Pupillometry sheds light on how caregivers scaffold infants’ learning

Jessica E. Kosie1,2, Avinash Bala3, & Dare A. Baldwin2

1 Department of Psychology, Princeton University

2 Department of Psychology, University of Oregon

3 Institute of Neuroscience, University of Oregon

Author note

## Data Availability

All de-identified data and code will be made available on the [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) page associated with this work.

## Acknowledgements

This article is based on a dissertation by the first author in partial fulfillment of the requirements for a doctoral degree in the Department of Psychology at the University of Oregon. The work was funded in part by a College of Arts and Sciences Dissertation Research Fellowship and a Beverly Fagot Dissertation Fellowship awarded to the first author. We would like to thank members of the first author’s dissertation committee: Dr. Caitlin Fausey, Dr. Lou Moses, and Dr. Eric Pederson. We also thank Dr. Terry Takahashi for theoretical and methodological support, Dr. Shahar Shirtz for technical assistance, Research Assistants in the Acquiring Minds Lab for help with recruitment, data collection, and coding, and the caregivers and infants who volunteered their time to participate in this work.

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

Infants’ learning hinges on early-emerging ability to interpret complex, dynamic human action: they not only recognize many actions, but predict, respond to, remember, and draw inferences from them. Such fluency depends on finding structure within unfolding activity, including identifying boundaries where units of action begin and end. Given that action unfolds rapidly and is just as quickly gone, how do infants identify such boundaries in the complex, fleeting stream of motion that they actually observe? Caregivers’ infant-directed action demonstrations might assist with this challenging learning task. In interactions with infants, caregivers modify their motion in a variety of ways that engage infants’ overall attention (i.e., “motionese;” (Brand, Baldwin, & Ashburn, 2002)). It seems plausible that these modifications promote infants’ detection of internal structure within dynamic activity. We harnessed pupillometry to investigate this possibility, comparing pupil size changes elicited when infants watched motionese versus adult-directed versions of a given activity stream. Infants’ pupil size (an indication of attention or cognitive engagement) increased in response to action boundaries, but only for motionese demonstrations. Thus, in addition to engaging overall attention, motionese facilitates infants’ detection of action’s internal structure. These findings showcase benefits of motionese for infants’ learning about action, and pupillometry for shining light on key developmental phenomena.

*Keywords:* action segmentation, action processing, event processing, motionese, pupillometry, infant cognition

*Word count:* X

Pupillometry sheds light on how caregivers scaffold infants’ learning

Much of what infants learn arises from observing others in action and interacting with them. Yet human activity – whether speech, action, gesture, or emotion (though we will focus specifically on action) - generates a dynamic stream of multimodal information that is complex, rapidly unfolding, and fleeting. Information that unfolds continuously must be processed and encoded “on the fly” before the opportunity is lost. Given these challenges, how do infants come to be able to interpret action in a way that supports learning?

Finding structure internal to continuous action – that is, discerning units of action within streaming activity – is key to making sense of the behavior stream (e.g., Baldwin & Baird, 2001; Kurby & Zacks, 2008; Newtson & Engquist, 1976; Zacks, Tversky, & Iyer, 2001). Indeed, such segmentation skill is linked not only to successful event comprehension, but also to memory (Flores, Bailey, Eisenberg, & Zacks, 2017; Sargent et al., 2013; Sonne, Kingo, & Krøjgaard, 2016, 2017; Swallow, Zacks, & Abrams, 2009), successful re-enactment (Bailey, Kurby, Giovannetti, & Zacks, 2013), social understanding (Baldwin & Baird, 2001; Zalla, Labruyère, & Georgieff, 2013), and language acquisition (Baldwin, 2015; Levine, Buchsbaum, Hirsh‐Pasek, & Golinkoff, 2019). One important question about infant learning, then, is how infants come to detect relevant internal structure within dynamic human activity streams.

Of course, infants aren’t entirely on their own as they confront the complexity of human action. In interactions with infants, caregivers spontaneously modify their actions in a variety of ways that engage infants’ attention and facilitate learning (e.g., Brand et al., 2002; Csibra & Gergely, 2009; Nagai & Rohlfing, 2009; Nagai & Rohlﬁng, 2007). It seems plausible that such “motionese” modifications might help infants find internal structure as activity unfolds continuously. To date, however, this possibility has not been directly tested. To address this gap, we used a novel, inexpensive, and infant-friendly system for measuring infants’ pupillary response (the Stimulus-Induced Pupil Response (SIPR) system; Patent Pending; Bala, Keller, Whitchurch, Baldwin, & Takahashi, 2016; Bala, Keller, & Takahashi, 2018) to explore the extent to which motionese influences infants’ ability to detect segmental structure within dynamic human activity.

Below, we briefly summarize what is currently known about infants’ action segmentation skills, as well as initial evidence that caregivers scaffold infants’ processing of human action. We additionally describe a body of evidence indicating that pupillometry offers potential insight into infants’ processing of segmental structure within dynamic action.

## Action segmentation in infancy

Infants’ sensitivity to segmental structure within continuous activity has been documented for a variety of everyday intentional-action sequences (Baldwin et al., 2001; Hespos, Grossman, & Saylor, 2010; Hespos, Saylor, & Grossman, 2009; Saylor, Baldwin, Baird, & LaBounty, 2007; see Levine et al., 2019 for a recent review). For example, Baldwin and colleagues (2001) familiarized 10- and 11-month-old infants to a video depicting an actor engaging in a series of everyday activities. At test, infants were shown the same videos with pauses of identical length at either action boundaries (i.e., the region within behavior at which one action ends and the next begins, such as the moment at which an actor grasps an object to pick it up) or at non-boundary junctures. Infants looked longer to the test videos that depicted pauses at non-boundary junctures, suggesting that they readily detect structure in unfolding intentional action, parse human behavior with respect to this structure, and are surprised when this structure is violated.

As well, action boundaries appear to hold privileged status in older infants’ and preschoolers’ processing of events (Kosie & Baldwin, under revision-b; Meyer et al., 2011a; Sonne et al., 2016, 2017), which implies sensitivity to segmental structure. For example, Sonne and colleagues (2017) found that, at a delay of ten minutes after viewing, 21-month-old infants more accurately remembered specific objects that had been presented at action boundaries than those presented at non-boundary junctures. Apparently, infants, like adults (e.g., Hard, Recchia, & Tversky, 2011; Kosie & Baldwin, 2019; Kurby & Zacks, 2008), selectively attend to boundaries within unfolding activity. Boundaries tend to be regions within behavior at which predictability dips; in this sense it is cognitively adaptive to selectively sample them (Baldwin & Kosie, n.d.). They are information-rich regions within the flow that play an out-size role in supporting comprehension (Newtson & Engquist, 1976), memory (Flores et al., 2017), and inference (Shin & Dubrow, n.d.), among other things. Just how infants begin to locate action boundaries within dynamically unfolding activity sequences remains an important open question, however.

On the one hand, statistical learning appears to facilitate infants’ detection of segmental structure in action. Several lines of evidence suggest that infants can track regularities within dynamic action to guide processing of the unfolding activity (e.g., Ambrosini et al., 2013; Kanakogi & Itakura, 2011; Monroy, Gerson, & Hunnius, 2017; Nagai, 2019). In particular, infants can capitalize on statistical regularities embedded within novel activity sequences to discover segments within the unfolding stream that were initially opaque to them, relying on low-predictability transitions within the stream to define boundaries between action units (Baldwin, 2012; Roseberry, Richie, Hirsh-Pasek, Golinkoff, & Shipley, 2011; Stahl, Romberg, Roseberry, Golinkoff, & Hirsh-Pasek, 2014). Moreover, infants can use these regularities to guide anticipatory gaze on later occasions when viewing the same sequences (Monroy et al., 2017). However, it is unclear as yet how much, or what kind of, repeated exposure is necessary before the statistics of a given novel activity sequence can be learned and used to guide infants’ subsequent processing. In infants’ day-to-day experience, some contexts might serve to enhance these statistics, promoting infants’ identification of attention-worthy regions of activity (i.e., action boundaries), thereby supporting infants’ rapid acquisition of action-processing skill. One particular context that might be especially influential in this regard occurs when caregivers specifically attempt to demonstrate novel activities to infants.

## “Motionese” might scaffold infants’ detection of structure in action

When demonstrating novel action to infants, caregivers modify their behavior in a variety of systematic ways that seem well suited to promoting infants’ processing of the dynamic activity stream. Recent research investigating this phenomenon provides initial confirmation that infants benefit from such “motionese” demonstrations (Williamson & Brand, 2014). It remains unclear, however, whether motionese specifically scaffolds infants’ detection of segmental structure within action, although this seems highly plausible.

A first study documenting motionese found that, when demonstrating novel objects to 6- to 13-month-old infants, caregivers exhibited increased interactiveness, proximity to their infant interactive partner, enthusiasm, range of motion, repetition, and simplicity in their actions (Brand et al., 2002). These modifications capture infants’ attention, in that infants prefer to watch action demonstrations in a motionese format over action characteristic of demonstrations directed toward adults (Brand & Shallcross, 2008). Toddlers are also more likely to imitate actions demonstrated using motionese (Williamson & Brand, 2014), and use of motionese increases 8- to 10-month-old infants’ subsequent object exploration, which can have downstream benefits for overall learning (Koterba & Iverson, 2009).

The motionese modifications just summarized parallel modifications in language directed to infants, commonly called “motherese” (Snow & Ferguson, 1977), and are likely part of a suite of infant-directed modifications jointly constituting a natural pedagogy phenomenon that has received extensive investigation in the developmental literature (e.g., Sage & Baldwin, 2010; Csibra & Gergely, 2009). Benefits of motherese in speech include facilitating infants’ attention (Fernald, 1985; ManyBabies Consortium, 2020), enhancing infants’ processing of the acoustic and segmental properties of speech (Kuhl, 2004), and promoting their detection of segmental structure within streams of fluent speech (Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989; Thiessen, Hill, & Saffran, 2005).

Given such findings regarding motherese, it seems highly plausible that motionese analogously promotes infants’ detection of segmental structure within dynamic action. In fact, certain documented features of motionese seem likely to specifically direct infants’ attention to action boundaries. For example, during object demonstrations to their 7- to 12-month-old infant, mothers’ infant-directed gaze is systematically aligned with boundary junctures (Brand, Hollenbeck, & Kominsky, 2013). Features of mothers’ infant-directed speech during action demonstrations is often aligned with action boundaries as well. For example, the onset and offset of mothers’ action-describing speech tends to be aligned with boundaries occurring at the initiation or completion of an action unit (Hirsh-Pasek & Golinkoff, 1996; Meyer et al., 2011b), and infants tend to group such packaged action into coherent “chunks” (Brand & Tapscott, 2007). At action boundaries, mothers also tend to speak with rising or falling intonation, perhaps signaling the completion of an action unit (Rohlfing, Fritsch, Wrede, & Jungmann, 2006). Features such as repetition (Brand et al., 2009) and turn taking (Brand et al., 2007) in infant-directed demonstrations also coincide systematically with action boundaries, which might additionally facilitate infants’ attention to the segmental structure of unfolding activity.

All in all, current evidence strongly suggests that motionese may assist infants in detecting action boundaries within continuously flowing activity, which would facilitate learning and memory. As yet, however, the possibility that motionese scaffolds infants’ detection of segmental structure within streaming activity has not been put to direct test, largely because most methods used to investigate infant perception and cognition are not well-suited to the task. In particular, existing techniques for investigating infants’ action processing have typically been limited to first exposing infants to action sequences and then, at later test, measuring infants’ recognition/discrimination with respect to the stimuli that they previously viewed (e.g., Woodward, 1998; Baldwin et al., 2001; Stahl et al., 2014). Fortunately, an increasingly popular technique - measuring ongoing involuntary changes in pupil diameter concomitant with cognitive engagement - provides a novel approach to exploring infants’ processing of unfolding action. This technique thus offers a novel window on ways in which motionese may scaffold such processing.

## Pupillometry as a promising solution

Pupil dilation (hereafter PD) occurs spontaneously in response to changes in luminance (Beatty & Lucero-Wagoner, 2000; Loewenfeld, 1993) as well as in response to a variety of cognitive stimuli (Beatty & Kahneman, 1966; Beatty & Lucero-Wagoner, 2000; Goldwater, 1972; Kahneman & Beatty, 1966; Laeng, Sirois, & Gredebäck, 2012; Sirois & Brisson, 2014). Among other things, these task-evoked pupil responses are thought to reflect the attentional demands imposed by a cognitive task (Beatty & Lucero-Wagoner, 2000; Goldinger & Papesh, 2012). In general, available findings provide strong confirmation of Kahneman’s (1973) suggestion that pupil diameter provides an online indication of the “intensity of attention” being allocated by an observer. Infancy researchers are increasingly taking advantage of pupillometry to investigate infant cognition; infants’ PD profiles have shown a range of similarities to those of adults’ (e.g., Addyman, Rocha, & Mareschal, 2014; Gredebäck & Melinder, 2010; Hepach & Westermann, 2016; Jackson & Sirois, 2009; Sirois & Jackson, 2012; Zhang, Jaffe‐Dax, Wilson, & Emberson, 2018).

In a particularly relevant adult study, Tanaka, Kosie, & Baldwin (in preparation) tracked pupil size while they presented adults with a series of short clips of sport activities, each containing a single coarse-level action boundary (e.g., when the athlete completed the primary goal, such as “locking out” in a weight-lifting maneuver, or striking a ball with a racquet in a tennis serve). Systematic changes in pupil diameter occurred in time-locked relation to action boundaries. Pupil size in adult viewers started to increase immediately prior to action boundaries, peak at or shortly after boundaries, and return toward baseline over an extended period thereafter. Boundaries elicited this characteristic dilation pattern regardless of the particular time-point (early or late) at which they occurred within the unfolding event. Further, systematic boundary-related PDs emerged even when controlling for luminance in statstical analyses, suggesting that dialtions were evoked by event boundaries themselves, and not by any changes in image intensity accompanying the boundaries.

Based on this collection of findings, we opted to employ pupillometry to test for comparable boundary-related PDs in infants, and more specifically, to compare possible boundary-related PDs as infants viewed novel activity sequences in motionese versus adult-directed formats. We recruited infants in the 9-12 month age range, because prior research has documented both that caregivers produce motionese to infants of this age (Brand et al., 2002), and that infants in this age range are known to be sensitive to the segmental structure of at least some everyday activity sequences (Baldwin et al., 2001). We addressed three main questions:

1. To what extent does infants’ previously-observed preference for “motionese” over adult-directed action replicate via pupillometry (in addition to standard looking-time measures as utilized in prior research)?
2. To what degree do infants spontaneously display systematic pupil-dilation response to action boundaries within streaming activity?
3. To what extent does motionese scaffold infants’ detection of action boundaries within continuous activity sequences relative to adult-directed action?

Regarding the first question, there is reason to suspect that motionese might elicit a general, overarching increase in pupil size in infants, relative to what adult-directed action induces. In pupillometry research, arousal manifests in increases in tonic (or sustained) pupil dilation (Kahneman & Beatty, 1966; Laeng et al., 2012). Such sustained increase in pupil size to highly arousing social stimuli has also been observed across a variety of infant and preschooler studies (e.g., Hepach, Vaish, & Tomasello, 2012, 2015; Geangu, Hauf, Bhardwaj, & Bentz, 2011; Martineau et al., 2011; Nuske, Vivanti, & Dissanayake, 2015; Nuske, Vivanti, Hudry, & Dissanayake, 2014). We therefore predicted that the previously documented looking-time preference for motionese (Brand & Shallcross, 2008) would be reflected in an enhanced pupillary tonic to motionese action sequences relative to that observed in relation to the adult-directed action sequences. As a direct replication of previous research, we additionally measured infants’ looking time to motionese and adult-directed activity (i.e., how long infants looked at the videos). Again, we expected to replicate prior research, predicting that infants would look longer to motionese than adult-directed action sequences.

In relation to the second question, we hypothesized that infants, like adults, privilege action boundaries in their processing. We thus predicted that they too would display a PD in relation to action boundaries that they are able to detect. As described earlier, a substantial body of prior evidence indicates that, even in the absence of caregiver scaffolding, infants can register action boundaries in at least some kinds of simple, familiar activity sequences (Baldwin et al., 2001; Hespos et al., 2010, 2009; Saylor et al., 2007). A study by Jackson and Sirois (2009) provided incidental evidence highlighting the plausibility of this prediction. They measured pupil diameter as infants viewed a train repeatedly entering and exiting a tunnel; infants’ pupil dilation profiles displayed a clear, systematic PD at the juncture when the train exited the tunnel (a juncture adults would likely judge to be an action boundary). Importantly, their findings could not be explained by a change in luminance alone. Although investigating infants’ pupil responses to action boundaries was not the focus of their research, their results nevertheless seem to provide suggestive evidence that infants exhibit a PD in response to action boundaries, thereby supporting our prediction that infants would display a systematic pupil response to action boundaries in human activity streams. We note, however, that the activity streams we presented to infants in the present study were relatively novel for infants, making it uncertain whether they would be able to detect segmental structure within the dynamic flow of action.

Concerning the third question, we hypothesized that motionese facilitates infants’ detection of action boundaries. For this reason, we predicted that infants would display an enhanced boundary-related PD in response to novel activity streams produced in motionese format, relative to adult-directed format. Evidence from the literatures regarding both motherese and motionese informed this prediction. Speech modifications that are characteristic of motherese have been shown to enhance infants’ ability to extract structure from dynamic streams of auditory stimuli (Kemler-Nelson et al., 1989; Thiessen et al., 2005). By analogy to motherese, motionese also may highlight action boundaries (e.g., Brand et al., 2013, 2009, 2007; Meyer et al., 2011a; Rohlfing et al., 2006). We thus expected motionse to have a similarly facilitative effect on infants’ detection of action boundaries. That is, we suspected that infants might display PD to action boundaries even in the non-motionese condition, but that motionese would produce a synergistic effect whereby a boundary-related PD would be larger when actions were demonstrated via motionese relative to adult-directed action.

To summarize, we predicted (a) an overall attentional-enhancement effect of motionese, such that tonic pupil size would be larger for actions demonstrated via motionese relative to those displayed via adult-directed action, and (b) phasic responses to action boundaries across both motionese and adult-directed demonstrations, but (c) larger phasic responses to action boundaries when actions were demonstrated in a motionese than an adult-directed format.

# Methods

## Participants

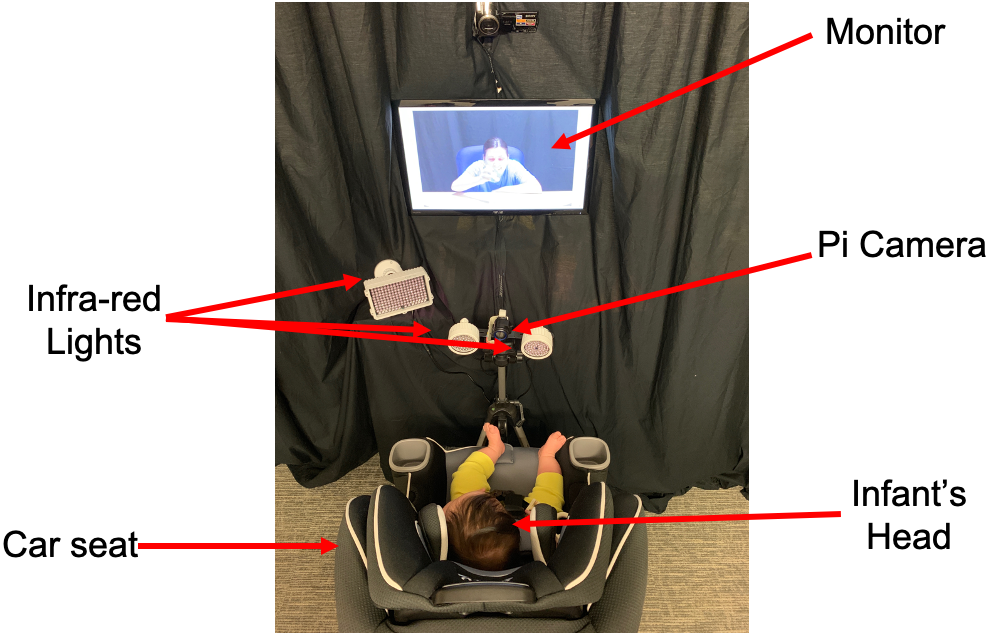
Thirty-six infants ranging from 9 to 12 months (19 females; Mean = 317.03 days; SD = 34.90 days) and their caregivers were recruited to participate from a database of community members with infants. Two infants were immediately excluded due to serious medical issues at birth. While our novel methodology and research question made it challenging to conduct a reliable power analysis, this sample size was selected for a number of reasons. First, it is substantially larger than the minimum sample size (of 24 infants per cell) suggested by Oakes (2017) for infant looking-time research and is larger than the samples collected in most infant pupillometry research. Additionally, we planned to collect a substantial amount of data from each individual infant (up to 36 trials). The large amount of data per infant was likely to provide a strong estimate of within-subject effects and thus increase our available power (DeBolt, Rhemtulla, & Oakes, 2020).

Race/ethnicity and maternal education level were representative of the local community; all participants (N = 36; 100%) identified as White, 6 participants (17%) additionally identified as another race/ethnicity, and 15 of mothers (42%) reported some level of graduate training. This research was approved by the University of Oregon IRB and parental informed consent was obtained for all infants.

## Apparatus

Infants were seated securely in a car seat approximately 82cm from a black floor-to-ceiling curtain, in front of which was a 58cm wide-screen monitor that presented stimuli at a size of 1920 x 1080 pixels. Infants’ movements in the car seat were not restricted, but padding near the head decreased movement.

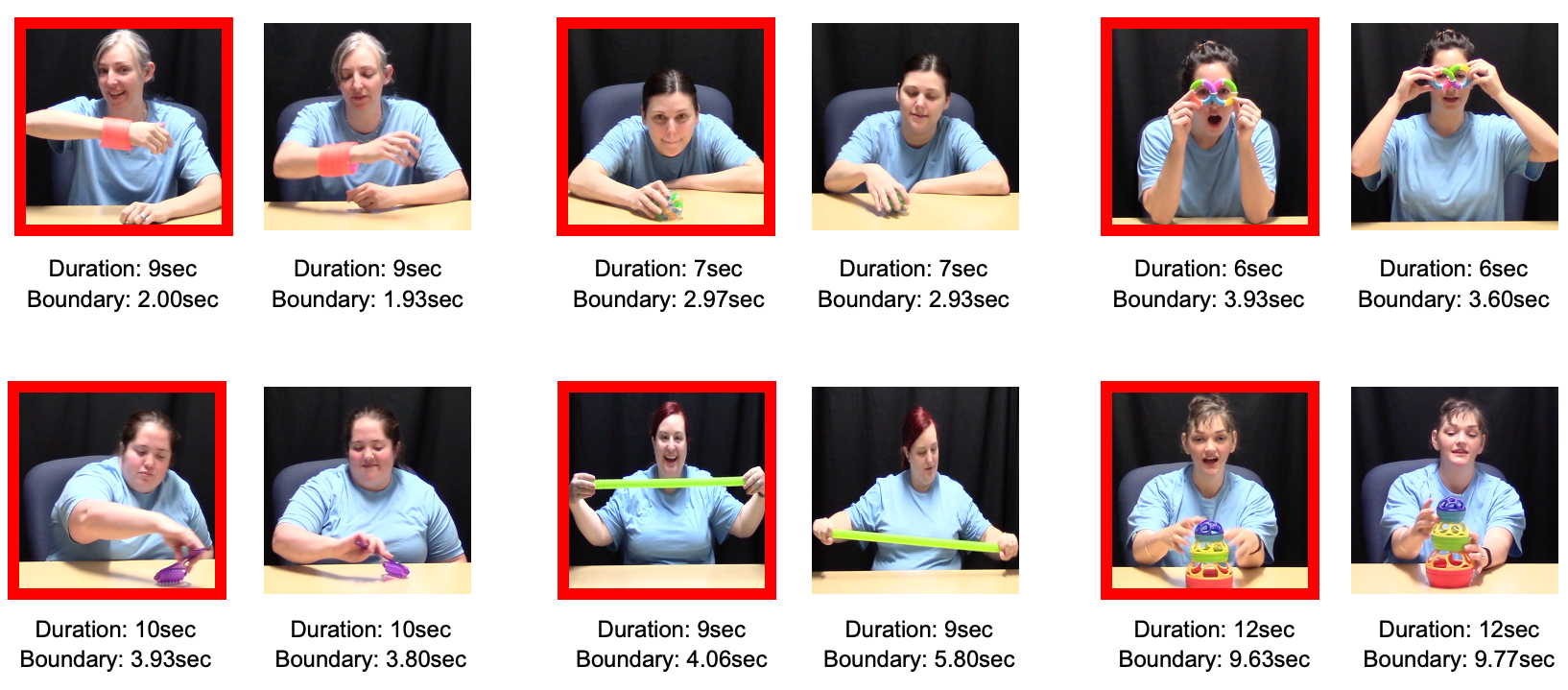
Pupil dilation was digitally recorded via a Raspbery Pi NoIR infrared camera placed approximately 38cm from the infant’s eyes, just out of reach. Video from the camera was recorded to a Raspberry Pi single-board computer at a rate of 30 frames per second. Two small infrared lights were placed on either side of the Pi camera and a third, larger, infrared light was placed immediately to the left of the Pi camera. These lights helped to illuminate the infant’s face and make the pupils more readily detectible on the resulting video recording. A second SONY video camera was placed above the monitor and zoomed in to gain a close view of the infant’s face. The video file to which this camera recorded was synchronized with the video being played to the infant, resulting in a recording of the infant’s face that also depicted what the infant was seeing. This second video was used for hand coding a subset of infants’ looking throughout the pupillometry session as a validity check. Figure 1 depicts the experimental setup.



*Figure* *1:*. Experimental setup.

## Stimuli

To investigate the extent to which caregivers’ modifications to infant-directed action (i.e., “motionese”) influenced infants’ processing of dynamic activity sequences, we selected a subset of videos from a new, open-source corpus of luminance-balanced motionese and adult-directed demonstrations of novel objects created by Kosie & Baldwin (2018, in preparation-a) (see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) for details about the corpus). We chose twelve clips (with six unique objects and six unique caregivers) to be used as stimuli for the pupillometry study. Our goal was to select videos in which infant- and adult-directed action were clearly distinct, while at the same time controlling for extraneous factors that might account for differences in infants’ attention to adult and infant-directed demonstrations. For example, we opted to have the same actor depicted in both the infant- and adult-directed clips involving interaction with the same object. Additionally, because one of our goals was to optimize the chance that each infant in the pupillometry study would attend to multiple presentations of each clip, we opted to keep the clips fairly short (i.e., six to twelve seconds in length). While we did allow length of the clips to vary across objects, each pair of infant- and adult-directed clips involving a given object were equated in length. Clips were selected and trimmed from full videos such that one major action boundary (e.g., the moment at which a caregiver finished placing the final item onto a tower) occurred at approximately the same location within a pair of adult and infant-directed demonstrations, but varied across demonstration pairs such that the boundary did not always occur at the same temporal location. See Figure 2 for a summary of the twelve video clips.



*Figure* *2:*. Example boundary images extracted from each of the twelve slideshows. Motionese images are outlined in red.

Before finalizing our stimulus choices, we examined the potential of luminance-related influence on PDR effects. For each frame in each of our twelve videos, we used a weighted average of pixel values across the red, green, and blue (RGB) channels as an index of photometric luminance (i.e., luminance corrected for perceived brightness by a human observer; see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) for additional details). This method of calculating luminance is standard when working with video files (e.g., Poynton, 2013) and has been used to measure and control for luminance of a stimulus in prior infant pupillometry research (e.g., Jackson & Sirois, 2009; Geangu et al., 2011; Hepach & Westermann, 2013). We used these values to directly test the extent to which luminance differed across: (1) videos of infant- versus adult-directed action, and (2) boundary, pre-boundary, and post-boundary regions. We also examined the extent to which any boundary-related luminance effects might differ across infant- versus adult-directed activity sequences. To briefly summarize our findings (see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) for additional detail), luminance was significantly lower at boundary regions than pre- and post-boundary regions ( *ps* < .02), but only for infant-directed action. Luminance did not differ significantly across (a) motionese versus adult-directed sequences on average, (b) pre-boundary, boundary, or post-boundary regions in adult-directed demonstrations, (c) across pre- and post-boundary regions in infant-directed demonstrations. In principle, low luminance could hypothetically be responsible for larger pupil dilation at boundary versus pre- and post-boundary regions in infant-directed action; thus, we controlled for luminance in statistical analyses reported below.

Behavioral coding of various kinds was undertaken to further validate the twelve video clips (see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) for details). First, a naïve group of coders displayed strong agreement with each other and with expert judgments regarding the location of the one major action boundary in each of the video clips. Next, coders scored all videos for the extent to which the actor displayed motionese characteristics. Motionese scores were significantly higher for infant-directed relative to adult-directed videos. Finally, videos were coded in terms of infants’ spontaneous looking to, and play with, the objects during an infant interaction task; no significant differences emerged in infants’ baseline interest across any of the included objects. Caregivers’ questionnaire responses further supported this finding; infants were relatively unfamiliar with all of the objects. Together, these findings established that the selected sub-sample of twelve infant- and adult-directed action video clips was appropriate for exploring the effects of motionese on infants’ processing of unfolding activity using the pupillometry methodology.

## Design

To avoid effects of familiarity with actors and objects and to keep videos engaging, each infant saw only one video from each pair; that is, they saw either the infant- or adult-directed version of action on a given object. As a result, each infant viewed six unique videos, three infant-directed and three adult-directed. A set of six videos (three infant-directed, three adult-directed) constituted one “block.” Infants viewed up to six total blocks (all six blocks were comprised of the same six videos); thus each infant viewed up to 36 total videos (six repetitions of the same six videos). Because we could not fully counter-balance the videos given our expected sample size, we opted to randomly choose two groups of videos (three infant-directed and three adult-directed in each) and assign an equal number of infants to each group. For each infant, the order of the selected videos was pseudo-randomly assigned such that they never saw more than two infant- or adult-directed videos in a row.

Stimulus presentation was programmed in PsychoPy (Peirce, 2007). As depicted in Figure 3, all blocks started with a brief video of a laughing baby as an attention-getter to help infants orient to the monitor, which played for eight seconds. After the laughing baby attention-getter, a set of moving concentric circles played for three seconds as infants heard a chime sound. The laughing baby and chimes stimuli were acquired via publicly shared materials from the ManyBabies study of infant-directed speech preference (ManyBabies Consortium, 2020). While the laughing baby attention-getter was only played at the start of each block, the circle-and-chimes attention-getter was played before each video. For three seconds at the beginning of each video, infants were presented with a grey screen signaling the start of the trial.1 After the grey screen, infants were presented with a three-second still image depicting the first frame of the action sequence. The still image was included to allow infants’ pupils to adapt to both the luminance and the characteristics of the visual scene (e.g., featured actor and object) that would be viewed in the upcoming video. After the three-second still, the action sequence began to play silently at a standard rate of 30 frames per second. Upon completion of a trial, infants again heard the chimes while viewing the concentric circles, and then the next trial started with a grey screen followed by a still frame. Once infants had completed their six unique trials the laughing baby played again, starting the next block. This sequence repeated for a total of six blocks or until infants became too fussy to continue.



*Figure* *3:*. Flow of events for a given trial (laughing baby marked the beginning of a block of trials). Six trials constituted a single block, and each block repeated a maximum of six times.

## Procedure

Caregivers were seated shoulder-to-shoulder with infants, but facing away from the monitor; they were asked to avoid looking at the stimuli or interacting with infants. This setup allowed infants to see caregivers should they look over, but avoided the possibility that infants would be influenced by any caregiver reaction to the stimuli. If infants started to fuss, caregivers were asked to simply put their hand on infants; we stopped the procedure if infants’ fussing intensified or caregivers requested. Once caregivers and infants were seated, the experimenter adjusted the focus of the Raspberry Pi NoIR camera to ensure a clear picture of infants’ pupil, then started the session from behind a curtain. If infants completed the entire set of six blocks, this part of the session lasted approximately 12 minutes.

## Inclusion Criteria

Trials were considered unusable if infants were fussy (as coded by the blinded experimenter) and/or not looking at the screen for at least 50% of the trial. The number of unusable trials was approximately equal for motionese and adult-directed activity. All infants in the current study contributed at least one block of data, thus none were completely excluded from analyses. In total, 897 trials across 34 infants were included in the pupillometry analyses (median number per infant was 29.50 out of 36 total possible trials).

## Data acquisition

We recorded a separate video for each infant via the Raspberry Pi NoIR camera and Raspberry Pi computer. Each video was run through a Matlab (MATLAB, 2019) pupil detection program designed to advance frame-by-frame through the video recorded by the Raspberry Pi camera, find circles, and measure their diameter (see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) for details).

Following previous research using this pupillometry system (Bala et al., 2018, 2016) as well as that of other studies using pupillometry (e.g., Unsworth & Robison, 2015; Bala, Whitchurch, & Takahashi, 2019; Miller, Gross, & Unsworth, 2019), we opted not to interpolate missing values and thus preserved the original data to the extent possible. Following standard practice, we z-scored pupil size measurements for each participant, including all relevant frames for each participant (i.e., data from the grey screen, still frame, and video across all blocks and trials but ignoring responses to the attention-getting stimuli) and used these same z-scored data across all analyses. Specifically, we calculated the mean and standard deviation of pupil size for each participant (across all blocks and trials), subtracted the individual’s mean from their pupil diameter at each frame, and divided this value by that individual’s standard deviation.

After z-scoring, raw pupil values were filtered to eliminate random fluctuations in the data. While there are multiple possible filters that can be used to smooth pupillometry data (see Hepach & Westermann, 2016 for a review in infancy research), we opted to use a hanning filter with a standard window size of 11 frames (see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) for details). The z-scored, filtered data are referred to simply as “pupil size” for the remainder of this manuscript.

After filtering the data, we generated baseline values for each participant on each trial. Following others (e.g., Geangu et al., 2011; Hepach & Westermann, 2013; Nuske et al., 2015), our measurement of baseline was the average pupil diameter in the one-second region before onset of the video, calculated separately on each trial (see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) for details). In line with methods used by Tanaka and colleagues (in preparation), we opted to control for baseline as a covariate in mixed-effects models rather than correct for baseline via subtracting or dividing pupil sizes as sometimes occurs in pupillometry research.

# Results

Our goals in the current analyses were to (1) examine effects of motionese on infants’ looking and pupil diameter, (2) explore the extent to which infants displayed a boundary-related PD, and (3) investigate the influence of motionese on any such boundary-related PD. We used the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (Team, 2019) with type III sums of squares (set using the afex package; Signmann, Bolker, Westfall, Aust, & Ben-Shachar, 2019) to estimate linear mixed-effects models. Significance for these models was assessed using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017; Luke, 2017) with Satterthwaite’s approximation for degrees of freedom. We have specified the exact fixed and random effects structure used for each model below. Baseline pupil size was controlled for in all analyses involving infants’ pupil diameter. The pupillometry study and analysis plan were preregistered (see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00) repository); however, a few minor deviations from preregistration occurred. None of these deviations influenced our general pattern of results, but nonetheless details of these deviations are described in further detail in the supplementary material (see [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00)).

## Did motionese enhance infants’ overall attention to action?

We first explored the extent to which motionese, relative to adult-directed action, influenced infants’ overall attention to unfolding activity by examining both (1) infants’ looking duration to motionese versus adult-directed action, and (2) their overall average pupil diameter (i.e., tonic pupil size) in response to motionese versus adult-directed action. These analyses included only data from frames corresponding to the video portion of each trial (i.e., ignoring the grey screen and still frames).

For analysis of infants’ looking to motionese versus adult-directed action, any frame for which the Pi/Matlab program detected a pupil was classified as “looking” and any frame for which Pi/Matlab did not detect a pupil was classified as “not looking.” We then created a proportion of time spent looking to each trial by summing the number of frames participants spent looking during each trial and dividing that value by the total number of frames in the trial. We conducted a linear mixed-effects model predicting the proportion of time spent looking to a given trial from a fixed effect of demonstration type (i.e., motionese versus adult-directed action) and random intercepts for subjects and videos. Because looking times were strongly positively skewed (i.e., most infants looked for almost the full trial duration), looking times were z-scored prior to analysis. In line with previous research (i.e., Brand et al., 2002), we found that infants looked significantly longer to motionese (*M* = 93.84%, *SD* = 10.40%) over adult-directed action (*M* = 92.19%, *SD* = 12.19%), (), .

To examine the influence of motionese versus adult-directed action on infants’ tonic pupil size, we ran a linear mixed-effects model predicting infants’ z-scored, filtered pupil diameter from a fixed effect of demonstration type (motionese versus adult-directed action) with random intercepts for subjects and videos, controlling for infants’ baseline pupil diameter. This time, contrary to our predictions, we did not find a significant effect of demonstration type, (), , though infants’ average pupil diameter tended to be larger in response to motionese ( *M* = 0.02, *SD* = 0.78) over adult-directed activity sequences (*M* = -0.11, *SD* = 0.80). Thus, while our analysis of infants’ *looking* to motionese versus adult-directed action replicated prior results, our analysis of infants’ *pupil diameter* (an attempt to replicate previous findings with a novel methdology) did not reveal a systematic motionese preference.

## Did infants display sensitivity to action boundaries in continuous activity sequences?

In our next set of analyses, we explored the extent to which infants displayed a PDR to the major action boundary within unfolding activity sequences, as indexed by changes in pupil diameter. For these analyses, we focused in particular on activity surrounding the one major action boundary depicted within each video. We first defined pre-boundary, boundary, and post-boundary regions in each video. The pre-boundary region covered the one second of activity (or 30 frames) occurring prior to the action boundary. The boundary region began at the action boundary and extended for the next one second (30 frames), and the post-boundary region began at the end of the boundary region and continued one additional second, or 30 more frames. We note that these are broader time windows than the comparable half-second windows used in Tanaka and colleagues’ (in preparation) for pre-boundary, boundary, and post-boundary regions, based on evidence that infants’ pupils respond to cognitive stimuli more slowly than adults’ (e.g., Verschoor, Spapé, Biro, & Hommel, 2013; Verschoor, Paulus, Spapé, Biro, & Hommel, 2015; Zhang et al., 2018).

To test for a possible boundary-related PDR across all videos (regardless of whether activity depicted was motionese versus adult-directed action), we ran a linear mixed-effects model predicting z-scored, filtered pupil diameter from a fixed effect of region (pre-boundary, boundary, post-boundary) and random intercepts for subjects and videos, controlling for baseline pupil size. Because we were specifically interested in boundary effects, the video frames included in these analyses were limited to those occurring in pre-boundary, boundary, and post-boundary regions. We found a significant main effect of region, , that followed both linear (), and quadratic trends (), . These effects remained when luminance was added to the model as a covariate in a follow-up analysis.

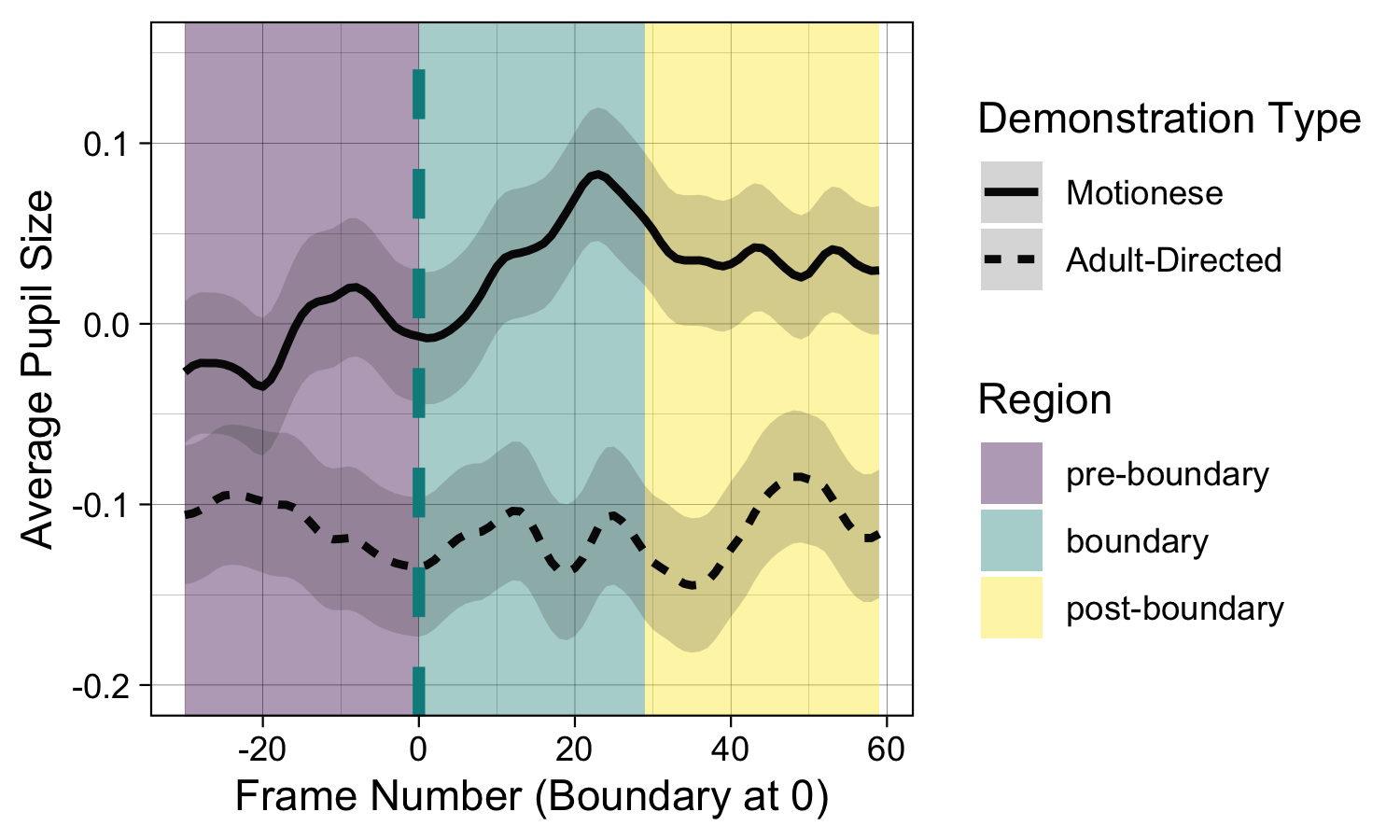
To explore the locus of the region effect, we ran a set of Bonferroni-corrected pairwise comparisons. Infants’ pupil diameter was significantly larger to boundary ( *M* = -0.04, *SD* = 0.80) and post-boundary regions ( *M* = -0.04, *SD* = 0.77) than to pre-boundary regions ( *M* = -0.06, *SD* = 0.82) ( and , respectively). However, pupil diameter to post-boundary regions did not significantly differ from pupil diameter to boundary regions (). To summarize, infants’ pupil diameter systematically increased to action boundaries as expected. However, infants’ pupils did not constrict in the one second region after boundaries.

## Did motionese enhance infants’ pupillary response to boundaries within continuous activity?

To explore the extent to which motionese influenced infants’ response to boundaries, we ran the same mixed-effects model described above, but now included fixed effects of demonstration type and an interaction between region and demonstration type, while still controlling for baseline pupil size. As in previous analyses, we found no significant main effect of demonstration type, , , but a significant main effect of region , . As predicted, we also found a significant interaction between demonstration type and region, , .

To explore this interaction, depicted in Figure 4, we ran two separate mixed-effects models for motionese and adult-directed demonstrations. In adult-directed demonstrations, there was no systematic effect of region, , . Pupil diameter did not differ significantly across pre-boundary ( *M* = 0.04, *SD* = 0.75), boundary ( *M* = -0.01, *SD* = 0.81), and post-boundary regions ( *M* = 0.04, *SD* = 0.78). In contrast, for motionese demonstrations we observed a significant effect of region, , that followed linear ( (), ) and quadratic trends ( (), .

We again followed up with a set of Bonferroni-corrected pairwise comparisons. In motionese demonstrations, pupil size during pre-boundary slides ( *M* = -0.01, *SD* = 0.81) was smaller than pupil size to boundary slides ( *M* = 0.04, *SD* = 0.78) and post-boundary slides ( *M* = 0.04, *SD* = 0.75), . However, pupil size did not differ significantly between boundary and post-boundary slides, . To summarize, in response to motionese demonstrations, infants’ pupil size increased, as predicted, within boundary regions (relative to pre-boundary regions) and remained high post-boundary. These effects were not observed in infants’ pupil size during adult-directed demonstrations. Again, a follow-up analysis controlling for luminance did not alter these patterns in any way.



*Figure* *4:*. Average z-scored, filtered pupil size in response to motionese (solid line) and adult-directed (dashed line) action. Video region is indicated by the background color of the plot, with the boundary occurring at time 0 on the x-axis. Shading around lines indicates +/- 1 SE.

# Discussion

Among other things, previous research has documented that (a) human caregivers spontaneously modify motion when demonstrating action to infants, (b) infants prefer such modified motion over adult-directed action, and (c) motionese promotes infants’ imitation of novel activity sequences. What had remained mysterious, however, were the precise ways in which motionese might alter infants’ processing of dynamically unfolding activity. This question has been difficult to address, in part, because existing methodologies have been ill-suited to probing infants’ moment-to-moment action processing. We harnessed pupillometry as a potentially valuable method to address the question.

We first tested for replication of the previous finding that infants prefer motionese over adult-directed demonstrations (e.g., Brand & Shallcross, 2008). In particular, we probed a possible preference in terms of a measure tested previously – infants’ looking time – and also with respect to a new measure – infants’ tonic pupil dilation. Our analysis of looking times replicated the previously observed motionese preference, although mean pupil size only trended in this direction.

A subsequent set of analyses examined whether infants displayed a pupil dilation response to major action boundaries across the videos, as previously documented for adults by Tanaka and colleagues (in preparation). We measured infants’ pupil size during pre-boundary, boundary, and post-boundary regions of unfolding activity sequences; the precise time-point at which these regions occurred varied across the videos. Overall, infants’ pupil diameter displayed a systematic increase in boundary relative to pre-boundary regions, and then remained elevated afterwards. However, the effect of region differed across motionese and adult-directed demonstrations, as evidenced by a significant interaction between action region and demonstration type. Specifically, the effect of increased pupil dilation at boundary regions held only for motionese demonstrations; it was absent when infants viewed adult-directed action. These patterns did not differ when luminance was added as a covariate in analyses. Together, these findings simultaneously (a) validate pupillometry as a useful technique for investigating infants’ on-line event segmentation, and (b) confirm our prediction that motionese not only enhances infants’ overall interest in novel activity streams, but also specifically facilitates their detection of segmental structure within unfolding activity sequences.

Our results offer the first evidence to date on several fronts. First, these findings showcase the value that pupillometry holds for investigating infants’ processing of unfolding events, and thus for shedding light on infants’ learning more generally. Second, our findings provide new evidence that motionese alters the details of infants’ online processing of novel activity sequences. Third, these results document that, with the aid of motionese, infants can detect internal structure within novel, dynamic activity streams. Moreover, the pupil dilation that infants displayed in relation to boundaries within motionese demonstrations implies enhanced cognitive engagement with those regions. Thus, motionese appears to entrain infants’ attention in a highly beneficial way – promoting enhanced attention to particularly informative regions within the complex stream of behavior (Baldwin & Kosie, n.d.).

These ground-breaking findings engender a range of additional questions, which we take up in turn:

## Why did looking time reveal a systematic motionese preference, whereas tonic pupil dilation did not?

As we intended, infants displayed overall high levels of looking to the stimulus videos in this research; nevertheless, they systematically looked longer to motionese than adult-directed demonstrations. It seems possible, however, that infants’ overall high level of engagement with the videos may have interfered with detecting a preference for motionese via the tonic pupil dilation measure. That is, restricted variance may have undercut sensitivity to overall differences in pupil dilation related to demonstration type. Additionally, if the degree of possible difference were attenuated for the above reason, our sample size may have been underpowered for detecting a motionese preference via the pupil size measure. Future research might best address these possible reasons for the absence of a motionese preference in overall pupil dilation by presenting a larger sample of infants with videos of activity sequences that are somewhat less compelling to them.

## To what extent was infants’ pupillary response to segmental structure in action comparable to that previously observed in adults?

Infants in this research displayed systematic pupil dilation in response to a major action boundary in motionese demonstrations, much like what Tanaka and colleagues (in preparation) observed in adults viewing sports events such as basketball free-throws, tennis serves, and weight-lifting clean-and-jerks. It is worth noting, however, that infants’ boundary-related PDR differed in other ways from the comparable pattern observed in this one previous study with adults. For one thing, infants’ boundary-related PDR during motionese demonstrations was slower (occurring on average within a one-second region after the boundary) than adults’ (occurring on average within a half-second region after the boundary). Infants’ PDR timing was consistent with other evidence that infants’ pupil response is generally slower than adults’ (e.g., Verschoor et al., 2013, 2015; Zhang et al., 2018). Another difference was that a linear trend provided the best characterization of infants’ boundary-related PDR. In contrast, Tanaka and colleagues (in preparation) observed both significant linear and quadratic trends, but the quadratic trend was stronger and suggested that adults’ pupil diameter began to return to baseline shortly after their boundary-related PDR.

## Why might infants’ pupil diameter have remained high after a major action boundary?

One likely explanation is that infants’ pupil response – including the return to baseline – may simply be slower than adults’, thus delaying an observable return to baseline for infants relative what was observed in adults. However, a different possibility is also worth considering: visual examination of the videos indicated that actors often displayed considerable post-boundary movement in motionese demonstrations. In particular, caregivers depicted in the videos frequently exaggerated their motions just after boundaries (e.g., spreading their arms wide apart, displaying florid emotional facial expressions) in what appeared to be an effort to emphasize the fact that a boundary had occurred. In support of this observation, we found that luminance (i.e., pixel values, which are sometimes used as an index of motion change; e.g., Hard et al., 2011; Loucks & Baldwin, 2009) was greater both before, and after, action boundaries in motionese relative to adult-directed demonstrations. These features of the stimuli are of course characteristic of motionese, and could serve to both promote pupil dilation in relation to boundaries and sustain infants’ arousal, thereby reducing a tendency for pupil diameter to return to baseline levels after a boundary.

## Precisely what is it about motionese that facilitates infants’ detection of action boundaries?

One plausible hypothesis is that something specific to motionese demonstrations enhanced infants’ discovery of action boundaries. In the related domain of motherese – infant-directed speech – it has been suggested that specific characteristics of caregivers’ speech promote infants’ ability to find structure within the complex auditory stream (e.g., Kemler-Nelson et al., 1989; Gleitman, Newport, & Gleitman, 1984; Kuhl, 2004). Perhaps characteristics of motionese similarly facilitate infants’ detection of structure within dynamic activity. It might be the case that caregivers move and manipulate objects in ways that highlight structure within dynamic activity. For example, there is some existing evidence (e.g., Brand et al., 2013) that caregivers’ gaze toward infants coincides systematically with action boundaries. Further, as described above, caregivers in our study often exaggerated motions (e.g., spreading their arms) both before and after a boundary occurred. It is thus unknown whether infants’ increased pupil size was related to boundaries *per se* or to features of motionese that co-occurred with boundaries. Nevertheless, our results indicate that motionese attracts infants’ attention to regions of activity that are key for processing unfolding action; regions that will eventually come to be understood as action boundaries.

A second possibility is that motionese simply heightens infants’ attention, which in turn increases the chances that they will detect structure within unfolding activity. Prior research supports this hypothesis: when infants are in an attentive state (as indexed by heart rate) during stimulus presentation, they can more readily recognize that stimulus at later test (Frick & Richards, 2001; Richards, 1997), suggesting that processing is enhanced during attentive moments. Motionese might further increase infants’ attention by indicating to infants that a given demonstration is “for me.” Couching information in a “for me” social context is likely to facilitate learning (e.g., Akhtar & Tomasello, 2000; Baldwin, 2000; Baldwin et al., 1996; Csibra & Gergely, 2009; Sage & Baldwin, 2010), and seems to be either illustrative of, or closely related to, what Kuhl and colleagues (e.g., Kuhl, Tsao, & Liu, 2003; Kuhl, 2007) call “social gating” following a similar phenomenon in bird-song learning (e.g., Doupe & Kuhl, 1999; Kuhl et al., 2003). Perhaps contributing to this effect, Gergely, Csibra, and colleagues (Csibra & Gergely, 2009, 2011; Gergely, Egyed, & Király, 2007) suggest that pedagogical cues, which abound in motionese, signal to infants that they are being taught and, consequently, infants adopt a “pedagogical stance” that primes them to learn. Motionese may likewise promote infants’ adoption of a pedagogical stance, and enhance their detection of segmental structure in unfolding activity sequences.

These potential explanations for infants’ enhanced sensitivity to action boundaries within motionese are of course not mutually exclusive. It seems plausible, and perhaps even likely, that some or all of these various mechanisms operated in concert in the present research to facilitate infants’ detection of segmental structure within the dynamic motionese action displays. It is additionally likely that this is not unique to the domain of action; infant-directed modifications in general (e.g., in infant-directed speech or emotion as well as in action) may function similarly to promote infants’ detection of structure across domains.

## Are the observed effects unique to infants?

It is not clear that our findings about benefits of motionese should be interpreted as specific to infants. That is, motionese might benefit action processing in learners of any age, especially for novel action sequences. The activity sequences featured in prior work assessing adults’ pupillary responses to action boundaries were relatively familiar, and highly predictable, sports activities (Tanaka et al., in preparation). However, the activity sequences in the current research were designed to be novel to infants. If adult observers were shown completely novel activity, they too might display increased pupil size to boundaries only when the action is demonstrated with characteristics of motionese. Finally, as mentioned earlier, infants’ pupil dilations were protracted in comparison to that of adults – remaining high after action boundaries rather than rapidly returning to baseline. It is unclear whether this is unique to infants or if adults too would display a protracted dilation when action is novel or when it is displayed in motionese format. In sum, motionese might have a general tendency to enhance observers’ detection - at any age - of action boundaries in complex, novel activity sequences. Further research is needed to better understand the interplay of age, familiarity, and expertise in response to both motionese and natural pedagogy more generally.

## Conclusion

To conclude, this research offers a signal advance on the methodological front, validating pupillometry as a valuable window on infants’ on-line processing of dynamic human action. These findings thus also open a door to using pupillometry more generally to investigate infants’ learning as it unfolds in real time.

As well, our findings materially extend current understanding of ways in which motionese benefits infants’ action processing and thus their learning more generally. In particular, motionese (a) promotes infants’ detection of segmental structure within continuously unfolding activity, (b) assists them in gaining purchase on individual action units amenable to encoding, subsequent recognition, generalization/inference, re-enactment, and generative recombination, and (c) elicits their enhanced cognitive engagement to highly informative regions within the complex, continuous flow of behavior. All in all, this research helps to account for the apparent ease and rapidity with which infants acquire knowledge as they confront the dynamic complexity of their everyday world.

Assisting infants’ discovery of action units implies, in turn, that motionese scaffolds infants’ encoding of action in memory for later recall, which likely promotes their efficient processing of similar activity sequences when subsequently encountered. All in all, this research helps to account for the apparent ease and rapidity with which infants acquire knowledge as they confront the dynamic complexity of their everyday world.

# End Notes

1. A secondary goal of the grey screen was to match the luminance of the first frame of the video. However, due to an inadvertent change in luminance of the grey screen during stimulus creation, this goal was not satisfied. Because we were not specifically interested in infants’ PDR to content immediately following the grey screen, this issue was not problematic for interpretation of our results (and is thus explained in further detail in supplementary materials on the [OSF](https://osf.io/8mzhf/?view_only=dd7ffa2a52824f3496cfd47d09541a00)).

# References

Addyman, C., Rocha, S., & Mareschal, D. (2014). Mapping the origins of time: Scalar errors in infant time estimation. *Developmental Psychology*, *50*(8), 2030–2035. <https://doi.org/10.1037/a0037108>

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In *Becoming a word learner: A debate on lexical acquisition* (pp. 115–121). Oxford, U.K.: Oxford University Press.

Ambrosini, E., Reddy, V., Looper, A. de, Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking Ahead: Anticipatory Gaze and Motor Ability in Infancy. *PLoS ONE*, *8*(7), e67916. <https://doi.org/10.1371/journal.pone.0067916>

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

Bala, A. D. S., Keller, C. H., & Takahashi, T. T. (2018). Infant hearing assessed using the Pupil Dilation Response. In *NorthWest auditory and verstibular research meeting*. Seattle, WA.

Bala, A. D. S., Whitchurch, E. A., & Takahashi, T. T. (2019). Human Auditory Detection and Discrimination Measured with the Pupil Dilation Response. *Journal of the Association for Research in Otolaryngology*, *21*(1), 43–59. <https://doi.org/10.1007/s10162-019-00739-x>

Bala, A., Keller, C., Whitchurch, E., Baldwin, D. A., & Takahashi, T. (2016). Pupillary dilation as a hearing screening in adults and infants. Poster, Portland, OR.

Baldwin, D. A. (2000). Interpersonal Understanding Fuels Knowledge Acquisition. *Current Directions in Psychological Science*, *9*(2), 40–45. <https://doi.org/10.1111/1467-8721.00057>

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. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Sciences*, *5*(4), 171–178. <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. <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*.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants’ Reliance on a Social Criterion for Establishing Word-Object Relations. *Child Development*, *67*(6), 3135–3153.

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

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science*, *5*(10), 371–372.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In *Handbook of psychophysiology* (pp. 142–162).

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for “motionese”: Modifications in mothers’ infant-directed action: Infant-directed action. *Developmental Science*, *5*(1), 72–83. <https://doi.org/10.1111/1467-7687.00211>

Brand, R. J., Hollenbeck, E., & Kominsky, J. F. (2013). Mothers’ Infant-Directed Gaze During Object Demonstration Highlights Action Boundaries and Goals. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 192–201. <https://doi.org/10.1109/TAMD.2013.2273057>

Brand, R. J., McGee, A., Kominsky, J. F., Briggs, K., Gruneisen, A., & Orbach, T. (2009). Repetition in infant-directed action depends on the goal structure of the object: Evidence for statistical regularities. *Gesture*, *9*(3), 337–353. <https://doi.org/10.1075/gest.9.3.04bra>

Brand, R. J., & Shallcross, W. L. (2008). Infants prefer motionese to adult-directed action. *Developmental Science*, *11*(6), 853–861. <https://doi.org/10.1111/j.1467-7687.2008.00734.x>

Brand, R. J., Shallcross, W. L., Sabatos, M. G., & Massie, K. P. (2007). Fine-Grained Analysis of Motionese: Eye Gaze, Object Exchanges, and Action Units in Infant-Versus Adult-Directed Action. *Infancy*, *11*(2), 203–214. <https://doi.org/10.1111/j.1532-7078.2007.tb00223.x>

Brand, R. J., & Tapscott, S. (2007). Acoustic Packaging of Action Sequences by Infants. *Infancy*, *11*(3), 321–332. <https://doi.org/10.1111/j.1532-7078.2007.tb00230.x>

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. <https://doi.org/10.1016/j.tics.2009.01.005>

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149–1157. <https://doi.org/10.1098/rstb.2010.0319>

DeBolt, M. C., Rhemtulla, M., & Oakes, L. M. (2020). Robust data and power in infant research: A case study of the effect of number of infants and number of trials in visual preference procedures. *Infancy*.

Doupe, A. J., & Kuhl, P. K. (1999). BIRDSONG AND HUMAN SPEECH: Common Themes and Mechanisms. *Annual Review of Neuroscience*, *22*(1), 567–631. <https://doi.org/10.1146/annurev.neuro.22.1.567>

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, *8*(2), 181–195. <https://doi.org/10.1016/S0163-6383(85)80005-9>

Flores, S., Bailey, H. R., Eisenberg, M. L., & Zacks, J. M. (2017). Event segmentation improves event memory up to one month later. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(8), 1183–1202. <https://doi.org/10.1037/xlm0000367>

Frick, J. E., & Richards, J. E. (2001). Individual Differences in Infants’ Recognition of Briefly Presented Visual Stimuli. *Infancy*, *2*(3), 331–352. <https://doi.org/10.1207/S15327078IN0203_3>

Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant Pupil Diameter Changes in Response to Others’ Positive and Negative Emotions. *PLoS ONE*, *6*(11), e27132. <https://doi.org/10.1371/journal.pone.0027132>

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental Science*, *10*(1), 139–146. <https://doi.org/10.1111/j.1467-7687.2007.00576.x>

Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, *11*(1), 43–79. <https://doi.org/10.1017/S0305000900005584>

Goldinger, S. D., & Papesh, M. H. (2012). Pupil Dilation Reflects the Creation and Retrieval of Memories. *Current Directions in Psychological Science*, *21*(2), 90–95. <https://doi.org/10.1177/0963721412436811>

Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin*, *77*(5), 340–355. <https://doi.org/10.1037/h0032456>

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

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

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young Children Are Intrinsically Motivated to See Others Helped. *Psychological Science*, *23*(9), 967–972. <https://doi.org/10.1177/0956797612440571>

Hepach, R., Vaish, A., & Tomasello, M. (2015). Novel paradigms to measure variability of behavior in early childhood: Posture, gaze, and pupil dilation. *Frontiers in Psychology*, *6*. <https://doi.org/10.3389/fpsyg.2015.00858>

Hepach, R., & Westermann, G. (2013). Infants’ sensitivity to the congruence of others’ emotions and actions. *Journal of Experimental Child Psychology*, *115*(1), 16–29. <https://doi.org/10.1016/j.jecp.2012.12.013>

Hepach, R., & Westermann, G. (2016). Pupillometry in Infancy Research. *Journal of Cognition and Development*, *17*(3), 359–377. <https://doi.org/10.1080/15248372.2015.1135801>

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. <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. <https://doi.org/10.1037/a0014145>

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The preferential looking paradigm reveals emerging language comprehension. In *Methods for Assessing Children’s Syntax* (pp. 105–124). Cambridge, MA: MIT Press.

Jackson, I., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science*, *12*(4), 670–679. <https://doi.org/10.1111/j.1467-7687.2008.00805.x>

Kahneman, D. (1973). *Attention and Effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil Diameter and Load on Memory. *Science*, *154*(3756), 1583–1585. <https://doi.org/10.1126/science.154.3756.1583>

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications*, *2*(1), 341. <https://doi.org/10.1038/ncomms1342>

Kemler-Nelson, D., Hirsh-Pasek, K., Jusczyk, P., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, *16*(1), 55–68.

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

Kosie, J. E., & Baldwin, D. A. (2018). The need for a diverse audiovisual corpus of infant-directed behavior. Poster, Phoenix, AZ.

Kosie, J. E., & Baldwin, D. A. (in preparation-a). A corpus of infant- and adult-directed action.

Kosie, J. E., & Baldwin, D. A. (under revision-b). Dwell times disambiguate preschoolers’ attention to motion versus goal structure as dynamic action unfolds.

Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants’ attention and object exploration. *Infant Behavior and Development*, *32*(4), 437–444. <https://doi.org/10.1016/j.infbeh.2009.07.003>

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843. <https://doi.org/10.1038/nrn1533>

Kuhl, P. K. (2007). Is speech learning “gated” by the social brain? *Developmental Science*, *10*(1), 110–120. <https://doi.org/10.1111/j.1467-7687.2007.00572.x>

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096–9101. <https://doi.org/10.1073/pnas.1532872100>

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

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

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. <https://doi.org/10.1177/1745691611427305>

Levine, D., Buchsbaum, D., Hirsh‐Pasek, K., & Golinkoff, R. M. (2019). Finding events in a continuous world: A developmental account. *Developmental Psychobiology*, *61*(3), 376–389. <https://doi.org/10.1002/dev.21804>

Loewenfeld, E. (1993). *The pupil: Anatomy, physiology, and clinical applications*. Detroit, MI: Wayne State University Press.

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

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494–1502. <https://doi.org/10.3758/s13428-016-0809-y>

ManyBabies Consortium. (2020). Quantifying Sources of Variability in Infancy Research Using the Infant-Directed-Speech Preference. *Advances in Methods and Practices in Psychological Science*, *3*(1), 24–52.

Martineau, J., Hernandez, N., Hiebel, L., Roché, L., Metzger, A., & Bonnet-Brilhault, F. (2011). Can pupil size and pupil responses during visual scanning contribute to the diagnosis of autism spectrum disorder in children? *Journal of Psychiatric Research*, *45*(8), 1077–1082. <https://doi.org/10.1016/j.jpsychires.2011.01.008>

MATLAB. (2019). *Version 9.7.0 (r2019b)*. Natick, Massachusetts: The MathWorks Inc.

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

Meyer, M., Hard, B., Brand, R. J., McGarvey, M., & Baldwin, D. A. (2011b). Acoustic Packaging: Maternal Speech and Action Synchrony. *IEEE Transactions on Autonomous Mental Development*, *3*(2), 154–162. <https://doi.org/10.1109/TAMD.2010.2103941>

Miller, A. L., Gross, M. P., & Unsworth, N. (2019). Individual differences in working memory capacity and long-term memory: The influence of intensity of attention to items at encoding as measured by pupil dilation. *Journal of Memory and Language*, *104*, 25–42. <https://doi.org/10.1016/j.jml.2018.09.005>

Monroy, C., Gerson, S., & Hunnius, S. (2017). Infants’ Motor Proficiency and Statistical Learning for Actions. *Frontiers in Psychology*, *8*, 2174. <https://doi.org/10.3389/fpsyg.2017.02174>

Nagai, Y. (2019). Predictive learning: Its key role in early cognitive development. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1771), 20180030. <https://doi.org/10.1098/rstb.2018.0030>

Nagai, Y., & Rohlfing, K. J. (2009). Computational Analysis of Motionese Toward Scaffolding Robot Action Learning. *IEEE Transactions on Autonomous Mental Development*, *1*(1), 44–54. <https://doi.org/10.1109/TAMD.2009.2021090>

Nagai, Y., & Rohlﬁng, K. J. (2007). Can Motionese Tell Infants and Robots “What to Imitate”? *Proceedings of the 4th International Symposium on Imitation in Animals and Artifacts*, 299–306.

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

Nuske, H. J., Vivanti, G., & Dissanayake, C. (2015). No Evidence of Emotional Dysregulation or Aversion to Mutual Gaze in Preschoolers with Autism Spectrum Disorder: An Eye-Tracking Pupillometry Study. *Journal of Autism and Developmental Disorders*, *45*(11), 3433–3445. <https://doi.org/10.1007/s10803-015-2479-5>

Nuske, H. J., Vivanti, G., Hudry, K., & Dissanayake, C. (2014). Pupillometry reveals reduced unconscious emotional reactivity in autism. *Biological Psychology*, *101*, 24–35. <https://doi.org/10.1016/j.biopsycho.2014.07.003>

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436–469.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>

Poynton, C. A. (2013). *Digital video and HDTV: Algorithms and interfaces*. San Francisco, CA: Morgan Kaufmann.

Richards, J. E. (1997). Effects of Attention on Infants’ Preference for Briefly Exposed Visual Stimuli in the Paired-Comparison Recognition-Memory Paradigm. *Developmental Psychology*, *33*(1), 22–31.

Rohlfing, K. J., Fritsch, J., Wrede, B., & Jungmann, T. (2006). How can multimodal cues from child-directed interaction reduce learning complexity in robots? *Advanced Robotics*, *20*(10), 1183–1199. <https://doi.org/10.1163/156855306778522532>

Roseberry, S., Richie, R., Hirsh-Pasek, K., Golinkoff, R. M., & Shipley, T. F. (2011). Babies Catch a Break: 7- to 9-Month-Olds Track Statistical Probabilities in Continuous Dynamic Events. *Psychological Science*, *22*(11), 1422–1424. <https://doi.org/10.1177/0956797611422074>

Sage, K. D., & Baldwin, D. (2010). Social gating and pedagogy: Mechanisms for learning and implications for robotics. *Neural Networks*, *23*(8-9), 1091–1098. <https://doi.org/10.1016/j.neunet.2010.09.004>

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. <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. <https://doi.org/10.1080/15248370709336996>

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

Signmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2019). Afex: Analysis of Factorial Experiments. Retrieved from <https://CRAN.R-project.org/package=afex>

Sirois, S., & Brisson, J. (2014). Pupillometry. *Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679–692.

Sirois, S., & Jackson, I. R. (2012). Pupil Dilation and Object Permanence in Infants: PUPIL DILATION AND OBJECT PERMANENCE. *Infancy*, *17*(1), 61–78. <https://doi.org/10.1111/j.1532-7078.2011.00096.x>

Snow, C. E., & Ferguson, C. A. (1977). *Talking to children*. Cambridge, UK: Cambridge University Press.

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. <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. <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. <https://doi.org/10.1111/cdev.12247>

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236–257. <https://doi.org/10.1037/a0015631>

Tanaka, Y., Kosie, J. E., & Baldwin, D. A. (in preparation). Implicit measure of event segmentation using pupillary response.

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/>

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, *7*(1), 53–71. <https://doi.org/10.1207/s15327078in0701_5>

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review*, *22*(3), 757–765. <https://doi.org/10.3758/s13423-014-0747-6>

Verschoor, S. A., Paulus, M., Spapé, M., Biro, S., & Hommel, B. (2015). The developing cognitive substrate of sequential action control in 9- to 12-month-olds: Evidence for concurrent activation models. *Cognition*, *138*, 64–78. <https://doi.org/10.1016/j.cognition.2015.01.005>

Verschoor, S. A., Spapé, M., Biro, S., & Hommel, B. (2013). From outcome prediction to action selection: Developmental change in the role of action-effect bindings. *Developmental Science*, n/a–n/a. <https://doi.org/10.1111/desc.12085>

Williamson, R. A., & Brand, R. J. (2014). Child-directed action promotes 2-year-olds’ imitation. *Journal of Experimental Child Psychology*, *118*, 119–126. <https://doi.org/10.1016/j.jecp.2013.08.005>

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

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, *130*(1), 29–58. <https://doi.org/10.1037/0096-3445.130.1.29>

Zalla, T., Labruyère, N., & Georgieff, N. (2013). Perceiving Goals and Actions in Individuals with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *43*(10), 2353–2365. <https://doi.org/10.1007/s10803-013-1784-0>

Zhang, F., Jaffe‐Dax, S., Wilson, R. C., & Emberson, L. L. (2018). Prediction in infants and adults: A pupillometry study. *Developmental Science*, e12780. <https://doi.org/10.1111/desc.12780>