The title

Jessica E. Kosie[1] & Dare Baldwin[2]

[1] Department of Psychology, Princeton University

[2] Department of Psychology, University of Oregon

Author Note

Correspondence concerning this article should be addressed to Jessica E. Kosie, Princeton, NJ 08540 USA. E-mail: jkosie@princeton.edu

Abstract

XXX abstract goes here.

Keywords: keywords

Word count: X

The title

Introduction

Our everyday experience of the world is one of events. For example, a simple trip to the grocery store might involve parking the car, entering the store, putting groceries into a cart, paying the cashier, and loading up the trunk. Segmenting action into these individual events - or units of time that occur at a particular location and are perceived to have a beginning and an end (???) - provides information about the structure of unfolding activity. The ability to find structure as activity unfolds is a skill necessary for understanding, remembering, talking about, and responding to events as well as predicting events that will happen in the future. Though event segmentation is typically experienced as an effortless and automatic process that proceeds naturally as part of ongoing perception (???), it is deceptively complex. The input from which we extract discrete events is not pre-segmented into individual units. Instead, visual input streams past rapidly and fluidly, with no systematic pauses or clear indications that one event unit has ended and another has begun. Just as quickly as it appears, event information is gone. Despite the challenge of rapidly extracting structure from dynamically unfolding stimuli, however, event processing fluency appears to be acquired early in development.

By adulthood, observers are remarkably consistent in how they segment events, generally agreeing on the location in an activity sequence at which one unit of action ends and another has begun (Kurby & Zacks, 2008, (???), (???), (???)). We refer to these moments as "event boundaries." Adults' agreement on the location of boundaries is apparent both when they are explicitly asked to mark moments that correspond to event boundaries (???, (???)) and in more implicit indicators of processing, like fMRI and self-paced slideshow viewing (???, (???), (???), (???)). Further, adults' ability to target boundaries is predictive of memory for event content (???), the ability to perform everyday sequences of action (???), and is impaired in disorders such as autism and dementia (???, (???)). Even infants as young as six months old are sensitive to the structure of at least some kinds of events (**???**, (**???**), (**???**), (**???**)). For example, infants are surprised when an action sequence is paused at midstream, but not at event boundaries, suggesting that the ability to segment action emerges early in development. Infants, like adults, have better memory for objects presented at event boundaries over objects presented at non-boundary regions (**???**). Further, infants' memory is impaired for activity sequences viewed with occlusions at event boundaries relative to the same sequences with occlusions occurring at non-boundary regions (**???**). Altogether, there exists a substantial body of evidence that, from a very young age, humans *can* extract event information from rapidly unfolding input, and such event processing fluency has important consequences for learning and memory. Relatively less research has focused on *how* humans acquire skill in processing dynamic events.

Recent work has begun addressing this important gap in understanding. Of particular interest is what kind of information observers use to guide their processing of activity as it unfolds. Across a number of studies, when asked to indicate junctures demarcating event boundaries, people tend to nominate regions that correspond to the initiation and completion of goals (e.g., **???**). For example, one might indicate that a boundary has occurred at the juncture at which an individual has closed the trunk after filling it with groceries, which corresponds to completion of the actor's goal of transferring the groceries from the cart to the trunk. It seems likely that such cognitive expectations about how an actor will achieve their goals provide strong cues to the presence of boundaries. However, event boundaries also frequently occur with large changes in motion (**???**, (**???**)). Thus, lower-level information about changes in motion trajectory might guide observers' perception of boundaries, and it is possible that adults solely rely on his information to guide their processing. Adjudicating between these two accounts enables more comprehensive understanding of how adults process unfolding activity, providing insight into the type of information that facilitates rapid and efficient processing of events.

Existing evidence suggests that, while adults are sensitive to motion changes in unfolding activity streams, they prioritize goal-related information. For example, adults asked to advance at their own pace through slideshows of activity sequences (via the dwell-time paradigm) spend significantly more time looking to (or "dwelling on") slides that correspond to boundary over non-boundary regions (**???**, (**???**), (**???**)). Slide-to-slide pixel change, an indication of changes in motion from one slide to the next, is additionally a significant predictor of increased dwelling. However, slide type (boundary vs. non-boundary) explains significantly more variance than pixel change alone, suggesting that adults' dwell times are responsive to more than just variability in motion trajectory across slides (**???**, (**???**)). Relatedly, tasks which assess adults' statistical learning of activity sequences allow boundaries to emerge only after repeated viewings, such that one and the same slide - with identical pixel change values - should only be treated as a boundary after the observer has had enough time to learn the structure of an event. As predicted, increased attention to boundary slides emerges only after an observer has had sufficient exposure to an event to learn that a particular slide is a boundary (**???**, (**???**)), suggesting that motion change alone cannot possibly explain observed increases in attention to boundary slides. Taken together, these results point to goal structure as a relatively more likely explanation for the increased attention to event boundaries that has been observed across a number of studies with adults.

Much less is known about the development of action processing fluency. On one hand, it might be the case that children, like adults, prioritize goal structure over motion trajectory as they view unfolding activity. A great deal of evidence supports this hypothesis, suggesting that, as early as six months of age, infants have expectations about actors' intentions and goals (Woodward, 1998, (**???**)). For example, once infants have been habituated to an actor reaching for a particular object, they expect the actor to continue reaching for that same object on subsequent trials, even if the location of the object changes. Additionally, infants make predictions about an actor's goals based on behaviors including where the actor looks, preferences they exhibit, and their intended goal in failed attempts at goal completion (**???**), and they are surprised when expected goals are violated (**???**, (**???**)). However, other evidence points to infants' attention to the motion trajectory with which an event unfolds (**???**, (**???**), (**???**)). In one study, infants were habituated to a hand reaching over a barrier for a ball. The barrier was then removed, and the actor either continued reaching up and over for the ball (in the absence of the barrier), or began reaching straight for the ball (which would be the most efficient route to achieve the goal of grasping the ball). Infants looked longer to the ill-formed reach (up and over in the absence of a barrier), suggesting both that: (a) as early as infancy, children are sensitive to the goal of a reaching action, and (b) infants pay attention to and track information about motion trajectory. It seems possible that, because infants and children have less experience with events, they may be less likely to have strong predictions about an actor's goals, and thus may rely more on motion-related cues when processing action.

The infancy research described thus far has used methods including familiarization, looking preference, and habituation/dishabituation, all of which are appropriate for this age group. However, simply asking whether, for example, infants prefer to view actions with pauses at boundaries versus non-boundary regions does not provide nuanced information about the cues that infants are using to guide their processing. Thus, existing measures are not amenable to investigations of the extent to which children prioritize goal structure versus motion trajectory when viewing events. The dwell-time paradigm, recently developed by Hard and colleagues (**???**), does provide this sort of nuanced information, and seems amenable for use with a younger population (e.g., **???**). In the dwell-time task, slideshows are constructed by extracting still frames at regular intervals (e.g., one frame per second) from videos of unfolding activity. Observers click a computer mouse to advance at their own pace through the slideshows, and latency between mouse clicks (i.e., "dwell times") index moment-to-moment changes in attention as activity unfolds. As mentioned previously, this paradigm has been successfully used to measure the extent to which adults' attention tracks goal structure versus motion trajectory as they view sequences of activity (**???**, (**???**)). While this paradigm is not currently amenable for use with infants, preschool-aged children readily engage with such computer-based tasks.

The current study.

Use of the dwell-time paradigm with preschoolers enables us to investigate the extent to which younger children prioritize goal-related or motion-related information as they view unfolding activity sequences. In the current study, preschoolers advanced at their own pace through one of three possible action sequences: (1) an actor reaching over a barrier to retrieve a ball, (2) an actor reaching up and over in the absence of a barrier to retrieve a ball, or (3) an actor reaching straight to retrieve a ball (without a barrier present). Thus, all three methods were equated in goal structure but differed in motion trajectory. If preschoolers are tracking only goal structure, we would predict dwell-time patterns to be equivalent across all three reaches. Specifically, dwell times should increase at the moment the actor achieves the goal of grasping the ball. If preschoolers are tracking motion trajectory, however, we would expect attention to differ across the reaches such that attention to actions involving reaching up and over (e.g., sequences 1 and 2) would elicit similar patterns of dwell time, while the straight reaching action would differ. A third prediction is that preschoolers may be sensitive to the path taken to achieve a goal and might be surprised when the actor reaches up and over in the absence of a barrier. If this is the case, we would predict dwell times to be similar across the reach over a barrier and the straight reach conditions, while dwell times would differ in response to the sequence in which the actor reaches up and over in the absence of a barrier. These results would conceptually replicate previous research with infants (e.g., **???**).

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Participants

Material

Procedure

Data analysis

We used R (Version 3.6.1; R Core Team, 2019) and the R-package *papaja* (Version 0.1.0.9842; Aust & Barth, 2018) for all our analyses.

Results

Discussion

References

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–79.

R Core Team. (2019). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

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

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–79.

R Core Team. (2019). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

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