Dwell times disambiguate preschoolers’ attention to motion versus goal structure as dynamic action unfolds

Jessica E. Kosie1,2 & Dare A. Baldwin2

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, Department of Psychology, Princeton University, Princeton, NJ 08540 USA. E-mail: [jkosie@princeton.edu](mailto:jkosie@princeton.edu)

Abstract

We harnessed Hard, Recchia, & Tversky’s (2011) dwell-time paradigm to disentangle the influence of motion patterns and goal structure on children’s analysis of events as dynamic action unfolds. Eighty-five preschoolers (aged 2.5-4.5 years) advanced at their own pace through one of three slideshows, all depicting an actor reaching toward, grasping, and retrieving a ball. However, motion patterns differed for one slideshow (straight reach) relative to the other two (arcing reach), and one of the arcing reach slideshows depicted a violation of canonical goal-related motion (arcing reach in the absence of any barrier). Dwell-time patterns revealed comparable segmental analyses across all slideshows, confirming children’s tendency to privilege goal structure over motion characteristics in their moment-to-moment analysis of continuously unfolding activity.

*Keywords:* event segmentation, event processing, goal structure, dwell-time paradigm, dynamic action

*Word count:* X

Dwell times disambiguate preschoolers’ attention to motion versus goal structure as dynamic action unfolds

Our everyday experience of the world centers on a sense of discrete events that occur across time. For example, we might experience a simple trip to the grocery store as a series of events that include parking the car, entering the store, putting groceries into a cart, paying the cashier, navigating a crowded parking lot, and loading up the trunk. Interestingly, the activity actually underlying this experience of events tends to be complex, dynamic, and evanescent. How do we transform the continuous, multi-modal stream that actually occurs into the experience of an event sequence? Answering this question has proven to be surprisingly challenging (Baldwin & Baird, 1999, 2001; Kurby & Zacks, 2008; Newtson, Engquist, & Bois, 1977; Zacks, Speer, Swallow, Braver, & Reynolds, 2007; Zacks & Tversky, 2001). Even more challenging is to answer how infants and young children come to acquire adult-like skill at rendering continuously flowing experience in terms of event sequences. One challenge that has slowed progress on these issues is methodological: a dearth of available techniques for examining children’s processing as information unfolds across time. In this paper, we showcase a valuable new method for this purpose, and report novel information about the nature of event-processing skill in preschool-aged children.

A recent surge of research on adults’ event processing offers starting information about the perceptual and cognitive skills at play in this mental rendering process. For one, adult observers display high levels of agreement when asked to consciously report on the location of event boundaries in continuously streaming activity sequences (Newtson, 1973); as well, implicit indicators of processing, like fMRI and self-paced slideshow viewing (Baldassano et al., 2017; Hard, Recchia, & Tversky, 2011; Kosie & Baldwin, 2019b, 2019a), reveal consistent patterns of event segmentation. Event boundary judgments tend to reflect adults’ analysis of the goal structure inherent in activity streams, although motion properties also influence such judgments (e.g., Newtson et al., 1977; Zacks, 2004; Zacks et al., 2006b). Further, adults’ success at achieving a segmental analysis of streaming activity carries functional significance; that is, accurately identifying event boundaries predicts both memory for event content (Sargent et al., 2013), and the ability to perform everyday sequences of action (Bailey, Kurby, Giovannetti, & Zacks, 2013).

Recent research also reveals information about how event-segmentation skills begin to emerge in children’s development, although many questions are yet unanswered. Even infants as young as six months appear to be sensitive to the segmental structure of at least some kinds of activity streams (Baldwin, 2012; Baldwin et al., 2001; Hespos, Grossman, & Saylor, 2010; Hespos, Saylor, & Grossman, 2009; Pace, Levine, Golinkoff, Carver, & Hirsh-Pasek, 2020; Saylor, Baldwin, Baird, & LaBounty, 2007; Stahl, Romberg, Roseberry, Golinkoff, & Hirsh-Pasek, 2014). Like adults, infants’ sensitivity to event units within continously flowing activity predicts their event memory (Sonne, Kingo, & Krøjgaard, 2016, 2017).

Among the unanswered questions, however, is the degree to which early event segmentation rides on knowledge of goal structure. On the one hand, a large literature provides extensive evidence that children, like adults, are sensitive to goal structure and often prioritize it in their processing of unfolding activity (e.g., Buresh & Woodward, 2007; Falck-Ytter, Gredebäck, & Hofsten, 2006; Gergely & Csibra, 2003; Gergely, Nádasdy, Csibra, & Bíró, 1995; Gredebäck & Melinder, 2010; Meltzoff, 1995; Olofson & Baldwin, 2011; Phillips, Wellman, & Spelke, 2002; Woodward, 1998, 2009). However, this body of research doesn’t provide insight into the extent to which children’s goal-related knowledge influences their segmental analysis of continuously unfolding events. In fact, traditional methods used in the bulk of research on children’s knowledge of goal structure (e.g., habituation-dishabituation paradigms) do not lend themselves well to answering questions about event segmentation. To do so requires a method that indexes children’s processing of continuously streaming activity as it unfolds.

Hard, Recchia, and Tversky’s (2011) dwell-time paradigm is just such a technique, and it has previously been successfully employed with preschool-aged children (e.g., Meyer, Baldwin, & Sage, 2011). A primary goal of the present research was to harness the dwell-time paradigm to probe the extent to which children’s analysis of goal structure influences their event segmentation. 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 slideshows. Latency between mouse clicks (i.e., “dwell times”) indexes moment-to-moment changes in attention as activity unfolds. Attentional profiles as revealed by dwell times reflect adults’ segmental analysis of unfolding activity, including: (a) surges in attention at event boundaries (“boundary advantage”); and (b) especially pronounced surges to boundaries that are at higher levels within a goal hierarchy (“hierarchical advantage”) (Hard et al., 2011; Kosie & Baldwin, 2019b, 2019a). These findings reveal that, as sensory information unfolds across time, adults as well as children modulate attention in a manner that reflects their segmental analysis of the dynamic activity stream.

Recent research begins to clarify mechanisms underlying the tendency to selectively focus attention on segment boundaries within dynamic activity. Segment boundaries represent information-rich regions for sampling event content: as one event unit ends, what will happen next is suddenly uncertain. Selectively attending to these regionss thus enhances detection of informative content and facilitates event comprehension (Baldwin & Kosie, n.d.; Shin & Dubrow, n.d.; Zacks, Kurby, Eisenberg, & Haroutunian, 2011)

Some initial evidence indicates that children, like adults, display a boundary advantage in dwell-time patterns measured for relatively familiar activity sequences (Meyer et al., 2011; Ross & Baldwin, 2015). However, another study failed to find these key patterns in preschoolers’ dwell times (Sage & Baldwin, 2015), engendering some uncertainty about the usefulness of the dwell-time paradigm for investigating children’s event segmentation. The present research was thus also designed to provide additional information about the value of the dwell-time paradigm for this purpose.

In the present research we specifically employed the dwell-time paradigm to examine the extent to which children’s knowledge of goal structure influences their segmental analysis of an unfolding activity stream, relative to the influence of differential motion patterns. We selected a type of event that even very young children are believed to understand with respect to goal structure; namely, the activity of reaching for, grasping, and retrieving a ball. In previous research, Phillips and Wellman (2005) employed this very event to examine 12-month-old infants’ ability to recognize the overarching goal of object retrieval independent of the specific motion pattern used to effect that goal. In this study, infants were habituated to an event in which an actor reached over a barrier to grasp and retrieve a ball. During subsequent test, infants displayed recovered interest when the actor displayed the identical motion pattern in a new context that appeared to violate canonical goal-oriented action (reaching over a non-existent barrier to grasp and retrieve the ball). In contrast, infants displayed continued habituation when the actor exhibited a novel motion pattern (reaching straight toward the ball) that preserved the canonical action properties with the same goal structure. These findings have been interpreted as confirming infants’ sensitivity to an actor’s goal independent of the specific motions (e.g., arcing versus straight reach) produced to satisfy the goal.

We employed Phillips and Wellman’s methodological design to investigate the relative degree to which goal structure, surface motion properties, and goal-structure violations influence how preschoolers deploy their attention as activity unfolds across time. We wondered what patterns would emerge in children’s dwell times as they advanced at their own pace through each of the Phillips and Wellman events. In one sense, segmental structure was the same across the three events: an actor reached for, grasped, and retrieved a ball. If goal-related segmental analysis is the primary determinant of children’s attentional deployment during processing, dwell times should exhibit surges at each of these event boundaries, but shouldn’t differ across the three events, despite differences in motion patterns. Alternatively, however, dwell-time patterns might differ across events, for two possible reasons. One possibility is that children’s dwell times would primarily pattern according to motion characteristics, in which case dwell times would be similar for events exhibiting the same motion characteristics (arcing reach over a barrier and arcing reach over no barrier) but different for events differing in motion characteristics (straight reach). Yet another possibility: given that canonical properties of goal-oriented actions were violated for one of the three events (an arcing reach with no visible barrier), dwell-time patterns might differ for this event relative to the other two events. In particular, the violation of expected motion characteristics might elicit enhanced dwelling in the relevant portion of the violating event stream (when the arc over the non-existent barrier actually occurred) relative to the comparable region in other events (the arcing motion over the barrier and the linearly progressing straight reach).

# Methods

## Participants

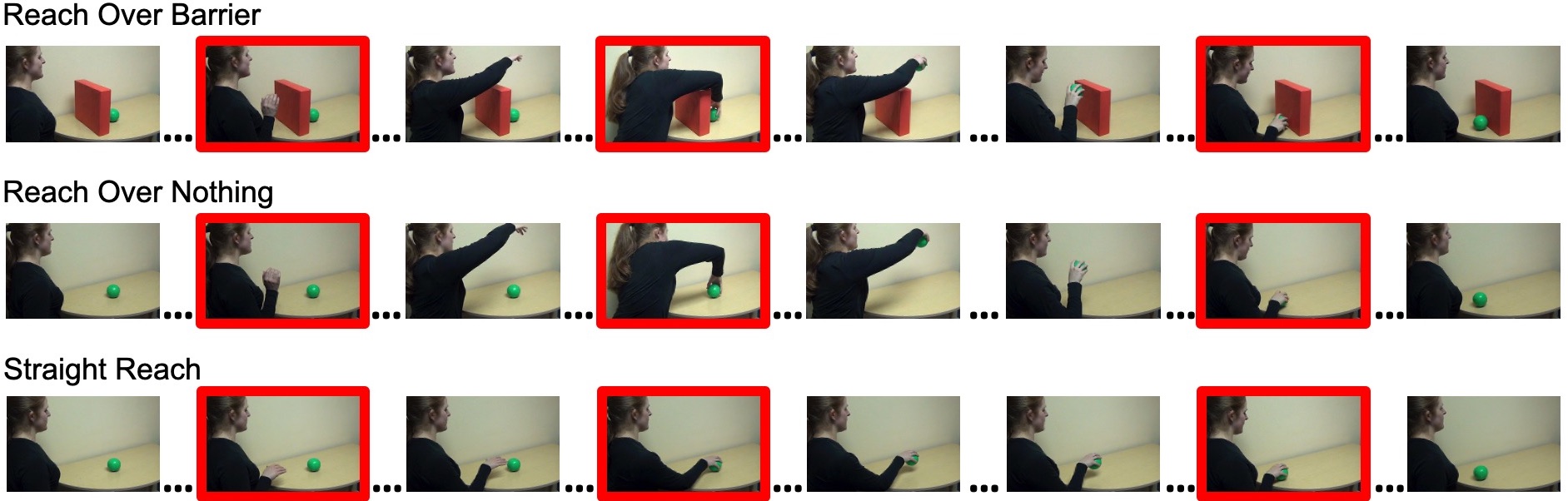
Eighty-eight preschool-aged children (*M* = 40.33 months; range = 2.5 to 4.5 years; 43 female) were recruited to participate from a database of community members with children. Demographics reflected those of the surronding community; particpants were from primarily white and middle-class families.

## Stimuli

Stimuli consisted of slideshows depicting an actor reaching for, and retrieving, a ball in three different ways (adapted from Phillips & Wellman, 2005). In the *barrier* slideshow, the actor reached over a barrier, grasped the ball, lifted it back over the barrier, and set it on the table in front of her. The *nothing* slideshow depicted the actor using the same action pattern displayed in the barrier slideshow, but in the absence of a barrier. In the *straight* slideshow, there was again no barrier present. However, this time the actor reached straight for the ball, grasped it, and pulled it straight back before setting it down on the table in front of her. During filming, the actor used a metronome to ensure that reaches unfolded at the same rate across the three slideshows. Slideshows were created by extracting still frames, at a rate of one frame per second, from these videos. Each slideshow contained 20 total slides, and participants viewed only one of the three slideshows (i.e., slideshow varied between subjects). As is standard practice across dwell-time studies (e.g., Kosie & Baldwin, 2019b, 2019a; Hard et al., 2011), data from the first and last slides were removed prior to analyses.

## Boundary Judgments

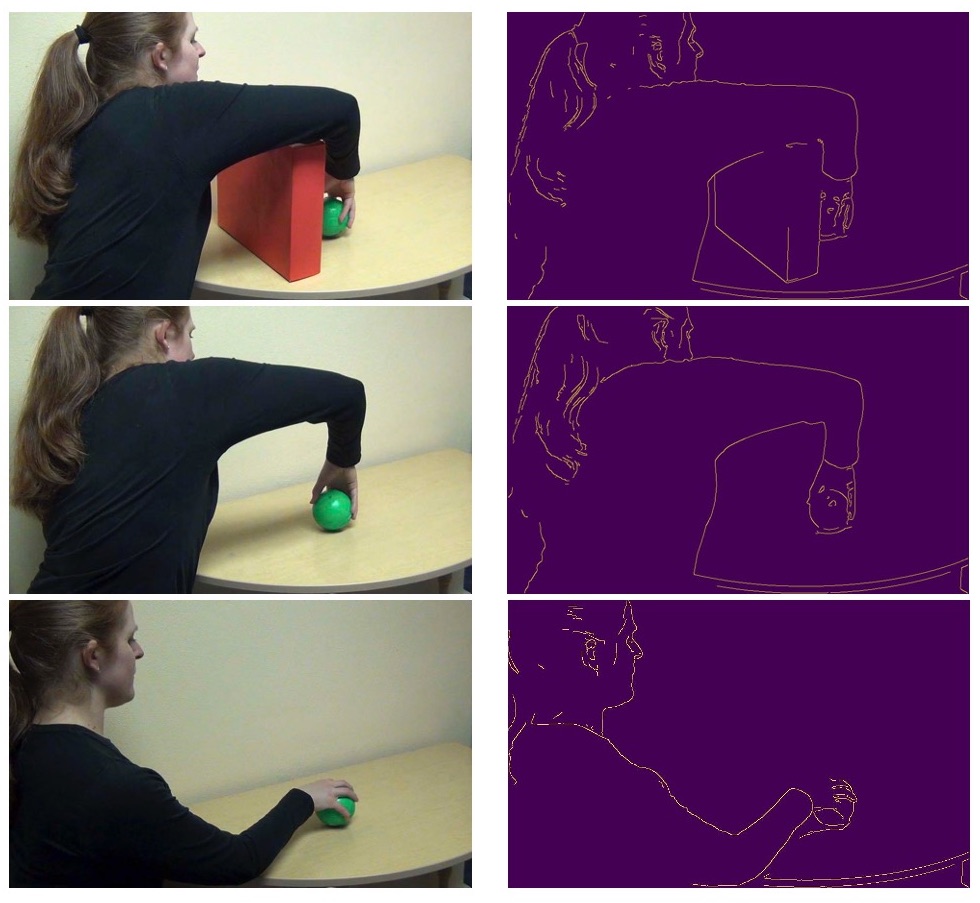
We classified each slide as a boundary or non-boundary. Three total boundary slides were identified per slideshow, corresponding to slides depicting initiation of the reaching action, grasping the ball, and setting the ball on the table. Because care was taken to temporally align the three reaching actions during filming (i.e., using the metronome), these boundaries occurred at the same slide number across the three slideshows. Boundary judgments were verified by a group of 37 naïve raters. A point-biserial correlation indexing agreement (following Zacks and colleagues’ technique, e.g. Zacks et al., 2006a; Kurby & Zacks, 2011) was high (*rpb*(52) = 0.96, , 95% CI[0.93, 0.98]). Example slides are depicted in Figure 1. The three slideshows and additional details about validation of boundary judgments are available on the Open Science Framework (OSF) (Link: [osf.io/m6q7r](https://osf.io/m6q7r); Kosie & Baldwin, 2020).



*Figure* *1:*. Examples of boundary and non-boundary slides from the *barrier*, *nothing*, and *straight* slideshows. Boundary slides are outlined in red.

## Motion Change

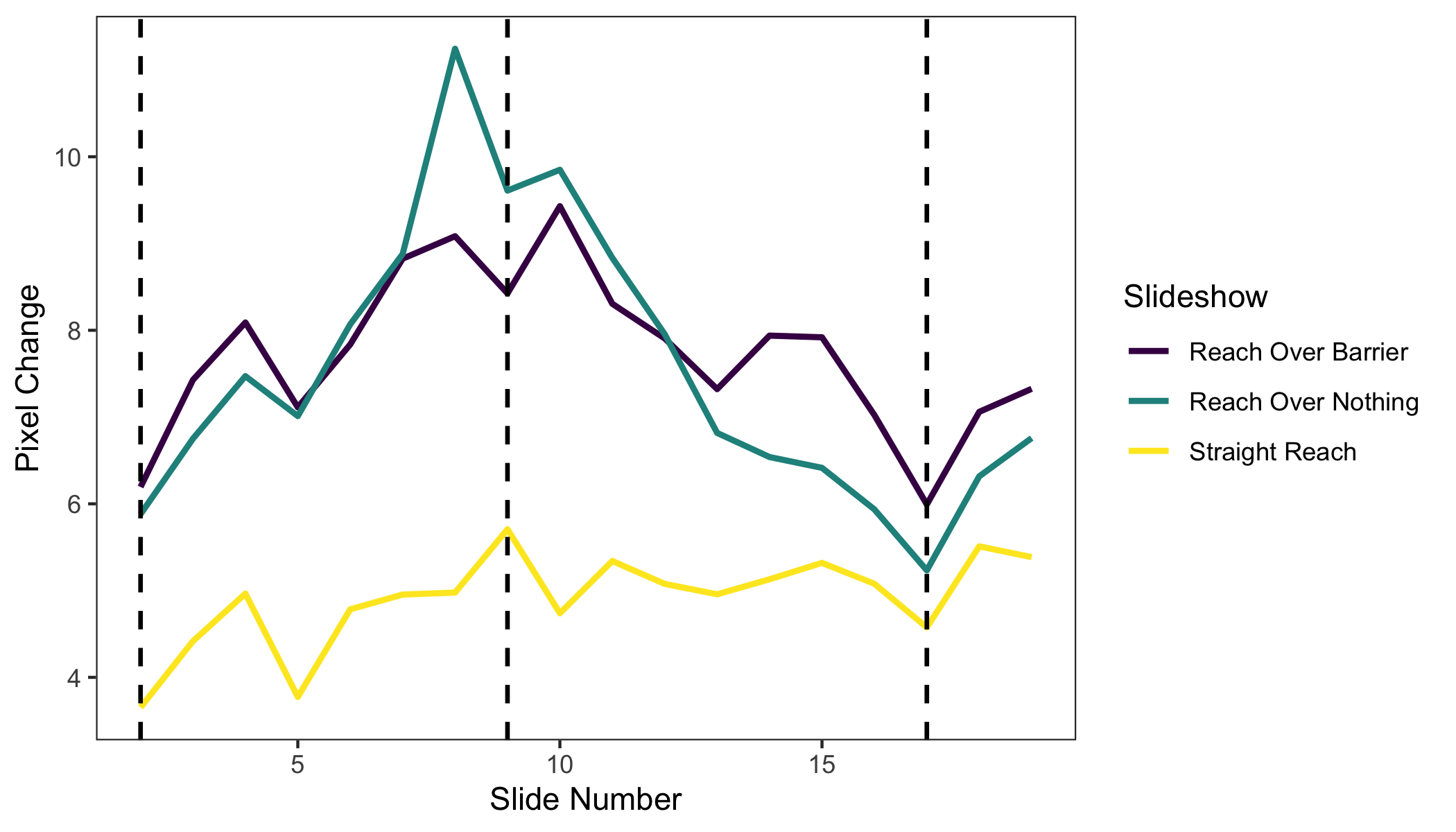
To gain an objective measure of motion-pattern differences across slideshows, we calculated pixel change as an index of slide-to-slide motion change in each of the three slideshows (following Hard et al., 2011). A Canny Edge Detector convolution filter (Canny, 1986) identified and highlighted high contrast areas of each slide image, corresponding to the edges of people and objects (see Figure 2). We then calcualted the degree of pixel change between adjacent convolution-filtered frames (Loucks and Baldwin (2009); see OSF for further details). The resulting values indicated the average amount of pixel change between a given slide and the slide immediately preceding it.



*Figure* *2:*. Example slides from the *barrier* (top), *nothing* (middle), and *straight* (bottom) slideshows before (left) and after (right) being passed through a convolution filter.

A one-way analysis of variance revealed that overall average pixel change values differed across the *barrier* (*M* =7.73, *SD* = 0.91), *nothing* (*M* = 7.53, *SD* = 1.60), and *straight* slideshows (*M* = 4.91, *SD* = 0.54), , . A set of Bonferroni-corrected pairwise comparisons confirmed the prediction that this difference was primarily driven by reduced pixel change for the *straight* slideshow relative to both the *barrier* or *nothing* slideshows (). Average pixel change did not differ between the *barrier* and *nothing* slideshows (). The difference in pixel change for boundary versus non-boundary slides did not reach statistical significance, , . Thus, if a boundary advantage in dwell times turned out to be observed in subsequent analyses, it would be unlikely that this effect could be solely explained by differences in motion change between boundary and non-boundary slides.

While there were overall differences in the magnitude of absolute pixel change between slideshows, the general patterns of pixel change were similar for the *barrier* and *nothing* slideshows, but differed between each of these slideshows relative to the the *straight* reach slideshow. Because the *barrier* and *nothing* reaches unfolded in the same manner (arcing up and over) while the *straight* reach did not, these patterns (depicted in Figure 3) were precisely what would be expected.



*Figure* *3:*. Degree of absolute pixel change across the *barrier*, *nothing*, and *straight* reach slideshows. Dashed lines indicate slides corresponding to event boundaries.

## Procedure

After consenting, children were invited to take a seat at the computer equipped with a wired mouse; a prominent sticker on the mouse indicated where to click. The experiment consisted of a training task and then the main dwell-time task (see OSF for additional details). The goal of the training task (a slideshow depicting a train moving from right to left across the screen) was simply to acquaint children with the experience of clicking the mouse to advance slides in order to see an event unfold. After completing the training task, children clicked through one of the three slideshows. After viewing the slideshow once, an image of a field and clouds was presented, and children then immediately clicked through the same slideshow a second time.

# Results

## Preparing Dwell Time Data

As described previously, dwell times index the latency between mouse clicks, or the total amount of time a given slide was visible on the screen. We used standard methods to prepare dwell-time data for analyses (e.g., Kosie & Baldwin, 2019b, 2019a), including log10 transforming dwell times and removing participants with excessively long dwell times. In the current study, three participants’ data were excluded for this reason, leaving 85 total participants (Nbarrier = 31, Nnothing = 24, Nstraight = 30). See OSF for further detail about the preparation of dwell-time data.

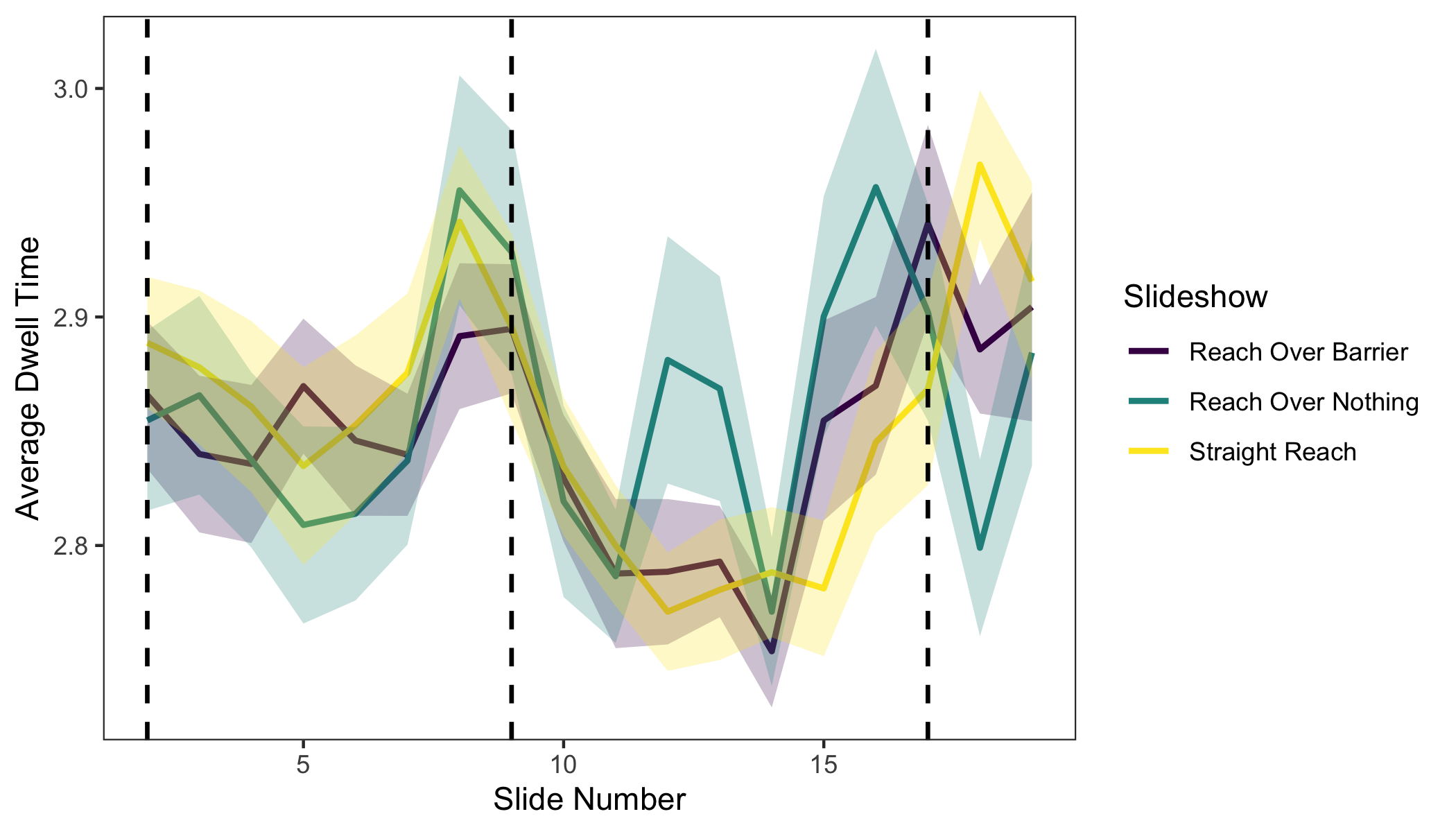
We fit linear mixed effects models with type III sums of squares using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (Team, 2019). The lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) was used to assess statistical significance using Satterthwaite’s approximation for degrees of freedom.

In a preliminary analysis, we asked whether there were systematic dwell-time differences across preschoolers’ viewings of the same slideshow and whether this effect differed across the three slideshow versions. To test this, we ran a mixed-effects model predicting dwell time from fixed effects of viewing and slideshow, random intercepts for subjects, and random slopes for viewing. This analysis revealed that, while overall average dwell times decreased from first (*M* = 2.87, *SD* = 0.24) to second viewing (*M* = 2.84, *SD* = 0.27) (, ), this effect did not reach statistical significance nor did it interact with slideshow (, ). Because the overall decline in dwell time from first to second viewing was similar across all three slideshows, dwell times were averaged across the two viewings.

## Preschoolers priortized segmental structure over motion trajectory.

In our first analysis, we examined the extent to which slide type (boundary vs. non-boundary), slideshow ( *barrier* vs. *nothing* vs. *straight*), and their interaction related to participants’ dwell times. Of particular interest were the extent to which: (a) preschoolers might display longer dwelling to boundary than non-boundary slides, and (b) this boundary advantage effect might differ in relation to which slideshow they viewed, pointing to different segmental analyses for different slideshows. In this mixed-effects model we predicted dwell time from fixed effects for slide type and slideshow, random intercepts for subjects, and random slopes for slide type. Replicating previous work with preschoolers and adults (e.g., Meyer et al., 2011; Hard et al., 2011; Kosie & Baldwin, 2019b, 2019a; Ross & Baldwin, 2015), preschoolers’ dwell times were significantly longer to boundary (*M* = 2.89, *SD* = 0.21) over non-boundary slides (*M* = 2.85, *SD* = 0.20), (), . Neither slideshow nor the interaction between slide type and slideshow were significant predictors of dwell time, all . These effects can be viewed in Figure 4. Thus, despite observed differences in motion trajectory, preschoolers’ tended to increase attention to the same boundary region across all slideshows , indicating that they prioritized goal structure over lower-level motion parameters in their segmental analysis of dynamically unfolding sequences.

To further test the extent to which motion patterns influenced preschoolers’ dwell times, slide-to-slide pixel change was included as a covariate in the above-described model. Pixel change turned out not to be a significant predictor of participants’ dwell times, (), . Additionally, when including pixel change as a covariate, the overall pattern of results did not change; slide type remained the only significant predictor of dwell time, (), . Visual inspection of slide-to-slide differences in pixel change (Figure 3) versus dwell time (Figure 4) supports the finding that dwell times did not in any direct way track pixel change. These results are consistent with the hypothesis that goal structure is a primary determinant of preschoolers’ segmental analysis of continuously unfolding dynamic activity, even when controlling for motion pattern differences.



*Figure* *4:*. Dwell times across the *barrier*, *nothing*, and *straight* slideshows. Dashed lines indicate slides corresponding to event boundaries. Shaded region represents one standard error.

## Exploratory analyses revealed a marginal increase in attention to causal violations.

While we did not find an overall significant effect of slideshow or pixel change, visual inspection of Figure 4 revealed a surge in mean dwell time between slides 12 and 13 of the *nothing* slideshow that seemed to be absent in the other two slideshows. This region corresponds to the juncture in the *nothing* slideshow at which the actor uses an arcing reach - despite the absence of a barrier - to bring the ball back and place it on the table in front of her. If, as described earlier, preschoolers are sensitive to this violation of the canonical properties of goal-oriented action, this juncture might be where we would anticipate a corresponding violation or novelty-related increase in dwell times (as has been observed in previous dwell-time studies with adults; Sage & Baldwin, 2014; Kosie & Baldwin, 2019b). If the effect were quite large, we would have observed it in our earlier analysis comparing participants’ dwell times across the three slideshows. However, it is possible that small but nevertheless systematic differences could be observable between the *nothing* versus *barrier* and *straight* slideshows if analyses were to focus narrowly on the relevant region - where the violation of canonical action specifically occurred Thus, we next compared participants’ dwell times within this region of causal violation (slides 12-13) across the *barrier*, *nothing*, and *straight* slideshows. Because decisions regarding which slides to include in analyses were made *after* viewing Figure 4, we conducted and interpreted these analyses in a purely exploratory manner. Numerically, mean dwell times were greater to the region of causal violation in the *nothing* slideshow (*M* = 2.87, *SD* = 0.25) than in the *barrier* (*M* = 2.79, *SD* = 0.16) or *straight* slideshows (*M* = 2.78, *SD* = 0.15). However, neither the omnibus test of slideshow nor direct comparisons between the *nothing*, *barrier*, and *straight* slideshows reached statistical significance, all . As in the above analysis, these effects did not differ when adding pixel change as a covariate. While the differences in means were suggestive that preschoolers - like the infants’ in Phillips and Wellman’s (2005) research - may be sensitive to violations to canonical goal-oriented action, clear conclusions to this effect could not be drawn from the current results.

# Discussion

Although prior research clarifies that young children often weight goal structure information more heavily than motion patterns in their interpretation and memory for human action, it wasn’t known how these two sources of information influence children’s segmental analysis as activity unfolds across time. To this effect, we examined the extent to which children’s knowledge of goal structure influenced their segmental analysis of unfolding behavior, relative to the influence of differential motion patterns. Preschoolers advanced at their own pace (via the dwell-time paradigm) through slideshows depicting three activity sequences. The three sequences had the same goal-related segmental structure, but differed in motion parameters. As preschoolers viewed the sequences, we measured their time spent “dwelling” on each slide, an index of attentional allocation. Across all three activity sequences, preschoolers’ dwell times increased at the same event boundaries, or junctures at which adult observers judge one unit of action to end and the next to begin. These results - using three new activity sequences - replicate previous research, and suggest that preschoolers, like adults, attend to goal-related segmental structure as they process unfolding activity (Hard, Meyer, & Baldwin, 2019; Hard et al., 2011; Kosie & Baldwin, 2019b, 2019a; Meyer et al., 2011; Ross & Baldwin, 2015).

We additionally engaged in several analyses to explore the extent to which preschoolers’ dwell-time patterns tracked the motion parameters of each sequence. First, we demonstrated that the observed boundary advantage did not differ across the three slideshows, suggesting that the boundary advantage is robust to variation in the precise path an actor takes to achieve a goal. Next, we found that slide-to-slide pixel change - an index of changes in degree of physical motion for each individual slide - were also not predictive of preschoolers’ dwell times. Finally, while we saw some indication that preschoolers up-regulated attention to a motion-related causal violation (e.g., Phillips & Wellman, 2005; Sage & Baldwin, 2014), this effect was weak and did not reach statistical significance. In sum, these results provide substantial evidence that goal-related segmental structure, rather than the specifics of motion parameters underlying the activity stream, are a primary determinant of preschoolers’ dwell-stime patterns.

In addition to providing new insight into how preschoolers modulate their attention when viewing everyday action, this work validates the dwell-time paradigm as a useful methodology with which to explore preschoolers’ processing of activity as it unfolds. The dwell-time paradigm is inexpensive, child-friendly, easy to use, and portable. Thus, in addition to use in the lab, this paradigm can be implemented in preschools, museums, and in the field. We are currently working to develop a version that can be used online, expanding our ability to recruit a large and diverse sample. Finally, because the dwell-time paradigm is an implicit measure of processing - requiring only that participants click through an unfolding slideshow - it provides the opportunity to run exactly the same experiment with adults and children. Comparing the dwell times of children and adults enables us to explore differences in processing across these two groups and holds potential to reveal nuanced information about how children’s processing becomes more adult-like over time.

In conclusion, we (a) replicated previous findings that preschoolers, like adults, attend to action’s goal-related segmental structure within dynamic activity, (b) extended these findings to demonstrate that preschoolers prioritize goal structure over motion properties in their segmental analysis of activity streams, and (c) showcased the value of the dwell-time paradigm as a useful methodology with which to investigate preschoolers’ processing of dynamically unfolding visual information. These results set the stage for further investigation into factors that influence preschoolers’ processing of events as their development and knowledge acquisition progress.

# References

Bailey, H. R., Kurby, C. A., Giovannetti, T., & Zacks, J. M. (2013). Action perception predicts action performance. *Neuropsychologia*, *51*(11), 2294–2304. doi:[10.1016/j.neuropsychologia.2013.06.022](https://doi.org/10.1016/j.neuropsychologia.2013.06.022)

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709–721.e5. doi:[10.1016/j.neuron.2017.06.041](https://doi.org/10.1016/j.neuron.2017.06.041)

Baldwin, D. A. (2012). Redescribing action. In *Navigating the social world: What infants, children, and other species can teach us.* New York: Oxford University Press.

Baldwin, D. A., & Baird, J. A. (1999). Action analysis: A gateway to intentional inference. In *Early social cognition: Understanding others in the first months of life* (pp. 215–240). Mahwah, NJ: Lawrence Earlbaum Associates.

Baldwin, D. A., & Baird, J. A. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Sciences*, *5*(4), 171–178. doi:[10.1016/S1364-6613(00)01615-6](https://doi.org/10.1016/S1364-6613(00)01615-6)

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

Baldwin, D. A., & Kosie, J. E. (n.d.). How does the mind render streaming experience as events? *Topics in Cognitive Science*.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:[10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)

Buresh, J. S., & Woodward, A. L. (2007). Infants track action goals within and across agents. *Cognition*, *104*(2), 287–314. doi:[10.1016/j.cognition.2006.07.001](https://doi.org/10.1016/j.cognition.2006.07.001)

Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *6*, 679–698.

Falck-Ytter, T., Gredebäck, G., & Hofsten, C. von. (2006). Infants predict other people’s action goals. *Nature Neuroscience*, *9*(7), 878–879. doi:[10.1038/nn1729](https://doi.org/10.1038/nn1729)

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naı̈ve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. doi:[10.1016/S1364-6613(03)00128-1](https://doi.org/10.1016/S1364-6613(03)00128-1)

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193. doi:[10.1016/0010-0277(95)00661-H](https://doi.org/10.1016/0010-0277(95)00661-H)

Gredebäck, G., & Melinder, A. (2010). Infants’ understanding of everyday social interactions: A dual process account. *Cognition*, *114*(2), 197–206. doi:[10.1016/j.cognition.2009.09.004](https://doi.org/10.1016/j.cognition.2009.09.004)

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

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

Hespos, S. J., Grossman, S. R., & Saylor, M. M. (2010). Infants’ ability to parse continuous actions: Further evidence. *Neural Networks*, *23*(8-9), 1026–1032. doi:[10.1016/j.neunet.2010.07.010](https://doi.org/10.1016/j.neunet.2010.07.010)

Hespos, S. J., Saylor, M. M., & Grossman, S. R. (2009). Infants’ ability to parse continuous actions. *Developmental Psychology*, *45*(2), 575–585. doi:[10.1037/a0014145](https://doi.org/10.1037/a0014145)

Kosie, J. E., & Baldwin, D. (2019a). Attentional profiles linked to event segmentation are robust to missing information. *Cognitive Research: Principles and Implications*, *4*(1), 8. doi:[10.1186/s41235-019-0157-4](https://doi.org/10.1186/s41235-019-0157-4)

Kosie, J. E., & Baldwin, D. (2019b). Attention rapidly reorganizes to naturally occurring structure in a novel activity sequence. *Cognition*, *182*, 31–44. doi:[10.1016/j.cognition.2018.09.004](https://doi.org/10.1016/j.cognition.2018.09.004)

Kosie, J. E., & Baldwin, D. A. (2020). OSF Repository: Dwell times disambiguate preschoolers’ attention to motion versus goal structure as dynamic action unfolds. *Open Science Framework*. Retrieved from <https://osf.io/m6q7r/>

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–79. doi:[10.1016/j.tics.2007.11.004](https://doi.org/10.1016/j.tics.2007.11.004)

Kurby, C. A., & Zacks, J. M. (2011). Age differences in the perception of hierarchical structure in events. *Memory & Cognition*, *39*(1), 75–91. doi:[10.3758/s13421-010-0027-2](https://doi.org/10.3758/s13421-010-0027-2)

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. doi:[10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13)

Loucks, J., & Baldwin, D. (2009). Sources of information for discriminating dynamic human actions. *Cognition*, *111*(1), 84–97. doi:[10.1016/j.cognition.2008.12.010](https://doi.org/10.1016/j.cognition.2008.12.010)

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*(5), 838–850.

Meyer, M., Baldwin, D. A., & Sage, K. D. (2011). Assessing Young Children’s Hierarchical Action Segmentation. *Proceedings of the Cognitive Science Society*, *33*, 3156–3161.

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

Newtson, D., Engquist, G., & Bois, J. (1977). The Objective Basis of Behavior Units. *Journal of Personality and Social Psychology*, *35*(12), 847–862.

Olofson, E. L., & Baldwin, D. (2011). Infants recognize similar goals across dissimilar actions involving object manipulation. *Cognition*, *118*(2), 258–264. doi:[10.1016/j.cognition.2010.11.012](https://doi.org/10.1016/j.cognition.2010.11.012)

Pace, A., Levine, D. F., Golinkoff, R. M., Carver, L. J., & Hirsh-Pasek, K. (2020). Keeping the end in mind: Preliminary brain and behavioral evidence for broad attention to endpoints in pre-linguistic infants. *Infant Behavior and Development*, *58*, 101425. doi:[10.1016/j.infbeh.2020.101425](https://doi.org/10.1016/j.infbeh.2020.101425)

Phillips, A. T., & Wellman, H. M. (2005). Infants’ understanding of object-directed action. *Cognition*, *98*(2), 137–155. doi:[10.1016/j.cognition.2004.11.005](https://doi.org/10.1016/j.cognition.2004.11.005)

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants’ ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*(1), 53–78. doi:[10.1016/S0010-0277(02)00073-2](https://doi.org/10.1016/S0010-0277(02)00073-2)

Ross, R. A., & Baldwin, D. A. (2015). Event processing as an executive enterprise. In *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource.* (pp. 1–13).

Sage, K. D., & Baldwin, D. (2014). Looking to the hands: Where we dwell in complex manual sequences. *Visual Cognition*, *22*(8), 1092–1104. doi:[10.1080/13506285.2014.962123](https://doi.org/10.1080/13506285.2014.962123)

Sage, K. D., & Baldwin, D. (2015). Children’s Use of Self-Paced Slideshows: An Extension of the Video Deficit Effect? *Journal of Research in Childhood Education*, *29*(1), 90–114. doi:[10.1080/02568543.2014.978919](https://doi.org/10.1080/02568543.2014.978919)

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., … Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, *129*(2), 241–255. doi:[10.1016/j.cognition.2013.07.002](https://doi.org/10.1016/j.cognition.2013.07.002)

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*(1), 113–128. doi:[10.1080/15248370709336996](https://doi.org/10.1080/15248370709336996)

Shin, Y. S., & Dubrow, S. (n.d.). Structuring memory through inference-based event segmentation. *Topics in Cognitive Science*.

Sonne, T., Kingo, O. S., & Krøjgaard, P. (2016). Occlusions at event boundaries during encoding have a negative effect on infant memory. *Consciousness and Cognition*, *41*, 72–82. doi:[10.1016/j.concog.2016.02.006](https://doi.org/10.1016/j.concog.2016.02.006)

Sonne, T., Kingo, O. S., & Krøjgaard, P. (2017). Bound to remember: Infants show superior memory for objects presented at event boundaries. *Scandinavian Journal of Psychology*, *58*(2), 107–113. doi:[10.1111/sjop.12351](https://doi.org/10.1111/sjop.12351)

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants Segment Continuous Events Using Transitional Probabilities. *Child Development*, n/a–n/a. doi:[10.1111/cdev.12247](https://doi.org/10.1111/cdev.12247)

Team, R. C. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statitsical Computing. Retrieved from <https://www.R-project.org/>

Woodward, A. (1998). Infants selectively encode the goal object of an actor’s reach. *Cognition*, *69*(1), 1–34. doi:[10.1016/S0010-0277(98)00058-4](https://doi.org/10.1016/S0010-0277(98)00058-4)

Woodward, A. L. (2009). Infants’ Grasp of Others’ Intentions. *Current Directions in Psychological Science*, *18*(1), 53–57. doi:[10.1111/j.1467-8721.2009.01605.x](https://doi.org/10.1111/j.1467-8721.2009.01605.x)

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*(6), 979–1008. doi:[10.1207/s15516709cog2806\_5](https://doi.org/10.1207/s15516709cog2806_5)

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction Error Associated with the Perceptual Segmentation of Naturalistic Events. *Journal of Cognitive Neuroscience*, *23*(12), 4057–4066. doi:[10.1162/jocn\_a\_00078](https://doi.org/10.1162/jocn_a_00078)

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*(2), 273–293. doi:[10.1037/0033-2909.133.2.273](https://doi.org/10.1037/0033-2909.133.2.273)

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006a). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging*, *21*(3), 466–482. doi:[10.1037/0882-7974.21.3.466](https://doi.org/10.1037/0882-7974.21.3.466)

Zacks, J. M., Swallow, K. M., Vettel, J. M., & McAvoy, M. P. (2006b). Visual motion and the neural correlates of event perception. *Brain Research*, *1076*(1), 150–162. doi:[10.1016/j.brainres.2005.12.122](https://doi.org/10.1016/j.brainres.2005.12.122)

Zacks, J. M., & Tversky, B. (2001). Event Structure in Perception and Conception. *Psychological Bulletin*, *127*(1), 3–21.