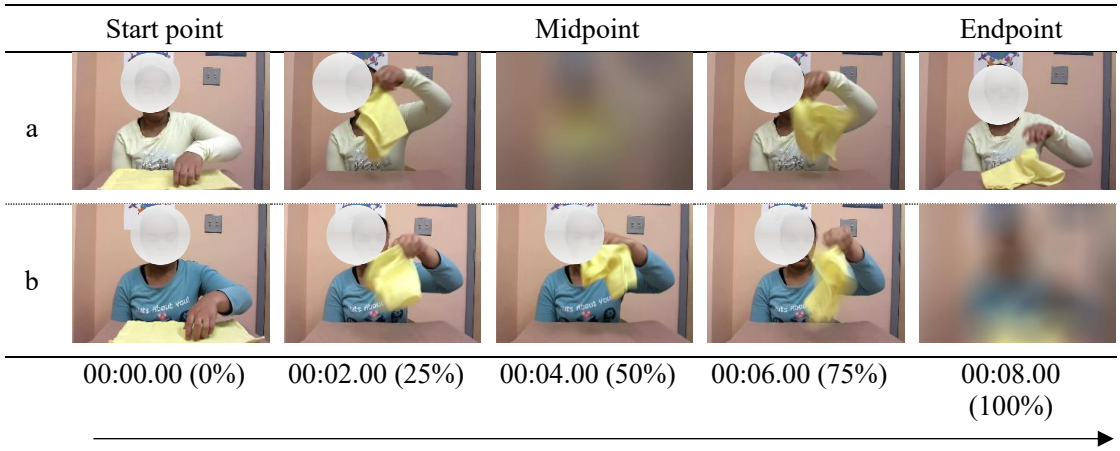

Figure 9 Examples of a training trial for the unbounded counterpart of the event in Figure 8 (wave a handkerchief) that includes the two versions of the event: (a) mid-interruption (actor in yellow shirt), (b) end-interruption (actor in blue shirt). The still-frame images were extracted from the video stimuli showing what happened at the beginning, when 25% of the time passed, at the midpoint, when 75% of the time passed and at the end respectively.

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 29-36 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 color was counterbalanced for that event across lists. Types of interruptions and shirt color 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.

3.2.3 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 color (blue or yellow), they would give a correct response around 50% of the time since shirt color was counterbalanced. As shown in Figure 10, the proportion of correct responses significantly differed from 50% in both Preference

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

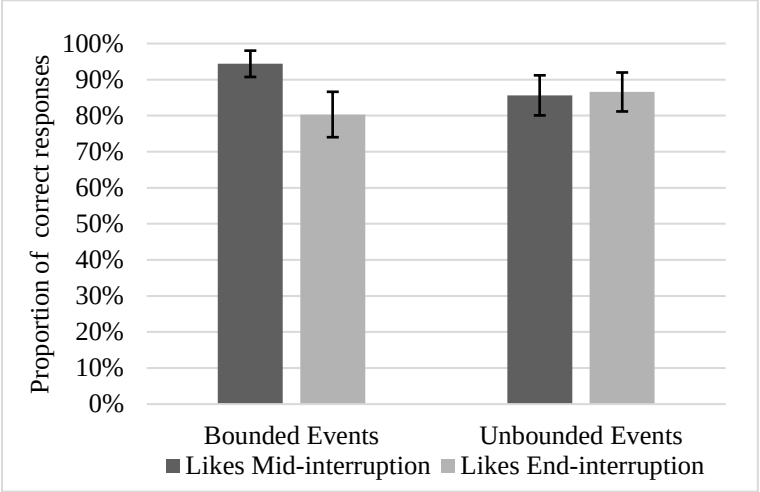

Figure 10 Proportion of correct responses in Experiment 6. Error bars represent ±SEM.

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 color (Blue vs. Yellow), contrast (Action vs. Affect Object), placement of the interruption (Mid vs. End), as well as List would influence accuracy. Adding them 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, Davidson, & Bates, 2008; Barr, 2008). As shown in Table 13, the model revealed an effect of Preference (β = -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 (β = -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 (β = 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%) (β = -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%) (β = 0.08, z = 0.34, p > .250).

Table 13 *Fixed effect estimates for multi-level model of identifying the preference for interruptions in Experiment 6.*

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

Recall that, at the end of the session, participants were asked to write down what kind of videos the girl liked. These responses are summarized in Table 14. 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 "Likes end-interruption" condition). A small number of participants responded on the basis of shirt color (e.g., "The girl seems to alternate on the basis of the color of the shirt"). As shown in Table 14, 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.

Table 14 *Counts (N) and percentages (%) of the two types of responses to the open question in Experiment 6.*

Event Type	Preference	Interruption		Shirt color	
		N	%	N	%
Bounded	Likes Mid-interruption	36/40	90%	5/40	12.5%
	Likes End-interruption	29/40	72.5%	11/40	27.5%
Unbounded	Likes Mid-interruption	37/40	92.5%	3/40	7.5%
	Likes End-interruption	36/40	90%	4/40	10%

3.2.4 Discussion

In Experiment 6, we tested the hypothesis that the abstract internal profile of events (or boundedness) 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.

3.3 Experiment 7

In this experiment, 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, Hirsh-Pasek, Pace, & Golinkoff, 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, Tversky, & Lang, 2006; Hard, Recchia, & Tversky, 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.

3.3.1 Method

3.3.1.1 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.

3.3.1.2 Stimuli

The four bounded events (stack 5 cups on the table, put some Q-tips together, erase a star, organize 3 pairs of socks by color) and four unbounded events (grind

biscuits, roll a ball back and forth, pull a towel, sprinkle pepper) created for the purposes of Experiment 3 were used in the practice phase of this experiment. The 32 unedited videos from the training and testing phase of Experiment 1 were used in the testing phase (see Table 1).

The total 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 11a and 11c). In the "reverse" version, each video was reversed (see Figure 11b and 11d).

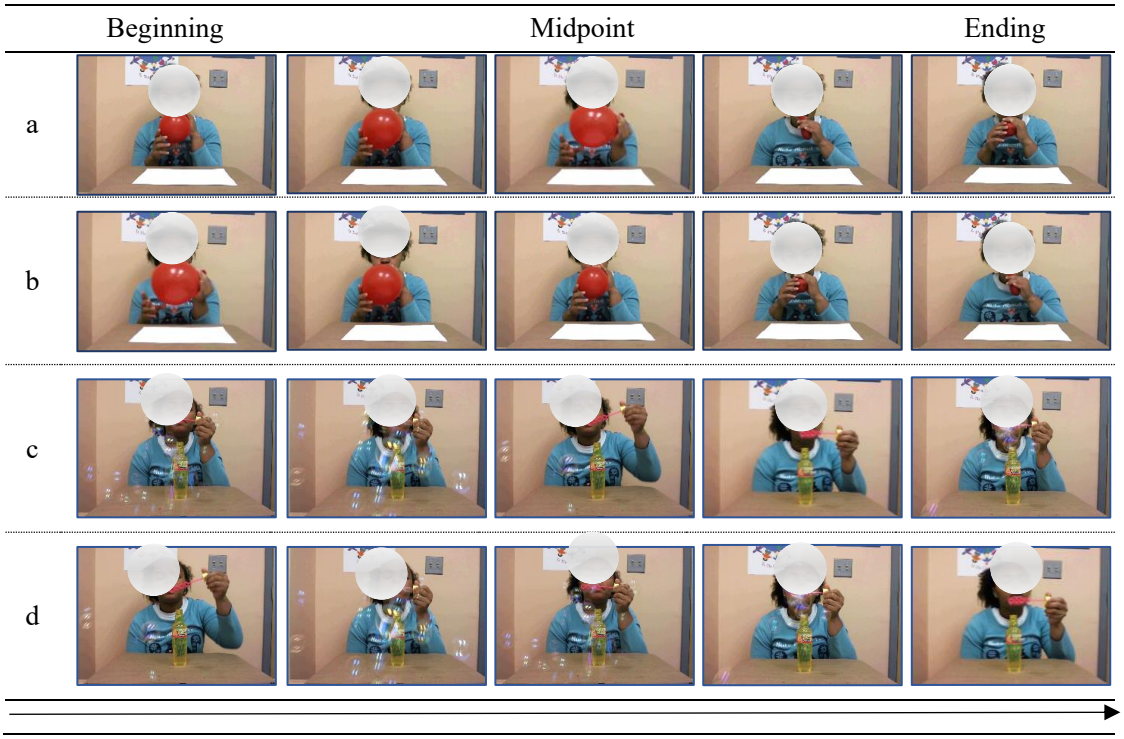

Figure 11 Edited videos in Experiment 7: (a) blow a balloon (bounded) – switched halves; (b) blow a balloon (bounded) – reverse; (c) blow bubbles (unbounded) – switched halves; (d) blow bubbles (unbounded) – reverse.

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.

3.3.1.3 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 familiarize 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.

3.3.2 Results

Results from Experiment 7 are shown in Figure 12. Participants’ rating of the videos were analyzed using multi-level mixed modeling 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 in Table 15, there was a fixed effect of Event Type (β = -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 (β =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 (β = 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.

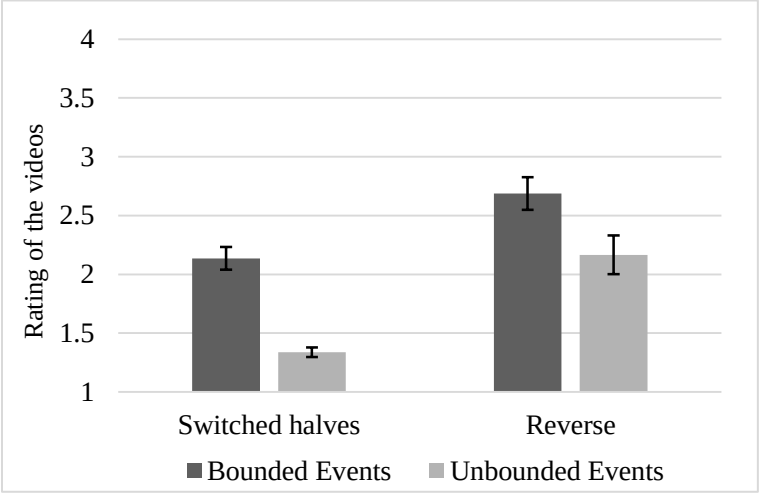

Figure 12 Average rating of the videos in Experiment 7 (on a 7-point scale). High ratings indicate highly confusing videos. Error bars represent ±SEM.

Table 15 *Fixed effect estimates for multi-level model of video rating in Experiment 7.*

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

3.3.3 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 result 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.

3.4 General Discussion

3.4.1 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 started with 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 (Ji & Papafragou, accepted). This hypothesis was consistent with linguistic treatments of aspect or telicity (e.g., Bach, 1986; Krifka, 1989, 1998) and prior experiments from non-linguistic categorization tasks (cf. Strickland et al., 2015). In two 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 6, 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 Experiment7, 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, in press; Strickland et al., 2015; see also Hard, Recchia, & Tversky, 2011; Hard, Meyer, & Baldwin, 2019; Loucks, Mutschler, & Meltzoff, 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 Folli & Harley, 2006; Hafri, Papafragou, & Trueswell, 2013; Lakusta & Landau, 2005; Malaia, 2014; Papafragou, 2015; Strickland et al., 2015; Tversky, Zacks, Morrison, & Hard, 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. This overall conclusion is consistent with findings from a recent eye tracking study showing that people pay more attention to the action and the object in resultative events (that, by definition, have an inherent endpoint) compared to non-resultative events (no inherent endpoint) at the video offset (Sakarias & Flecken, 2019).

Furthermore, our data bear on Event Segmentation theory (e.g., Zacks & Swallow, 2007; Zacks et al., 2009) 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., waving a handkerchief), internal development can be predicted from the very beginning. By contrast, for bounded events with a non-homogeneous structure (e.g., folding up a handkerchief), 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 handkerchief is folded up), are highly salient to observers and can be actively used for both identifying the event category and reasoning about it. 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. 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). 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, Altmann, Kalenik, & Thompson-Schill, 2012; Solomon, Hindy, Altmann, Thompson-Schill, 2015). The idea that object state changes play a critical role for event cognition is related to our findings, since bounded events in many cases involve pronounced object changes. 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). Furthermore, unlike this theory, our approach does not tie the definition of event to object state changes: indeed, within the class of unbounded events in our study, some stimuli (e.g., scratch one's hair, pat a teddy bear, etc.) involve no observable change in the affected object at all.

3.4.2 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 (as opposed to strategic or more reflective processing that could not offer a unified interpretation of our findings across three disparate experiments), 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).

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 conceptualized 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; Hard et al., 2006; Zacks, 2004; Zacks & Tversky, 2001). In other cases, whether an event is construed as bounded or not might depend on one's situational knowledge (Filip, 2001). Future research needs to address how perceptual information and higher-level contextual knowledge are integrated in determining the boundedness profile of an event.

3.4.3 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.

Chapter 4

CHILDREN TRACK THE INTERNAL TEMPORAL STRUCTURE OF EVENTS

4.1 Introduction

The previous two chapters demonstrate that adults are sensitive to the homogeneity property in event temporal structure, and care about whether an event has an inherent endpoint. In this chapter, we ask whether young children track the internal temporal structure of events.

Inspired by adult models of event segmentation (Newtson, Engquist, & Bois, 1977; Magliano, Miller, & Zwaan, 2001; Zacks, 2004; Zacks, Speer, Swallow, Braver, & Reynolds, 2007), developmental research has shown that infants and young children use multiple cues to segment the continuous flow of action into discrete evens (Hespos, Saylor & Grossman, 2009; Meyer, Baldwin, & Sage, 2011; Stahl, Romberg, Roseberry, Golinkoff, & Hirsh-Pasek, 2014; Baldwin, Baird, Saylor, & Clark, 2001; Saylor, Baldwin, Baird, & Labounty, 2007; Pace, Carver, & Friend, 2013).

A key finding from the literature on event cognition is that endpoints are critical event components for both children and adults. For instance, infants' perception of action is guided by whether the endpoint of action coincides with the achievement of a goal (Henrichs, Elsner, Elsner, Wilkinson & Gredebäck, 2014; Woodward, 1998, 1999; Csibra, 2008; Csibra, Bíró, Koós, & Gergely, 2003).

Relatedly, in both memory and language, both children and adults encode the goal of motion events (i.e., the physical endpoint of a motion, as in *A girl walked into the classroom*) more accurately compared to the source (*A girl walked from her home*; Lakusta & Landau, 2005, 2012; Lakusta, Spinelli & Garcia, 2017; Papafragou, 2010; Regier & Zheng, 2007). More generally, in adults, event boundaries are privileged in memory and have a larger influence on understanding and describing events compared to other time-points (Boltz, 1995; Swallow, Zacks, & Abrams, 2009; for a review, see Radvansky & Zacks, 2017).

Despite the importance of event endpoints for event cognition, what counts as an endpoint (and thus an event itself) is not straightforward. Events have different internal structures and different ways in which they can be said to end. Currently, the most prominent event segmentation accounts in both adult and developmental studies offer a mechanism for chunking dynamic input from experience into discrete event units but do not address the representational content of each event unit (including what counts as an event endpoint). Here we propose that the way language encodes event structure provides researchers with a novel, powerful framework for defining events and their endpoints and opens up new ways of understanding how learners (and adults) construe the temporal profile of events.

4.1.1 Internal structure of events in language

Language distinguishes between two types of events that have different internal temporal structures and differ in how they come to an end (Bach, 1986;

Jackendoff, 1991; Vendler, 1957; see Filip, 2012 for an overview). *Bounded* events (e.g., *A girl drew a balloon*) can be considered as developments towards a "built-in terminal point" (Comrie, 1976), "climax" (Vendler, 1957) or "culmination" (Parsons, 1990). The endpoint of a bounded event is inherent: it can be predicted from the outset and realized if the event is not interrupted (Mittwoch, 2013). In our example, the event of drawing a balloon ends when a picture of a balloon comes into being. By contrast, *unbounded* events (e.g., *A girl drew circles*) have a largely undifferentiated internal structure and thus can terminate arbitrarily. The endpoint of such events is unspecified (e.g., the event described by the utterance *A girl drew circles* could end at any moment).[3] In this sense, bounded events can be considered "atomic" or individuated (Bach, 1986), and unbounded events unspecified with respect to these dimensions.

One way of characterizing the difference in the temporal structure of bounded and unbounded events is given by the semantic property of homogeneity (Hinrichs, 1985; Krifka, 1998). Bounded events have a non-homogeneous structure: the developments from the beginning toward the climax cannot be alike (in the example *A girl drew a balloon*, each phase of the event could be a distinct step). Unbounded events have a homogeneous structure: "any part of the process is of the same nature as the whole" (Vendler, 1957, p. 146). In the example *A girl drew circles*, any moment of the girl's action can still be described as an event of drawing circles. Linguistically,

[3] The terms "bounded" and "unbounded" are also known as "telic" and "atelic" in the linguistics literature (e.g., Dowty, 1979; Garey, 1957; Smith, 1991). Boundedness, or telicity, is a semantic property of situation aspect or lexical aspect, i.e., the linguistic encoding of the temporal contour of an event.

the two types of event behave differently. Descriptions of bounded events are congruent with delimited temporal phrases indicating that the inherent endpoint has been achieved within a certain amount of time (*A girl drew a balloon in half an hour*) but incongruent with durative temporal phrases (*??A girl drew a balloon for half an hour*). The opposite holds for unbounded events (*A girl drew circles for half an hour* is fine but not *A girl drew circles in half an hour*; Dowty, 1979; Smith, 1991).[4]

Children show some sensitivity to the linguistic bounded-unbounded contrast (e.g., Wagner, 2006; Van Hout, 2016). Around the age of 2, when children start to talk about their dynamic experience, they tend to describe a bounded event that has reached its inherent endpoint (e.g., *fell*) or an unbounded event that is still ongoing (e.g., *dancing*), but rarely encode an ongoing bounded event (e.g., *falling*) or the termination of an unbounded event (e.g., *danced*) (Labelle, Godard, & Longtin, 2002; Li & Bowerman, 1998; Shirai & Andersen, 1995). Furthermore, 3-to-5-year olds give different answers to questions with different aspectual profiles. For instance, when seeing a video of a girl eating up a cookie in three bites, and hearing a question including a bounded event description such as "How many times was the cookie eaten?", children at age 3 would count how many inherent endpoints had been achieved and answer "One"; if the question involved an unbounded description such

[4] Boundedness is distinct from goal-directedness, since the latter is neither necessary nor sufficient for an event to be bounded. On the one hand, bounded events may happen by accident (e.g., *Mom broke a dish when washing the tableware; Ariel kicked over a bottle while running across the street*). On the other hand, many unbounded events involve an intentional agent doing a goal-directed action (e.g., *Mom walked the dog in the dark; George brushed his teeth after dinner*).

as “How many times did the girl eat?”, children would count each cessation and answer “Three” (Wagner, 2006; Wagner & Carey, 2003; cf. Barner, Wagner & Snedeker, 2008; Wellwood, Hespos, & Rips, 2018). Nevertheless, children do not fully acquire the way boundedness is expressed in their language until school age (e.g., Gentner, 1978; Gropen, Pinker, Hollander & Goldberg, 1991; Hodgoson, 2001; Jeschull, 2007; Ogiela, 2007; Penner, Schulz & Wymann, 2003; van Hout, 1998; for an overview, see van Hout, 2018).

Importantly for present purposes, linguistic aspect is often assumed to rely on pre-linguistic notions of event structure (e.g., Filip, 1993; Folli & Harley, 2006; Malaia, 2014). In support of this view, there is evidence that both adults and children interpret linguistic boundedness as indicating different perspectives on events (Barner, et al., 2008; Malaia, Renaweera, Wilbur, & Talavage, 2012; Strickland et al., 2015; van Hout, 2007, 2016, 2018; Wagner, 2006, 2012; Wagner & Carey, 2003; Wellwood, et al., 2018). In a more direct test of the role of boundedness in event cognition, Ji and Papafragou (in press) showed that adults could identify categories of bounded and unbounded events on the basis of visual stimuli alone. In that study, sensitivity to boundedness emerged even when viewers were prevented from encoding the stimuli linguistically and thus could not rely on the online use of aspectual distinctions (cf. also Sakarias & Flecken, 2019). Nevertheless, at present, little evidence exists about whether children are sensitive to the bounded-unbounded distinction when processing visual events.

4.1.2 Cognitive representation of boundedness in children

We propose that probing children's sensitivity to (linguistically inspired) boundedness categories in non-linguistic event cognition can offer a novel, more precise way of construing events and event boundaries, and thereby enrich developmental theories of event cognition. Currently, the most prominent such theories focus on mechanisms of event segmentation (e.g., Baldwin et al., 2001; Hespos et al., 2009; Meyer et al., 2011; Stahl et al., 2014; Saylor et al., 2007; cf. Radvansky & Zacks, 2014; Zacks & Tversky, 2001) but do not address the representational content within event boundaries or the way segmented events are related to one another to form event types. These accounts were not designed to distinguish different ways that events may come to an end (e.g., an event of drawing a balloon culminates while an event of drawing circles involves mere cessation). Investigating children's cognitive ability to distinguish bounded from unbounded events can advance our understanding of how children track, understand and categorize dynamic input. Most broadly, the cognitive ability to distinguish between bounded and unbounded events is foundational for how children represent temporal entities in the world, just like the cognitive ability to distinguish between objects and substances is foundational for how children represent spatial entities (see also Barner & Snedeker, 2006; Wagner & Carey, 2003; Wellwood, et al., 2018).

Furthermore, understanding the role of boundedness in cognition can lead to novel predictions for how children process unfolding events. As alluded to earlier, the literature on event cognition typically assumes that event boundaries have a privileged

status in memory and provide anchors for later learning and describing (Swallow et al., 2009). On this view, event endpoints, in particular, are critical for how events are represented by both children and adults (see He & Arunachalam, 2016; Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007; Strickland & Keil, 2011). This literature has mostly drawn its examples from bounded events but without explicitly distinguishing such events from their unbounded variants. If (un)boundedness is psychologically real and underlies children's conception of events, however, it follows that event endpoints should be treated differently depending on the nature of the event: endpoints should be privileged only for bounded events (for which they represent an inherent boundary) but not for unbounded events (for which the endpoint indicates simply the moment that the event stopped).

Finally, a better understanding of events and event boundaries can throw light onto the cognitive prerequisites for the acquisition of linguistic aspect (van Hout, 2007, 2016, 2018; Wagner, 2012). Most broadly, evidence that children are sensitive to the boundedness distinction in cognition can provide evidence for the idea that linguistic-semantic representations are homologous with and draw from underlying conceptual representations (Gleitman, 1990; Jackendoff, 1990; Lakusta & Landau, 2005; Papafragou & Grigoroglou, 2019; Pinker, 1989). Even though this idea is a central tenet of modern cognitive and developmental science, it has nevertheless not undergone extensive direct empirical testing.

In the present study, we build on prior work on adult event cognition to investigate cognitive representations of event temporal structure in children in ways

that address the above themes. To our knowledge, this is the first study on how children process the internal temporal profile of events using exclusively non-linguistic tasks. We focus on 4-to-5-year-olds since children at this age are still developing their linguistic knowledge about event boundedness (e.g., Wagner, 2006; Wagner & Carey, 2003; van Hout, 2016). In Experiments 8 and 9, we ask whether 4-5-year-olds distinguish between bounded and unbounded event structure through a category identification task. In Experiment 10, we test the psychological implications of the boundedness distinction by comparing how children at that age process temporal slices (specifically, midpoints and endpoints) of bounded and unbounded events.

4.2 Experiment 8

Experiment 8 explored whether 4-5-year-old children (and adults) represent categories of bounded and unbounded events in cognition using a categorization task similar to that in Ji and Papafragou (in press).

4.2.1 Method

4.2.1.1 Participants

Forty 4-5-year-old children (age range: 4;1-5;4, mean age: 4;9, 19 female, 21 male) and forty adults (age range: 18;5-21;6, mean age: 19;7, 19 female, 21 male) participated in the experiment. Children were recruited from the Early Learning Center and Laboratory Preschool, both of which were affiliated to the University of

Delaware in Newark, DE. These children were mostly from middle-class families and lived near the campus. They were identified as Caucasian (72.5%, *n* = 29), African American (17.5%, *n* = 7), or Asian (10%, *n* = 4). Adults were undergraduates at the University of Delaware and received course credit for participation. The adult sample was 77.5% Caucasian (*n* = 31), 10% African American (*n* = 4), 5% Asian (*n* = 2), and 7.5% unreported (*n* = 3). Data from an additional group of three children were collected but excluded: one child was distracted and did not finish the test; two children exhibited a response bias (they consistently assigned a star to the video stimuli throughout the test). Approval for testing these participants has been obtained from the University of Delaware Institutional Review Board (IRB protocol number: 165481). Our sample size was based on prior adult event categorization studies (Ji & Papafragou, in press; Mereu, Zacks, Kurby, & Lleras, 2014; Strickland & Keil, 2011; Strickland & Scholl, 2015).

4.2.1.2 Stimuli

The same event stimuli that showed16 pairs of bounded and unbounded events as in the training and testing phase of Experiment 1 were used. To ensure that all video stimuli would illustrate the bounded-unbounded contrast, we conducted a norming study with 20 English monolingual adults (age range: 18;0-21;9, mean age: 19;4, 10 female, 10 male) and 20 English monolingual children (age range: 4;4-5;5, mean age: 4;10, 12 female, 8 male). Participants were requested to watch a subset of the video clips and describe what happened. For this norming task, the events in Table 1 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 participants were randomly assigned to one of the two lists. Their descriptions were coded for the verb phrase which was composed of the verb (or verb particle) used to describe the action and the noun phrase used to describe the affected object(s). The verb phrases underwent linguistic *in some time* vs. *for some time* test for boundedness (see Dowty, 1979; Smith, 1991; Vendler, 1957; also Introduction) by 4 research assistants who were native English speakers (2 assistants for each list). Overall coders agreed in 90.6 % of their boundedness judgments. Discrepancies were resolved by the research assistants through discussion. Among adults, as expected, stimuli of bounded events elicited bounded verb phrases 94.4% of the time; those responses included change-of-state verbs or verb particles (e.g., *roll up a towel*) and quantified count noun phrases (e.g., *peel a banana*). Stimuli of unbounded events elicited unbounded verb phrases 85.6% of the time; responses included verbs of activity (e.g., *twist a towel*) or unquantified noun phrases (bare plurals or mass nouns: e.g., *crack/shell peanuts*). As for children, bounded stimuli elicited bounded descriptions 82.5% of the time and unbounded stimuli elicited unbounded descriptions 63.8% of the time. As found in previous literature (Gentner, 1978; Gropen, Pinker, Hollander & Goldberg, 1991), children used verbs denoting bounded events such as *mixing* instead of verbs denoting unbounded events such as *stirring* for unbounded event stimuli.[5]

[5] Children gave a bounded description for unbounded stimuli 29.4% of the time. Incomplete, unintelligible

The video stimuli were arranged in the same way as in Experiment 1.

4.2.1.3 Procedure

Participants were randomly assigned to one of two conditions. In the Bounded condition, a star appeared under the videos of bounded events but not under their unbounded counterparts in the training phase. In the Unbounded condition, the placement of the star was reversed.

In the training phase for both conditions, participants were invited to watch two videos appearing side by side on the screen each time. After each video played, the last frame would remain on the screen. After both videos finished, a star appeared under one of the two videos, and the experimenter said, "Look! This video gets a star!" The participants' task was to figure out what kind of videos could get a star. In the testing phase, participants saw a set of new videos, each played in the center of the screen. After each video, they were asked: "Could the video get a star or not?" Child responses were recorded by a research assistant. Adult participants recorded a *Yes/No* response on an answer sheet.

4.2.2 Results

Results from Experiment 8 are shown in Figure 13. The data from this experiment (and all subsequent experiments) were analyzed using multi-level mixed

---

utterances occurred 6.8% of the time.

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. The binary accuracy data were submitted to a logit model examining the fixed effects of Age Group (Children vs. Adults) and Condition (Bounded vs. Unbounded). Random intercepts were provided for each Subject and each Item. Including Age Group significantly improved the model fit (χ^2 (1) = 29.79, $p < .001$); similarly, Condition significantly contributed to the model fit (χ^2 (1) = 25.60, $p < .001$); last, the interaction between the two predictors also led to a significant increase in the fit (χ^2 (1) = 15.61, $p = .003$).[6] As shown in Table 16, the model revealed an effect of Age Group (β = -0.92, z = -2.76, p = .006): children (M = 67.5%) were less successful in identifying the target event category than adults (M = 86.3%). There was also a significant effect of Condition (β = -0.76, z = -2.24, p = .025): participants were better at forming the category of bounded events (M = 85%) compared to unbounded events (M = 68.8%). Moreover, a significant interaction between Age Group and Condition was found (β = -0.87, z = -2.52, p = .015): adults were more accurate in the Bounded (M = 90.6%) than in the Unbounded Condition (M = 81.9%) (β = -0.78, z = -2.28, p = .022), and the same pattern was observed in children (Bounded: M = 79.4%, Unbounded: M = 55.6%; β = -1.12, z = -4.45, $p < .001$), but the difference between

[6] We also examined whether Source of Boundedness (Action vs. Affect Object), Testing List, as well as the Gender of participants influenced accuracy. We did not have a theoretically-driven hypothesis for the role of any of these factors. Neither these factors nor their interactions with other factors reliably improved the model fit (based on chi-square tests of change in -2 restricted log likelihood, all *ps* > .150). Therefore, they were not included in the final model.

the two conditions was greater in children than in adults. A further analysis showed that, unlike adults, the performance of children in the Unbounded Condition was not above chance levels (t (19) = 1.506, p = .148).

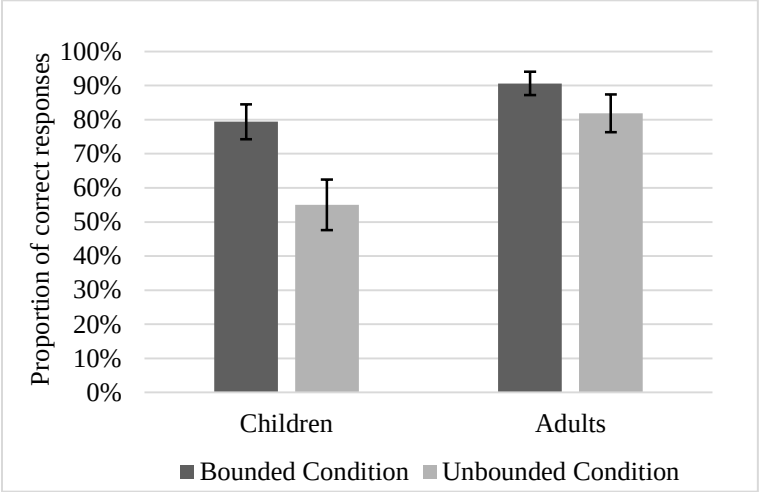

Figure 13 Proportion of correct responses in Experiment 8. Error bars represent ±SEM.

Table 16 *Fixed effect estimates for multi-level model of identifying the target category in Experiment 8.*

Effect	Estimate	SE	z value
(Intercept)	2.27	0.27	8.37***
Age Group (Children vs. Adults)	-0.92	0.33	-2.76**
Condition (Bounded vs. Unbounded)	-0.76	0.34	-2.24*
Age Group * Condition	-0.87	0.42	-2.52*

Note. Formula in R: Acc ~ 1 + (1|Subject) + (1|Item) + AgeGroup + Condition + AgeGroup: Condition
* p < .05, ** p < .01, *** p < .001

4.2.3 Discussion

Experiment 8 led to two major findings. First, 4-to-5-year-old children are sensitive to the internal temporal profile of events: after watching a few contrastive examples of bounded and unbounded events, children could form the bounded event category and extend it to new examples. These results comport with and extend recent work on event cognition with adults (Ji & Papafragou, in press; see also Sakarias & Flecken, 2019; Strickland et al., 2015). Because bounded and unbounded stimuli were equally visually similar to each other and equally intentional (at least according to adults in our norming studies), the ability to place events into classes did not depend on either visual similarity or intentionality.

Second, bounded events were treated differently from unbounded events: both age groups in our sample could identify the category of bounded events with greater ease compared to that of unbounded events, even though the difference between the two event categories was more pronounced in children's performance. In fact, unlike adults, 4-to-5-year olds did not succeed in extracting the unboundedness category (i.e., the class of events that is characterized by the *lack of* an inherent endpoint). This result is reminiscent of a similar advantage for the bounded category in the adult data of Ji and Papafragou (in press).

How can the asymmetry be explained? We propose that a well-defined endpoint acts as an anchor for individuating and comparing bounded events, and helps form a generalization about the bounded category; in other words, the development towards a predictable endpoint in a bounded event structure is easier to track. By

contrast, the *lack* of internal development in an unbounded event structure, together with a less predictable endpoint, makes it harder to track unbounded events. This account is in line with models arguing that significant changes in event features (i.e., moments of culmination within bounded events) are associated with increasing processing activity in the brain of event observers and help with event understanding (Swallow et al., 2009; Zacks et al., 2007). We revisit the bounded-unbounded difference in later sections.

4.3 Experiment 9

An alternative explanation for the patterns in Experiment 8 was that children (and adults) simply tracked whether the actor in the videos completed the depicted action: in bounded events, it was clear that the girl was done since the inherent endpoint was always realized; in unbounded events (that lacked an inherent endpoint), it was not clear whether the girl finished her action. Completion (realizing an endpoint) and boundedness (having an inherent endpoint, whether realized or not) are different notions (Dahl, 1981; Ji & Papafragou, in press). To exclude this possibility, Experiment 9 had the same training phase as Experiment 8; however, in the testing phase half of the videos showed an entire event, as in the previous experiment (full videos), and the other half only showed the very beginning of the event (truncated videos). Full and truncated videos differed in terms of whether the events were actually completed but not in terms of whether the events were bounded. If children (and adults) in the Bounded Condition based their conjectures on event completion

after watching full videos during training, they should only accept full videos as members of the target category at test. However, if participants based their conjectures on whether the event had a potential, but not necessarily realized, endpoint, they should generalize the target category to both full and truncated videos.

4.3.1 Method

4.3.1.1 Participants

A new group of forty 4-to-5-year old children (age range: 4;1-5;5, mean age: 4;9, 22 female, 18 male) and forty adults (age range: 18;6-21;4, mean age: 19;4, 19 female, 21 male) participated in the experiment. Children were recruited at the same local preschools as children in Experiment 8. They were identified as Caucasian (75%, n = 30), African American (12.5%, n = 5), or Asian (12.5%, n = 5). Adults were undergraduates at the University of Delaware and received course credit for participation. The adult sample was 80% Caucasian (n = 32), 10% African American (n = 4), 7.5% Asian (n = 3), and 2.5% unreported (n = 1). Data from two additional children were collected but excluded because they did not finish all the questions.

4.3.1.2 Stimuli

In the training phase, videos were the same as in Experiment 8. 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 8 such that only the first 25% of each video remained. As a result,

truncated bounded events were incomplete while truncated unbounded events were unterminated. The full videos consisted of 8 unpaired videos of 4 bounded events (stack 5 cups on the table, put some Q-tips together, erase a star, organize 3 pairs of socks by color) and 4 unbounded events (roll a ball back and forth, grind biscuits, sprinkle pepper, pull a towel). The truncated videos were arranged into 2 lists in the same way as in Experiment 8. Then the 8 full videos were added to each list. Each of the resulting testing lists was composed of 16 videos; truncated and full videos were intermixed. Whether the event was bounded or unbounded and whether the video was truncated or full was counterbalanced.

4.3.1.3 Procedure

The procedure was identical to that of Experiment 8.

4.3.2 Results

Results from Experiment 9 are shown in Figure 14. The binary accuracy data was analyzed with a mixed logit model, examining the fixed effects of Age Group (Children vs. Adults), Condition (Bounded vs. Unbounded) and Video Type (Truncated vs. Full). Inclusion of Age Group significantly improved the model fit (χ^2 (1) = 22.15, $p < .001$); Condition also led to a significant increase in the fit (χ^2 (1) = 30.72, $p < .001$). Video Type did not significantly contribute to the model fit (χ^2 (1) =

0.32, $p > .250$) but its interaction with Condition did (χ^2 (1) = 7.38, $p = .007$).[7] As shown in Table 17, there was a significant effect of Age Group (β = -0.76, z = -5.61, $p < .001$): adults (M = 81.6%) were overall better than children (M = 68.1%). There was also a significant effect of Condition (β = -0.41, z = -2.16, p = .031): performance in the Bounded Condition (M = 81.7%) was better than in the Unbounded Condition (M = 68.0%). Importantly, no significant effect of Video Type was found (β = -0.32, z = 1.30, p = .195): full and truncated videos were treated similarly by participants. In addition, a significant interaction between Video Type and Condition was detected (β = -0.73, z = -2.69, p = .007): for full videos, there was a better performance in the Bounded (M = 79.4%) than the Unbounded condition (M = 72.2%) (β = -0.41, z = -2.13, p = .033), but the effect of Condition was greater for truncated videos (Bounded condition: M = 84.1%, Unbounded condition: M = 63.8%; β = -1.14, z = -5.83, $p < .001$).

Both adults' and children's performance was significantly above chance level across combinations of conditions and video types (all *ps* < .001), with the exception of children's performance with truncated videos in the Unbounded condition (t(19) = .82, $p > .250$).

[7] We also examined non-theoretically driven factors – testing list and gender. Neither these factors nor their interactions with other factors reliably improved the model fit (all *ps* > .230). Therefore, they were not included in the final model.

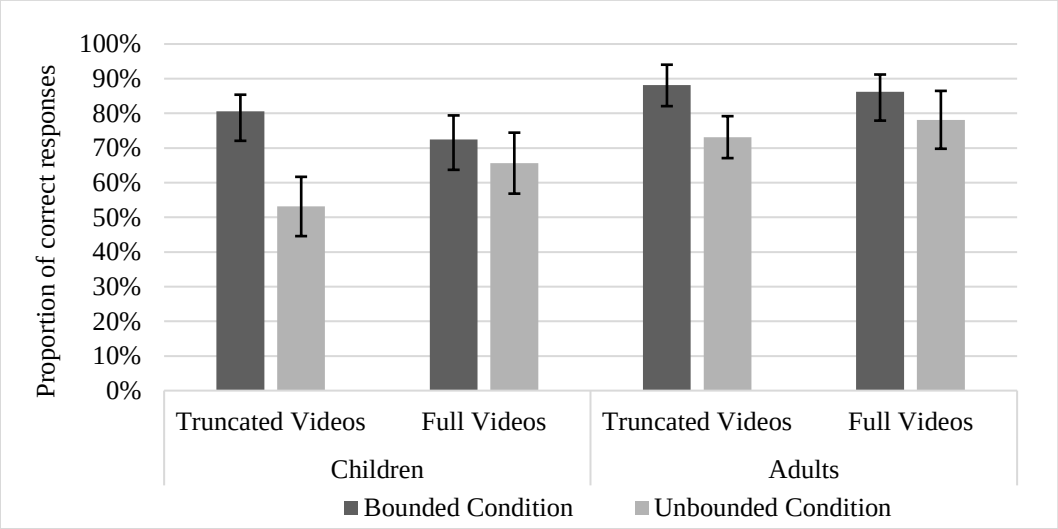

Figure 14 Proportion of correct responses in Experiment 9. Error bars represent ±SEM.

Table 17 *Fixed effect estimates for multi-level model of identifying the target category in Experiment 9.*

Effect	Estimate	SE	z value
(Intercept)	1.79	0.19	9.44***
Age Group (Children vs. Adults)	-0.76	0.14	-5.61***
Condition (Bounded vs. Unbounded)	-0.41	0.19	-2.16*
Video Type (Truncated vs. Full)	0.32	0.25	1.30
Condition*Video Type	-0.73	0.27	-2.69**

Note. Formula in R: Acc ~ 1 + (1|Subject) + (1|Item) + AgeGroup + Condition + VideoType + Condition: VideoType
Only interactions that significantly improved model fit were included in the final model and reported here.
* $p < .05$, ** $p < .01$, *** $p < .001$

4.3.3 Discussion

The results of Experiment 9 suggest that both 4-to-5-year-old children and adults can rely on boundedness independently of completion when categorizing events. Here, unlike Experiment 8, children could form the category of unbounded events

(perhaps because the testing phase was longer). The interaction between Video Type and Condition suggests that there was a cost of generalizing (un)boundedness to a new type of video stimuli (i.e., the truncated videos), particularly when unbounded events were the target category. As in Experiment 9, there was a bounded-unbounded asymmetry in performance. We conclude that boundedness is not a specifically linguistic property but is present in the (non-linguistic) cognitive representation of events.

4.4 Experiment 10

Experiments 8 and 9 showed that 4-to-5-year-old children regard bounded and unbounded events as different categories. Experiment 10 tested the psychological consequences of (un)boundedness for the way children process dynamic events, especially in the context of previous literature arguing that event endpoints are particularly salient compared to earlier points in the event timeline (e.g., Lakusta & Landau, 2005, 2012; Papafragou, 2010; Strickland & Keil, 2011; Zacks & Swallow, 2007). Specifically, we introduced interruptions at different time points during bounded vs. unbounded events and examined children's reactions to these interruptions. We adopted a variant of the "picky puppet task" (Waxman & Gelman, 1986). During training, 4-to-5-year-old children and adults watched videos containing an interruption that blurred either the endpoint or the midpoint of an event, and were introduced to a picky girl who liked only one of these blurred versions but not the

other. At test, participants were asked whether the girl would like new videos containing mid- or end-interruptions.

We expected that the ability to learn the picky girl's (ostensibly arbitrary) preference would be driven by the 4-to-5-year-olds' (and adults') own biases when processing bounded vs. unbounded events. Recall that the two types of events differ in terms of homogeneity (e.g., Vendler, 1957): in bounded events, a non-homogeneous internal structure leads to a climax but in bounded events, no such finely differentiated internal structure exists (i.e., event development is homogeneous). We therefore hypothesized that, for bounded events, both 4-to-5-year-olds and adults would have different reactions to the girl's preference for midpoint vs. endpoint interruptions; furthermore, for both groups, it would be harder to accept that she prefers endpoint compared to midpoint interruptions, since endpoint interruptions occlude the climax of the event. For unbounded events that have a largely undifferentiated internal structure, however, the endpoint-midpoint difference should diminish or disappear altogether.

4.4.1 Method

4.4.1.1 Participants

A new group of eighty children (age range: 4;0-5;6, mean age: 4;9, 34 female, 46 male) and eighty adults (age range: 18;1-21;5, mean age: 19;5, 40 female, 40 male) participated in the experiment. Participants were recruited from the same populations as in previous experiments. Child participants were identified as Caucasian (76.3%, *n* = 61), African American (21.3%, *n* = 17), or Asian (2.4%, *n* = 2); adult participants

identified their race as Caucasian (73.8%, *n* = 59), African American (18.8%, *n* = 15), Asian (6.2%, *n* = 5), or American Indian (1.2%, *n* = 1). Data from 5 additional children and 1 adult were collected but excluded because their responses consisted exclusively of *Yes* responses.

4.4.1.2 Stimuli

The video stimuli were composed of twenty bounded events and twenty closely paired unbounded events (see Table 18). The events were similar to those used in previous experiments. Paired bounded-unbounded videos (i.e., each row in Table 18) had the same duration (range: 4.5-13s, M = 7.2s), and showed events that differed in either the action (e.g., open a towel vs. twist a towel), or the affected object (e.g., paint a heart vs. paint stuff). 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 clothes of a different color.

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: 0.9-2.6s). 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 15). Each video was edited twice, once to create a mid-interruption and once to create an end-interruption. The mid-

interruption was centered on the video midpoint, while the end-interruption blocked the last 20% of the video.

Table 18 *Event stimuli in Experiment 10. Paired videos appear in the same row.*

Phase	No.	Bounded Events	No.	Unbounded Events
Training	1	paint a heart	21	paint stuff
	2	draw a fish	22	draw circles
	3	drink up a cup of juice	23	drink/sip water
	4	cover a slice of bread with peanut butter	24	cover slices of bread with peanut butter
	5	cut a ribbon in half	25	cut ribbon from a roll
	6	peel an orange	26	crack peanuts
	7	open a towel	27	twist a towel
	8	fill a glass with milk	28	shake a bottle of milk
Testing	9	eat a candy	29	eat cheerios
	10	break a biscuit (into two pieces)	30	grind biscuits
	11	blow a balloon	31	blow bubbles
	12	tear a paper towel	32	tear slices off paper towels
	13	tie a knot	33	tie knots
	14	clean a mirror	34	clean/brush teeth
	15	make a letter with Ketchup on a plate	35	squeeze Ketchup on a plate
	16	roll up a towel	36	pull a towel
	17	fold up a handkerchief	37	wave a handkerchief
	18	crack an egg	38	beat an egg
	19	scoop up yogurt	39	stir yogurt
	20	stack 5 cups on the table	40	roll a cup back and forth

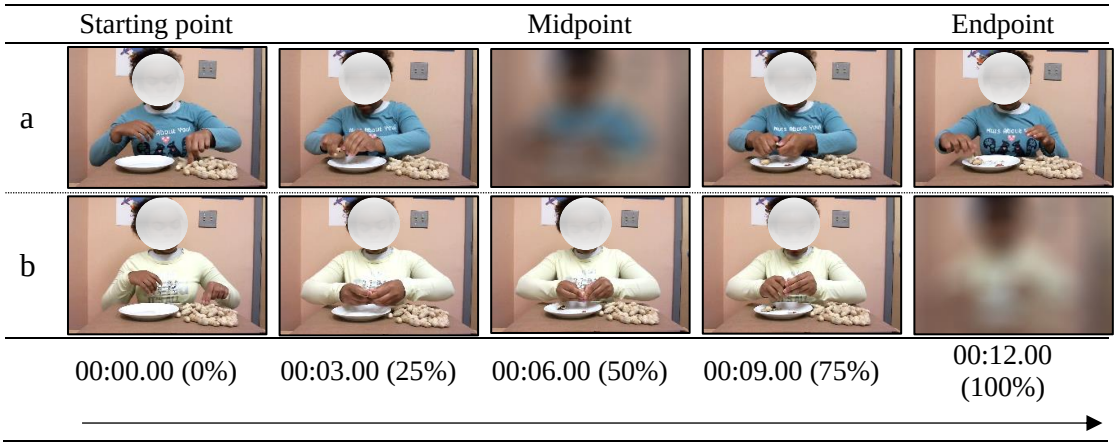

Figure 15 Example of a training trial for an unbounded event (crack peanuts) that includes the two versions of the event: (a) mid-interruption (actor in blue shirt), (b) end-interruption (actor in yellow shirt). The still-frame images were extracted from the video stimuli at different time-points corresponding to different stages (0-100%) of the unfolding event.

4.4.1.3 Procedure

Participants were randomly assigned to one of two Event type conditions depending on the event type (Bounded or Unbounded) that they were exposed to throughout the experiment. The experiment was composed of a training and a testing phase. At the beginning of the training phase, the experimenter invited participants to watch a couple of videos and told participants that the girl in the 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 watched a total of 16 videos. These were comprised of 8 events (No. 1-8 in Table 18 for bounded events, or 21-28 for unbounded events, presented in a random order),

each with two versions shown in succession in the center of the screen. The two versions differed in terms of both the actor's clothes color and in terms of the placement of the interruption (mid-interruption vs. end-interruption, see Figure 15 for an example). Within this phase, half of the time mid-interruptions occurred with one clothes color and the rest of time they occurred with the other color. Even though our hypothesis targeted detection of a mid- vs. end-interruption, we added the change of clothes color 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 clothes colors, was counterbalanced.

After each version the experimenter said either, "The girl likes the video", or "The girl doesn't like the video". To highlight the picky girl's attitude towards the video, a smiley face or an angry face appeared on the screen respectively. Within the Bounded and Unbounded group, participants were randomly assigned to one of two Preference conditions. In the "Likes mid-interruption" condition, they always heard that the picky girl liked the video and saw a smiley face after mid-interruptions, but heard that the girl did not like the video and saw an angry face 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 12 videos showing 12 new events (No. 9-20 in Table 18 for bounded events, or 29-40 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 clothes color was counterbalanced for that event across lists. Types of interruptions and clothes color changes were evenly split within each list. After watching each video, participants were asked: "Will the girl like this video or not?" Child responses were recorded by a research assistant. Adult participants put down a *Yes/No* response on an answer sheet.

4.4.2 Results

Correct responses indicated that participants could identify the type of interruptions that the picky girl liked. For instance, in the "Likes mid-interruption" condition, the correct answer was *Yes* for a video with a mid-interruption, and *No* for a video with an end-interruption. If participants failed to connect the girl's preference with the placement of the interruption, the proportion of the correct responses would be around chance level. Furthermore, if participants based their conjecture on the color of the girl's clothes, they should give a correct response around half of the time (since color was counterbalanced). As shown in Figure 16, for both children and adults, the proportion of correct responses significantly differed from chance level in both Preference conditions and across both Event types (all $ps < .01$). This suggests that, overall, participants could track the placement of interruptions as expected and were not distracted by other factors such as the change in the color of the girl's clothes.

The binary accuracy data were analysed with a mixed logit model examining the fixed effects of Age Group (Children vs. Adults), Preference (Likes mid-interruption vs. Likes end-interruption), and Event Type (Bounded vs. Unbounded). Adding Age Group and Preference significantly improved the model fit (Age Group: χ^2 (1) = 88.66, $p < .001$; Preference: χ^2 (1) = 5.06, $p = .024$). Inclusion of Event Type did not lead to a significant increase in the model fit (χ^2 (1) = 0.14, $p > .250$), but its interaction with Preference did (χ^2 (1) = 10.44, $p = .001$).[8] As shown in Table 19, the model revealed an effect of Age Group (β = -1.57, z = -9.69, $p < .001$); adults (M = 91.0%) were overall better at figuring out the picky girl's preference than children (M = 70.8%). There was also a significant effect of Preference (β = -0.99, z = -4.47, $p < .001$): participants were more accurate when the picky girl liked mid-interruptions (M = 81.5%) compared to end-interruptions (M = 76.3%). Moreover, there was a significant effect of Event Type (β = -0.47, z = -2.17, $p = .030$): participants watching videos of bounded events (M = 80.1%) performed better than those watching videos of unbounded events (M = 78.1%). Importantly, a significant interaction between Preference and Event Type was detected (β = 0.98, z = 3.37, $p < .001$): participants exposed to bounded events were better at identifying a preference for mid-interruptions (M = 86.7%) compared to a preference for end-interruptions (M = 73.5%) (β = -2.30, z = -4.16, $p < .001$); by contrast, participants exposed to unbounded events

[8] We also examined whether Clothes Color, Placement of the Interruption (Mid vs. End) and List would influence accuracy. Adding these non-theoretically driven factors to the model did not reliably improve the fit (all *ps* > 250). Therefore, they were not included in the final model.

were equally good at identifying that the girl liked mid-interruptions (M = 78.1%) and end-interruptions (M = 78.2%) (β = -0.15, z = -0.44, p > .250).

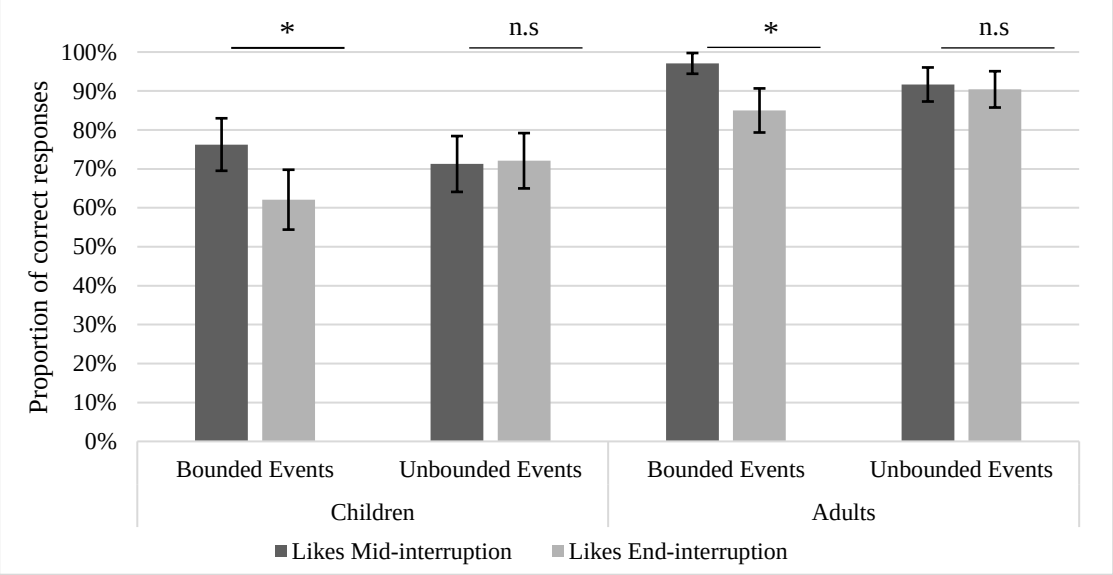

Figure 16 Proportion of correct responses in Experiment 10. Error bars represent ±SEM.

Table 19 *Fixed effect estimates for the multi-level model of accuracy in identifying the preference for interruptions in Experiment 10.*

Effect	Estimate	SE	z value
(Intercept)	2.99	0.22	13.90***
Age Group (Adults vs. Children)	-1.57	0.16	-9.69***
Preference (Likes Mid- vs. End-interruption)	-0.99	0.22	-4.47***
Event Type (Bounded vs. Unbounded)	-0.47	0.22	-2.17*
Preference * Event Type	0.98	0.29	3.37***

Note. Formula in R: Acc ~ 1 + (1|Subject) + (1|Item) + AgeGroup + Preference + Event Type + Preference: Event Type
Only the interactions that significantly improved the model fit were included in the final model and reported here.
* p < .05, ** p < .01, *** p < .001

4.4.3 Discussion

The data from Experiment 3 show that 4-to-5-year-old children (as well as adults) processed bounded and unbounded events differently. Most obviously, as in previous experiments, children exposed to bounded event stimuli performed better than those exposed to unbounded event stimuli. More crucially for present purposes, children had more difficulty learning that someone liked end-interruptions compared to mid-interruptions of bounded events, but no such endpoint-midpoint difference was found for unbounded events. These results support the hypothesis that the boundedness profile of an event (specifically, its degree of homogeneity) has consequences for how the event and its various temporal stages are apprehended. Furthermore, these results raise the possibility that, during the event categorization task of Experiments 8 and 9, children (and adults) relied on the property of homogeneity to form the class of unbounded events and, conversely, the property of non-homogeneity to form the class of bounded events.

4.5 General Discussion

Our findings show that 4-to-5-year-old children track the internal temporal structure of events, i.e., they are sensitive to the distinction between bounded events (those with a non-homogeneous structure leading toward an inherent moment of culmination) and unbounded events (those with a homogeneous structure without a specified endpoint) (Experiment 8). Furthermore, children at this age are able to treat boundedness (i.e., the availability of an inherent endpoint) as different from

completion (i.e., the actual realization of the endpoint) (Experiment 9). Lastly, children treat temporal landmarks such as midpoints and endpoints differently depending on the homogeneity of the event structure – they react differently to midpoint and endpoint interruptions in a bounded (non-homogeneous) event, but not in an unbounded (homogeneous) event (Experiment 10). Below we discuss implications of these results for early event cognition and language.

4.5.1 Delimiting events and their boundaries

Our results bear on the mechanisms and development of event cognition, since they suggest that young children's event representations (including the precise nature of event endpoints) integrate a highly abstract construal of event temporal structure. Our findings show continuity in mechanisms of event cognition, since the shape of results was overall similar between 4-to-5-year-olds and adults despite lower accuracy in children's performance (cf. Ji & Papafragou, in press).

Unlike past developmental work on event segmentation that has focused on how children segment the continuous flow of action into discrete evens (Hespos et al., 2009; Meyer et al., 2011; Stahl et al., 2014; Baldwin et al., 2001; Saylor et al., 2007; cf. Radvansky & Zacks, 2014; Zacks & Tversky, 2001), the present approach reveals children's generalizations about the representational content of individual events and the way events are related to one another. We show that 4-to-5-year-old children are able to form an event *category* across visually dissimilar events (e.g., drawing a balloon and rolling up a towel vs. drawing circles and twisting a towel) based on the

way these events unfold over time and come to an end; they can further generalize the category to examples where the endpoint is not visible. 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 (given the visual variability of our stimuli) nor is it identical to higher-level conceptual features such as intentionality (since videos of both bounded and unbounded events depicted intentional actions to the same degree; see norming data in Experiment 1). As already alluded to, we take the cognitive ability to distinguish between bounded and unbounded events to be foundational for how children represent temporal entities in the world, just like the cognitive ability to distinguish between objects and substances is foundational for how children represent spatial entities.

The parallel between events and objects, long noted in both the linguistics and philosophical literature (see Bach, 1986; Jackendoff, 1991; cf. also Barner & Snedeker, 2006; Wagner & Carey, 2003; Wellwood et al., 2018), can help explain one of our most robust findings across all three experiments: namely, the consistent advantage of bounded over unbounded events for both children and adults (cf. also Ji & Papafragou, in press). We propose that the advantage of the bounded event category can be attributed to atomicity, a property that bounded events share with objects (cf. Introduction). Bounded events (expressed by bounded verb phrases such as *peel a banana*), just like objects (expressed by count nouns such as *a peanut*), are naturally individuated "atoms". Similarly to objects that are measured by counting their number

(e.g., one can count how many peanuts there are in a jar), bounded events can be quantified by counting the number of inherent endpoints that have been achieved (e.g., one can measure the occurrence of banana-peeling events by counting how many bananas have been peeled; Barner et al., 2008; Wagner & Carey, 2003). Unbounded events (expressed by unbounded verb phrases such as *crack peanuts*), just like substances (expressed by mass nouns such as *sugar*), are unspecified for atomic features. Neither unbounded events nor substances can be directly counted; additional criteria are needed to turn these entities into something countable (e.g., sugar can be measured and compared by weight; cracking peanuts can be quantified by the duration of the activity, among other possibilities). Our data show that, across development, temporal entities that are considered individuated or "atomic" (bounded events) behave differently from temporal entities that are unspecified for these features (unbounded events) – for instance, they are more easily tracked and abstracted over.

4.5.2 Processing event midpoints and endpoints

Previous literature on event cognition has converged on the finding that children – as well as adults - encode endpoints as a critical event component in both language and memory (e.g., Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007; Strickland & Keil, 2011; Zacks & Swallow, 2007). The present work expands on and offers a framework for explaining this finding by showing that the role of event endpoint shifts depending on the overall event architecture (see especially Experiment 3). Specifically, the salience of endpoints is connected with the non-

homogeneous structure of bounded events; in unbounded events that have a homogeneous structure, endpoints are treated largely similarly to other temporal pieces such as midpoints. Thus the internal temporal contour of events determines how different event segments are processed as the events unfold.

We should emphasize that, even though our stimuli were constructed to be readily perceived in either bounded or unbounded terms, the same experience can often be construed from both a bounded and an unbounded perspective (e.g., you can view a single event as either cracking five peanuts or cracking peanuts; Wagner, 2009). Thus it is the learner's mind that applies boundedness categories to (sometimes identical) streams of sensory information. The present perspective predicts that midpoints and endpoints *even for the same dynamically unfolding input stream* should be processed differently depending on whether the event is conceptualized as bounded or unbounded.

4.5.3 (Un)boundedness in early cognition and language

Finally, our results demonstrate that boundedness is not simply a linguistic property but a foundational conceptual distinction that represents a basic commitment to the kind of event a young learner has in mind. Early sensitivity to boundedness could lay the foundations for the acquisition of aspect in language (Filip, 1993; Wagner, 2009). In that sense, our results support the claim that event language is parallel to and draws from event cognition (see also Cohn, Paczynski, & Kutas, 2017; Jackendoff, 2007; Folli & Harley, 2006; Hafri, Papafragou, & Trueswell, 2013;

Lakusta & Landau, 2005; Malaia, 2014; Papafragou, 2015; Strickland et al., 2015; Tversky, Zacks, Morrison, & Hard, 2011).

Given that boundedness is encoded in different ways across languages (Bar-el, Davis, & Matthewson, 2005; Botne, 2003), and that these differences affect the acquisition of aspect (Hacohen, 2012; Van Hout, 2007), it remains to be seen how children's non-linguistic boundedness generalizations interface with the development of aspectual distinctions. Relatedly, at present, it is an open question whether boundedness could be computed by learners without exposure to a conventional linguistic system (e.g., deaf homesigners; Feldman, Goldin-Meadow, & Gleitman, 1978; Zheng & Goldin-Meadow, 2002).

4.5.4 Summary

We conclude that young children are sensitive to foundational and abstract properties of event temporal structure in ways that align with the encoding of the temporal profile of events in language. Such sensitivity to the internal temporal profile of events can further support the way children conceptualize and process their dynamic experience.

Chapter 5

CONCLUSIONS

Events can be defined as temporal segments that are considered by observers to have a beginning and an end (Zacks & Tversky, 2001). This dissertation explores whether people are sensitive to the structure of these temporal segments, particularly the temporal structure expressed by event language. The point of departure in our studies is the linguistic distinction between bounded events that can be considered as non-homogeneous developments leading to an inherent endpoint and unbounded events that are homogeneous temporal entities without an inherent endpoint. We examined whether boundedness makes a difference in non-linguistic representations of events. Our findings show that bounded and unbounded events are treated as different categories by event viewers (Chapter 2). Moreover, (un)boundedness, particularly, the (non)homogeneity property in the event temporal structure affects how people process and weigh temporal slices of different events (Chapter 3). The sensitivity to abstract properties of event temporal structure is also found in 4-to-5-year-old children, although they have not fully acquired linguistic expressions of bounded and unbounded events at this age.

Our findings contribute to two separate fields of inquiry. Within the field of event cognition, our findings provide direct evidence for the psychological reality of the boundedness distinction in conceptual structure, a finding that speaks to the basic

building blocks of cognitive event representation. Our results inform theories of event segmentation by characterizing the temporal structure within an event and distinguishing different types of event boundaries. The homogeneity of the event structure bears on Event Segmentation theory that links event boundaries to low predictability of upcoming happenings (e.g., Zacks & Swallow, 2007; Zacks et al., 2009). Our data indicate that in a non-homogeneous event structure, the ending is less predictable, providing more information for identifying the event category and comprehending the event. Our results go beyond Event Segmentation theory by taking the *nature* of event boundaries into account. Event observers treat the achievement of an inherent endpoint as different from mere cessation. The abstract properties of boundedness and homogeneity in event structure have further implications for event processing. Our results explain and place constraints on the privileged status of event boundaries found in a number of prior studies (e.g., Lakusta & Landau, 2005, 2012; Papafragou, 2010; Strickland & Keil, 2011; Zacks & Swallow, 2007).

Within psycholinguistics, the findings that viewers are sensitive to the linguistically-inspired notion of boundedness provide new evidence for the presence of homologies between event representations in language and cognition (e.g., Folli & Harley, 2006; Malaia, 2014; Tversky, et al., 2011). An important direction for future research is to examine the nature of the connection between event language and event cognition. According to the universalist view, the notions "bounded" and "unbounded" are cognitive primitives and linguistic distinctions are built upon these concepts (e.g., Filip, 1993; see also Pinker, 1989; Papafragou, 2015; Shipley & Zacks, 2008). By

contrast, according to the linguistic relativity hypothesis, the way language encodes boundedness may influence, or even shape the way that people observe and distinguish between bounded and unbounded events (see Sapir, 1924; Whorf, 1956 for the relativist proposal and Flecken, von Stutterheim, & Carroll, 2014; von Stutterheim & Nüse, 2003 for some empirical evidence). To test these possibilities, we can examine whether viewers who speak different languages encode the bounded-unbounded contrast cognitively in similar ways. Due to the outbreak of the pandemic, data collection with Mandarin-speaking adults and children is ongoing. The results are expected to show whether language-specific properties of expressing boundedness would affect non-linguistic representations of bounded vs. unbounded events.

Another open question concerns the components that determine event boundedness. We designed our stimuli based on two sources, i.e., the nature of the action and the affected object. We did not plan to compare between the two sources in the present experiments, but it is worthwhile to figure out which component could be more important for the presence of an inherent endpoint. Within linguistic studies of aspect, it has been proposed that predicate telicity (elements related to the verb, such as particles as in *fold up*) is easier to acquire compared to compositional telicity which requires projecting the quantificational properties of the affected object onto the event (van Hout, 2016). This proposal is supported by many studies showing that children as young as 3 know that telicity is expressed by particles but sensitivity to quantization of the object appears relatively late (e.g., Jeschull, 2007; Ogiela, 2007; van Hout, 1998, 2007). One might expect that, within the nonlinguistic apprehension of events, the

nature of the action might similarly offer a more straightforward path to boundedness compared to the nature of the affected object. To test this hypothesis, we are running a version of our category identification task where people have to rely on only one source to extract the bounded-unbounded contrast. If the hypothesis is correct, people watching contrastive events in the nature of the action during training would have a better performance compared to those watching contrastive events in the nature of the affected object.

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.

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

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., & Snedeker, J. (2005). Quantity judgements and individuation: Evidence that mass nouns count. *Cognition, 97,* 41-46. doi: 10.1207/s15473341lld0203_2

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.

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

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

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.

Cohn, N., Paczynski, M., & Kutas, M. (2017). Not so secret agents: Event-related potentials to semantic roles in visual event comprehension. *Brain and Cognition, 119,* 1-9. doi: 10.1016/j.bandc.2017.09.001.

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]

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, 107, 705-717. doi:10.1016/j.cognition.2007.08.001

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively, *Cognitive Science, 27*, 111-133. doi:10.1016/S0364-0213(02)00112-X

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.

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

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

Feldman, H., Goldin-Meadow, S., & Gleitman, L. R. (1978). Beyond Herodotus: The creation of language by linguistically deprived deaf children. In A. Lock (Ed.), *Action, symbol, and gesture: The emergence of language* (pp. 351–414). New York: Academic Press.

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

Filip, H. (2001). Nominal and verbal semantic structure: analogies and interactions. *Language Sciences, 23*, 453-501. doi: 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). Washington, DC: LSA.

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

Flecken, M., von Stutterheim, C., & Carroll, M. (2014). Grammatical aspect influences motion event perception: Evidence from a cross-linguistic non-verbal recognition task. *Language and Cognition*, 6, 45-78.

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.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development, 49,* 988-998. doi: 10.2307/1128738

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition, 1,* 3– 55. doi: 10.1207/s15327817la0101_2

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

Gropen, J., Pinker, S., Hollander, M., Goldberg, R. (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition, 41,* 153-195. doi:10.1016/0010-0277(91)90035-3

Hacohen, A. (2012). On the acquisition of compositional telicity in Hebrew. In E. Labeau, & I. Saddour (Eds.), *Tense, aspect and mood in first and second language acquisition* (pp. 1-25). New York: Rodopi.

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

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

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

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

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

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

Henrichs, I., Elsner, C., Elsner, B., Wilkinson, N., & Gredebäck, G. (2014). Goal certainty modulates infants' goal-directed gaze shifts. *Developmental Psychology, 50,* 100–107. doi: 10.1037/a0032664

Hespos, S. J., Sayor, M., & Grossman, S. R. (2009). Infants' ability to parse continuous actions. *Developmental Psychology, 45*, 575-584. doi: 10.1037/a0014145

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.

Hodgson, M. (2001). The acquisition of Spanish reflexive *se* in transitive verb-argument structures and singularity. In B. Skarabela, S. Fish, & A. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 302-313). Somerville: Cascadilla Press.

Jackendoff, R. (1990). *Semantic structures.* Cambridge, MA: MIT Press.

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.

Jeschull, L. (2007). The pragmatics of telicity and what children make of it. In A. Belikova, L. Meroni, & M. Umeda (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America* (pp. 180-187). Somerville, MA: Cascadilla Proceedings Project.

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.

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). Austin, TX: Cognitive Science Society.

Ji, Y. & Papafragou, A. (in press). Is there an end in sight? Viewers' sensitivity to abstract event structure. *Cognition*. doi: 10.1111/tops.12475

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.

Labelle, M., Godard, L., & Longtin, C. M. (2002). Grammatical and situational aspect in French: a developmental study. *Journal of Child Language, 29,* 301-326. doi: 10.1017/s0305000902005056

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

Lakusta, L., Spinelli, D., & Garcia, K. (2017). The relationship between pre-verbal event representations and semantic structures: The case of goal and source paths. *Cognition, 164,* 174-187. doi: 10.1016/j.cognition.2017.04.003

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.

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, 43,* 916-927. doi: 10.1037/xlm0000364

Li, P., & Bowerman, M. (1998). The acquisition of lexical and grammatical aspect in Chinese. *First Language, 18,* 311-350. doi: 10.1177/014272379801805404

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,* 1904-1933. DOI: 10.1111/cogs.12446

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

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

Meyer, M., Baldwin, D. A., & Sage, K. (2011). Assessing young children's hierarchical action segmentation. In L. Carlson, C. Hoelscher & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Society* (pp. 3156-3161). Austin, TX: Cognitive Science Society.

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*. Dordrecht: Springer.

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

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

Newtson, D., Engquist, G., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology, 35,* 847-862. doi: 10.1016/0022-1031(76)90076-7

Ogiela, D. (2007). *Development of telicity interpretation: Sensitivity to verb-type and*

determiner-type. (Unpublished dissertation). Michigan State University, East Lansing, MI.

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. doi: 10.1016/j.cognition.2019.01.002

Pace, A., Carver, L. J., & Friend, M. (2013). Event-related potentials to intact and disrupted actions in children and adults. *Journal of Experimental Child Psychology, 116,* 453-470. doi: 10.1016/j.jecp.2012.10.013

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.

Penner, Z., Schulz, P. & Wymann, K. (2003). Learning the meaning of verbs: What distinguishes language impaired from normally developing children? *Linguistics*, *41*, 289–319. doi: 10.1515/ling.2003.010

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.

Radvansky, G., & Zacks, J. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences, 17,* 133-140. doi: 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,* 705-719. doi: 10.1080/15326900701399954

Richmond, L. L., & Zacks, J. M. (2017). Constructing experience: Event models from perception to action. *Trends in Cognitive Sciences, 21,* 962-980. doi: 10.1016/j.tics.2017.08.005

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

Sapir, E. (1924). The grammarian and his language. Repr. in Sapir, E. (1949), *Selected writings of Edward Sapir in language, culture, and personality* (pp. 150-159). Ed. by D. Mandelbaum. Berkeley/Los Angeles, CA: University of California Press.

Saylor, M. M., Baldwin, D. A., Baird, J. A. & LaBounty, J. (2007). Infants' on-line segmentation of dynamic human action. *Journal of Cognition and Development, 8,* 113-128. doi: 10.1080/15248370709336996

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.

Shirai, Y., & Andersen, R. (1995). The acquisition of tense-aspect morphology: A prototype account. *Language, 71,* 743-762. doi: 10.2307/415743

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

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M. & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development, 85,* 1821-1826. doi:10.1111/cdev.12247

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

Taylor, B. (1977). Tense and continuity. *Linguistics and Philosophy, 1,* 199-220. doi: 10.1007/BF00351103

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.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review, 94,* 3-15. doi: 10.1037/0033-295X.94.1.3

Van Hout, A. (1998). On the role of direct objects and particles in learning telicity in Dutch and English. In A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh (Eds.), *Proceedings of the 22nd Annual Boston University Conference on Language Development* (pp. 397-408). Somerville: Cascadilla 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

Von Stutterheim, C., & Nüse, R. (2003). Processes of conceptualization in language production: Language-specific perspectives and event construal. *Linguistics, 41*, 831–881.

Wagner, L. (2006) Aspectual bootstrapping in language acquisition: Telicity and transitivity. *Language Learning and Development, 2*, 51-76. doi: 10.1207/s15473341lld0201_3

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). Somerville, MA: Cascadilla Press.

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

Waxman, S. & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1*, 139-156. doi: 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: Oxford University Press.

Whorf, B. L. (1956). *Language, thought and reality.* In J. B. Carroll (Ed.), *Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Wilder, D. A. (1978). Predictability of behaviors, goals, and unit of perception. *Personality and Social Psychology Bulletin, 4,* 604-607. doi: 10.1177/014616727800400422

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

Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34. doi: 10.1016/s0010-0277(98)00058-4

Woodward, A. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development, 22*, 145–160. doi: 10.1016/S0163-6383(99)00007-7

Zacks, J. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 26,* 979-1008. doi: 10.1207/s15516709cog2806_5

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

Zheng, M., & Goldin-Meadow, S. (2002). Thought before language: How deaf and hearing children express motion events across cultures. *Cognition, 85,* 145–175. doi: 10.1016/S0010-0277(02)00105-1

Appendix A

IRB APPROVAL FOR RESEARCH ON HUMAN SUBJECTS

UNIVERSITY OF DELAWARE
RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: April 5, 2016

TO: Anna Papafragou
FROM: University of Delaware IRB

STUDY TITLE: [165481-13] The interface between spatial cognition and language
IRB REFERENCE #:
SUBMISSION TYPE: Continuing Review/Progress Report

ACTION: APPROVED
APPROVAL DATE: April 5, 2016
EXPIRATION DATE: April 7, 2017
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # *(7)*

Thank you for your submission of Continuing Review/Progress Report materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: April 5, 2017

TO: Anna Papafragou
FROM: University of Delaware IRB

STUDY TITLE: [165481-16] The interface between spatial cognition and language

SUBMISSION TYPE: Continuing Review/Progress Report

ACTION: APPROVED
APPROVAL DATE: April 5, 2017
EXPIRATION DATE: April 7, 2018
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # *(7)*

Thank you for your submission of Continuing Review/Progress Report materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: April 2, 2018

TO: Anna Papafragou
FROM: University of Delaware IRB

STUDY TITLE: [165481-19] The interface between spatial cognition and language

SUBMISSION TYPE: Amendment/Modification

ACTION: APPROVED
APPROVAL DATE: April 2, 2018
EXPIRATION DATE: April 7, 2019
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # *(7)*

Thank you for your submission of Amendment/Modification materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE: March 19, 2019

TO: Anna Papafragou
FROM: University of Delaware IRB

STUDY TITLE: [165481-22] The interface between spatial cognition and language

SUBMISSION TYPE: Continuing Review/Progress Report

ACTION: APPROVED
APPROVAL DATE: March 19, 2019
EXPIRATION DATE: April 7, 2020
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # *(7)*

Thank you for your submission of Continuing Review/Progress Report materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

Appendix B

REPRINT PERMISSION FOR CHAPTER 2