Seeing Is Not Enough for Sustained Visual Attention

Lei Yuan, Tian Linger Xu, Chen Yu and Linda Smith
Department of Psychological and Brain Sciences, and Cognitive Science Program, Indiana University
Bloomington, IN, 47405 USA

Abstract

Sustained visual attention is crucial to many developmental outcomes. We demonstrate that, consistent with the developmental systems view, sustained visual attention emerges from and is tightly tied to sensory motor coordination. We examined whether changes in manual behavior alter toddlers' eye gaze by giving one group of children heavy toys that were hard to pick up, while giving another group of children perceptually identical toys that were lighter, easy to pick up and hold. We found a tight temporal coupling between the dynamics of visual attention and the dynamics of manual activities on objects, a relation that cannot be explained by interest alone. In the Heavy condition, toddlers looked at objects just as much as did toddlers in the Light condition but did so through many brief glances, whereas in Light condition looks to the objects were longer and sustained. We discuss the implication of hand-eye coordination in the development of visual attention.

Keywords: Sustained visual attention; hand-eye coordination; multimodal; perception action; manual behavior; developmental systems

Introduction

The ability to focus attention on an individual object or event for a period of time, often in the face of distractions, is predictive of learning and general cognitive capacities (Lansink, Mintz, Richards, 2000; Ruff & Lawson, 1990). The ability to sustain visual attention undergoes substantial developmental change from infancy to early childhood with a steady increase in both total duration and the ability to resist distractions (Ruff & Lawson, 1990; Kannass, Oakes & Shaddy, 2006). Prior research on the development of visual attention has focused on both the effect of low-level stimulus-driven properties (exogenous) and the emergence of top-down internal control of attention (endogenous) (Colombo, 2001). However, like the development of many other cognitive capacities, visual attention interacts with and is influenced by other sensory modalities within the developmental system (Thelen & Smith, 1994). The ability to sustain attention may not emerge directly from the development of internal controls but rather externally—from the coupling of vision with physical action.

Within this view, visual attentional skills are not built solely on the development of vision; but rather are influenced, altered, and coordinated with other sensory modalities (Yu, Smith L, Shen, Pereira, & Smith T, 2009). One apt example is the demonstration that deaf children performed worse on a non-auditory visual attention task than their age-matched controls; but, deaf children who had cochlear implant for at least one year performed similarly to hearing children (Quittner, Smith, Osberger, Mitchell, & Katz, 1994). Because the visual attention task did not rely on auditory process at all, the deficit shown by deaf children without the implant was solely attributable to an impoverished capacity of visual attention. A history of having auditory experience with the aid of cochlear implant helped to build visual attention, which was then successively recruited to perform a task that did not rely on auditory information. Thus, visual information alone is not enough for building visual attention; the interaction of multiple sensory modalities may be critically involved in the pathways to internal control of attention.

We focus here on the role that manual behavior plays in the control of visual attention. It has long been recognized that the development of perception is driven by the development of motor behaviors (Gibson, 1979). For example, as infants achieve motor milestones (e.g., sitting, crawling and walking), they are able to receive different perceptual experiences (e.g., stably held objects, optical flow), leading to the development of various perceptual abilities such as object recognition and depth perception. Research has also shown that changes on the affordance of objects (or how they can be held) alters the visual input infants receive, which in turn alters the outcome of object recognition (Pereira, James, Jones, & Smith, 2010). Thus, changes in manual behavior may alter infants' visual attention on objects through the coordination between hands and eyes.

Recent research suggests that infants' hands and eyes are dynamically coupled during toy play (Pereira, Smith, & Yu, 2014; Yu & Smith, 2014, 2016) and this coupling may play a causative role in sustained attention. The natural learning environment is complex, often presenting multiple visually interesting objects in a cluttered setting. In these visually complex contexts, infants may rely on manual behaviors to externally select and maintain attention on a target of interest. For example, Pereira, Smith, & Yu (2014) have shown that infants own manual actions on objects help them to select target, reduce visual clutter, and create larger input images in the visual field, leading to sustained visual attention on objects, better object recognition and early word learning. In this sense, manual action helps to regulate and sustain visual attention. Conversely, it has been found that irregular attentional patterns in atypical development co-occur with perturbations between the visual and manual modalities (Koterba, Leezenbaum, & Iverson, 2014).

Both between- and within-person hand and eye coordination may contribute to a more mature control of visual attention in social contexts. Yu and Smith (2016) demonstrated that parent's visual attention often follows infant's hands to the object to which the infant was directed; this between person hand and eye coordination substantially prolonged infant's sustained attention on the same object during toy play. Thus, visual attention is not a sole product of vision or perhaps even the individual but also influenced by cross-person sensory-motor coordination. Consistent with this idea, recent findings suggest that joint attention is also dependent on the child's hand-eye coordination. One- to two-year-olds who showed more tightly coordinated hands and eyes were also better able to coordinate their attention with their parents and better able to sustain joint attention with a parent (Yu & Smith, 2013, 2014).

All previous findings linking hand-eye coordination in toddlers to sustained visual attention were correlational. Here we attempt to show a causal link. We manipulated manual behavior by giving one group of children heavy toys that were hard to pick up but that could be poked and touched in various interesting ways. We gave another group of children perceptually identical toys that were lighter and easy to pick up and hold. The expectation is that the duration of each individual hand contact will be less for the heavy toys than the light toys that can be picked up and held. However, if the toys are equally engaging—which we designed them to be—the total amount of hand contact may not differ between the two conditions. The key expectation then is on the dynamics of individual contacting events: more briefer touches (pokes and touches) in the heavy case and fewer but longer touches (poking and touching while holding) in the light case.

We illustrate the expectations under the two hypotheses in Figure 1. First, if infants' hands and eyes are dynamically coupled—when hands are on an object, eyes are more likely to be on the same object—then the different dynamic properties of the manual behaviors caused by the weights of object should lead to different dynamic patterns of visual attention, with less sustained attention in the heavy condition (Fig. 1, H1). Second, and in opposition, if visual attention is independent of hand actions (if visual properties of objects solely determine gaze) then, when presented with novel and interesting toys, infants would visually look at them and for similar durations at each looking event, irrespective to whether the object can be held or not (Fig. 1, H2). We expect that the results will support Hypothesis 1: children from both conditions will manually handle and visually attend to the objects for the same total amount of time over the whole play session, but the dynamic properties of gaze will differ considerably and aligned with the different dynamics of the hands.

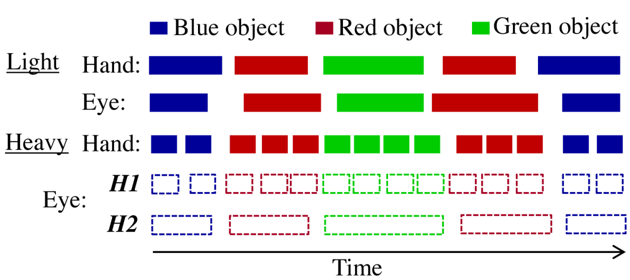

Figure 1. Two hypotheses for the looking patterns in the heavy condition

Methods

Participants

The final sample consisted of thirty-one parent – toddler (mean age = 21 months old, range = 18-25) dyads. Roughly half (16) of the dyads were assigned to play with light weight toys, while the other half (15) played with heavy weight toys. Children were recruited from a population of working and middle class families in a Midwestern town.

Stimuli

Two sets of six novel toys (12 in total) were developed from extensive pilot work to be engaging for manual play with moveable elements, openings, and possible actions. They were made of hardened clay, painted in red, blue or green, and were roughly the same size (9.5 x 6.5 x 5cm). The two sets were identical in terms of shape, size and color, with the only difference being their weights. The heavy set of toys was on average 1.4lbs, seven times heavier than the average weight of the light set, which was 0.21lbs.

Apparatus

Parent and child sat across a small table (61cm x 91cm x 64cm) (see Fig. 2). The child was strapped loosely into a small chair and the parent sat cross-legged on a pillow. Both participants wore head-mounted eye trackers with a sampling rate of 30 hz (positive science, LLC; also see Franchak et al., 2011). The eye tracker consists of a scene camera that captures the egocentric view of the participant, and an infrared camera that is mounted on the head, points to the right eye of the participant, and records the eye-in-head position (x and y) in the captured scene. Another high-resolution camera (recording rate 30 frames per sec) was mounted above the table and provided a bird's eye view that was independent of participants' movements.

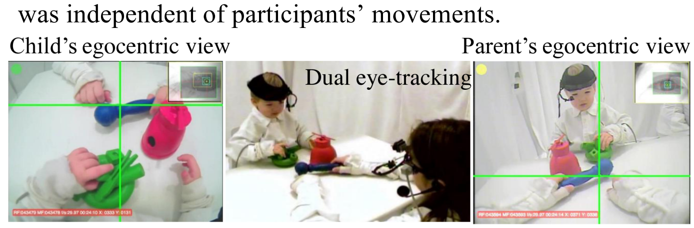

Figure 2. The experimental setup.

Procedures

To place the eye tracker on the child head, one experimenter attracted the child's attention with an interesting toy, while another experimenter put the eye-tracking gear low on the child's forehead. To calibrate the eye tracker, the experimenter directed the child's eyes toward an interesting toy, which were repeated 15 times while the toy was placed at various locations on the table. Parents were instructed to place the eye tracker on their heads. Parents' eye

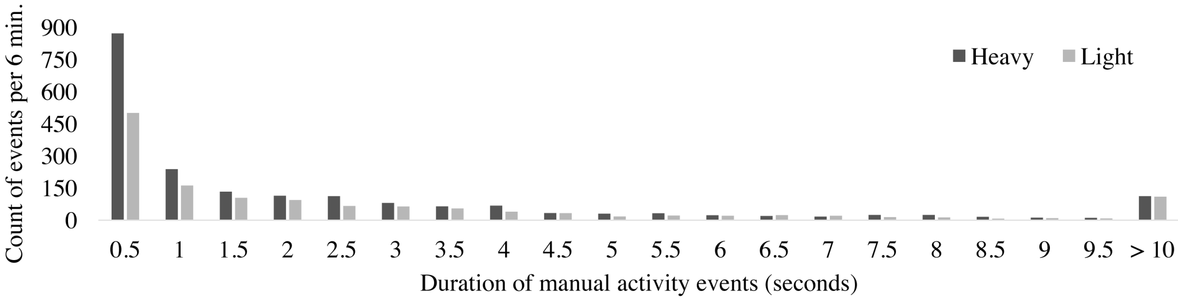

Figure 3. Frequency of manual activity events that last for different durations

tracker was calibrated in a similar way. After this initial set up, parents were told that the goal of the experiment is to study how parents and their toddlers interact during toy play, and were instructed to play with their toddlers as naturally as possible.

The free play session lasted for a total of 6 minutes that was composed of four trials with each lasted 1.5 minutes. The six novel toys were grouped into two sets (A and B) with each set having three different colored objects (red, blue and green). The sets were interleaved with the order of the sets counterbalanced across dyads (ABAB or BABA). At the end of each trial, the experimenter signaled parent with a clicking sound, and quickly replaced the old set of toys with a new set.

Coding

Three regions-of-interest (ROI) were defined for both the eye tracking data and the manual action data: the green, blue and red object. These ROIs were coded manually by coders who annotated frame-by-frame when the cross-hairs overlapped with any of the three ROIs. Another coder independently coded 10% of the frames with 95% agreement between coders. The final dataset consisted of a total of 203,316 frames.

Results

Because our manipulation was on the weight of objects, we first analyzed toddlers' manual activity. We then turned to visual attention as measured by gaze patterns. Finally, we examined the hand-eye coordination as a possible mechanism that drives the observed effects.

Manual activity

We defined manual activity event as any event during which the toddles' hands were in contact with any of the three objects (data from two hands were coded individually and then combined with a manual contact defined as either or both hands). Results showed that children in the heavy condition handled the objects for a comparable amount of total time as those in the light condition (Fig. 4), suggesting that overall the Heavy and Light versions of the toys were both manually engaging. There was no significant difference in the proportion of total time children in the light (M = 84%, SD = 7%) and heavy condition (M = 87%, SD = 6%) were in manual contact with the objects, t (29) = .24, p = .8. This is important to rule out the possibility that due to object weight, children in one of the conditions were more interested in the objects and played with them more than the other condition.

Children in the heavy condition (M = 21.31, SD = 7.26) produced manual activities at a higher frequency (count of events per minute) than those in the light condition (M = 15.48, SD = 4.42), t (29) =2.67, p = .01 (Fig. 3 & 4). But, children in the heavy condition (M = 2.63s, SD = .99s) spend less time in each manual activity event than those the light condition (M = 3.46s, SD = 1.38s), t (29) = 1.91, p = .06. Thus, it appears that children in the light condition would pick up and hold objects, resulting in many long manual activity events. In contrast, children in the heavy condition generated more short manual activity events because they can't hold the objects for a long time if at all, and would probably more often touch the object that sat on the table. This prediction was confirmed by the data: during manual activity events, compared to the heavy condition, children in the light condition had on average a larger visual image size (the size of the object in proportion to the entire visual field captured by the ego-centric view recording in the eye tracker), *Light*: M = 5.84%, SD = .99%, *Heavy:* M = 4.21%, SD = 1.17%, t (29) = 4.24, p < .001.

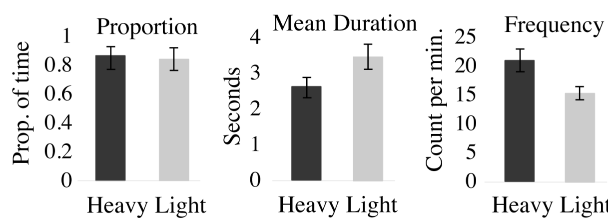

Figure 4. The proportion of time, mean duration and frequency of manual activity events for both conditions.

Despite the similarity in the total duration of manual activity, the way children handled the objects were different between the two conditions. Because previous studies have used 3 seconds as the threshold of sustained attention (Ruff & Lawson, 1990; Yu & Smith, 2016), here we defined sustained manual activity as any manual action that lasted for more than 3s. Consistent with our prediction, children in the heavy condition had significantly more short (less than 3 seconds) manual activity events per six minutes (session length) than did children in the light condition (*Heavy* = 1553, *Light* = 994); in contrast, the number of sustained manual action events per six minutes were comparable between conditions (*Heavy* = 493, *Light* = 399). Chi-square test of independence indicated that there was a significant

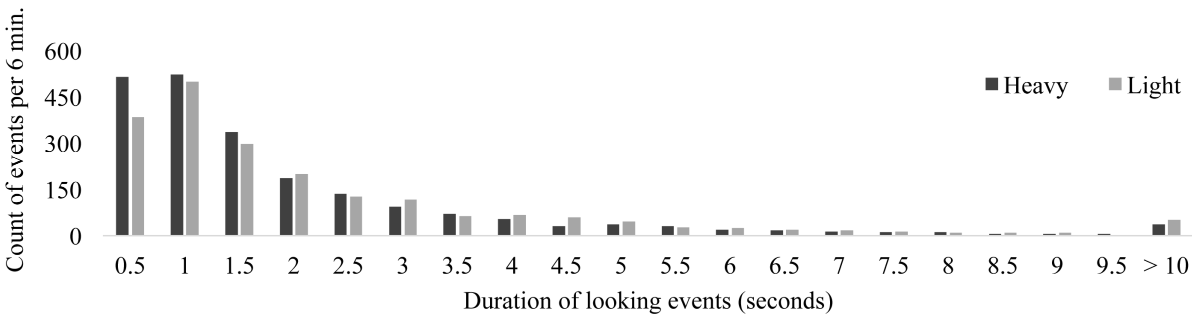

Figure 5. Frequency of looking events that last for different durations

relationship between the number of sustained manual activity events and the weight of objects, χ^2 (1, N = 3439) = 8.92, *p* = .002.

These results set the stage for answering the key question: given that hand dynamics differ, do eye dynamics—and sustained attention episodes—differ as well?

Visual attention

To analyze children's visual attention, we first examined all looking events during which the child had fixated on any of the objects (the ROIs). There was no significant difference in the proportion of total time children in the Light (*M* = 67%, *SD* = 2%) and Heavy conditions (*M* =65%, *SD* = 2%) looked at the objects, *t* (29) = .51, *p* = .61. Thus, children from both conditions were visually interested in the objects by this measure.

The mean duration of looking events was significantly lower in the heavy condition (*M* = 2s, *SD* = 0.44s) than the light condition (*M* = 2.43s, *SD* = 0.63s), *t* (29) = 2.21, *p* = .03. However, the looking events in the heavy condition (*M* = 20.55, *SD* = 5) had a slightly higher frequency (count per minute) than those in the light condition (*M* = 17.64, *SD* = 4.6), although this difference was not statistically significant, *t* (29) = 1.67, *p* = .1. Similar to the manual activity analysis and to previous research (Ruff & Lawson, 1990; Yu & Smith, 2015), we defined sustained looking as any looking event that lasted for more than 3 seconds. As shown in Fig. 5, children in the heavy condition had significantly more short (less than 3 seconds) looking events per six minutes (session length) than did children in the light condition (*Heavy* = 1789, *Light* = 1625); in contrast, the number of sustained looking events per six minutes were comparable between conditions (*Heavy* = 350, *Light* = 425). Chi-square test of independence indicated a significant relationship between the number of sustained looking events and the weight of objects, χ^2 (1, N = 4189) = 13.25, p = .0003.

Overall, the results of the looking patterns mirror the results from the manual activity: children in the heavy condition produced more rapid but frequent manual activity events, as well as more rapid but frequent looking events. By our hypothesis, the dynamic hand-eye coordination is responsible for the corresponding differences in the hand and eye patterns in the two conditions.

Hand-eye coordination

We propose that the result—that heavy condition had more short and rapid manual activity events, as well as more short and rapid looking events than the light condition—is driven by the hand-eye coordination of the child. In other words, because child's hands and eyes are closely coupled such that when hands are on the object, the eyes are also more likely to be on the same object—*sustained hand actions create and support sustained visual attention.* To demonstrate this link, we measured the durations of joint hand-eye to the same object. If this is the case, then we would expect to see more short but rapid hand-eye coordination events—the hands and eyes of the child were on the same object—in the heavy than the light condition.

As predicted, the mean duration of hand-eye coordination events was significantly lower in the heavy condition (*M* = 1.04s, *SD* = 0.25s) than the light condition (*M* = 1.33s, *SD* = 0.44s), *t* (29) = 2.26, *p* = .03. However, the hand-eye coordination events in the heavy condition (*M* = 17.83, *SD* = 3.93) had a significantly higher frequency (count per minute) than those in the light condition (*M* = 14.19, *SD* = 3.26), *t* (29) = 2.55, *p* = .01. Again, we used 3 seconds as the threshold to define sustained hand-eye coordination event and found that children in the heavy condition had significantly more short hand-eye coordination events per six minutes (session length) than did children in the light condition (*Heavy* = 1577, *Light* = 1116); in contrast, the number of sustained hand-eye coordination events per six minutes were comparable between conditions (*Heavy* = 134, *Light* = 151). Chi-square test of independence showed a significant relationship between the number of sustained hand-eye coordination events and the weight of objects, χ^2 (1, N = 2978) = 14.05, *p* = .0002.

General Discussion

When actively engaged with objects—the context for much real-world learning and problem solving—infants' visual attention is dynamically tied to their hand actions. The implications of this for the development of visual attention and for the underlying brain mechanisms are profound: this sensory motor coordination could be a core driving force for visual development, setting up the behavioral and neural networks for the mature control of visual attention (Byrge, Smith, & Sporns, 2014).

The direct connection between bodily movement, gaze direction and internal cognitive processing has been supported in many studies of adults' cognition. For example, it has been shown that bodily movement or direction of eye gaze serves as the basis for establishing deictic (pointing) reference to objects as well as the spatial relations between objects, suggesting that visual attention and action may share overlapping spatial referent frames (Ballard et al, 1997; Yuan, Uttal, & Franconeri, 2016). Manual actions can also directly guide or bias visual attention. The position of hands elicits unique neural responses in several brain areas and serves to prioritize visual attention (Makin, Holmes, & Zohary, 2007). Using a visual covert-orienting paradigm, for example, Reed, Grubb and Steele (2006) have shown that placing a hand on the side of the screen where a target would appear facilitated target detection, but the presence of visual anchors did not produce the same effect. This result suggests that adults have a hand-centered representation within peripersonal space (i.e., space that is close to a person's body), raising the possibility that children may have a similar or even stronger hand-centered representation in near space as they had shorter arms than adults and often hold objects very close to their body.

Manual action is a crucial way through which infants select and learn about the visual properties of objects in the world. Despite the complex and often cluttered real-world learning environment, the ego-centric view of infants suggests that they often attend to one dominant object at one time (Yu et al., 2009), which is crucial for developing visual attention to the detailed properties of objects. Importantly, although social partner occasionally brings an object in front of an infant's face, the predominant pathway through which infants create this optimal learning moments is through his or her own hand actions—it is hand actions that bring objects closer to the body and eyes, allowing for close examination of the various properties of the object, multiple sampling of the dynamic views of objects, leading to sustained visual attention and helping to build representations of the three-dimensional structure of objects (Bambach, Crandall, Smith, & Yu, 2016; Soska, Adolph, and Johnson, 2010).

The current study offers another pathway through which manual action exerts influence on visual attention—by changing the frequency and duration of looking events. Because hands and eyes are closely synchronized during play, the temporal characteristics of manual actions can influence those of vision. An analogous example is the demonstration that auditory input, particularly the rhythm of sounds, can facilitate visual learning for both adults (Iordanescu, Guzman-Martinez, Grabowecky, & Suzuki, 2008) and children (Bahrick & Lickliter, 2000). This multimodal learning not only provides redundant information to recruit sustained attention, but also capitalizes on the interconnection among sensory modalities—activities in one domain can influence and promote that of another domain. In this sense, the sensory motor coordination is the core driving force for the development of cognitive capacities.

Hand-eye coordination can help to build and integrate multiple neural networks that underpin cognitive development. Time-locked signals from perception and action not only afford the direct mapping between the physical properties of the object to the neuronal activity of the visual network, between the physical properties of the object to the neuronal activity of the haptic system, but also allow for cross-modality integration and enrichment: activity of the visual system and the activity of the haptic system are directly mapped to each other (Edelman, 1987; Smith & Gasser, 2005). For example, as one holds and manipulates an object, the neuronal activity of the visual system is time-locked to the activity of the haptic system—each different hold is linked to each unique visual representation of the object. As a result, a particular sight of an object may elicit its corresponding neuronally mapped action. For instance, in one visual recognition task, adults were shown a picture of a pitcher and answered the question "Is this a pitcher" by pressing either a left or a right button. Adults responded faster when the "yes" button was on the same side of the pitcher's handle, suggesting that the sight of the object may have elicited corresponding motor activity, facilitating the motor execution of button press on the same side (Ellis & Tucker, 2000).

This multimodal learning mechanism has important implications for development and learning. For example, manual actions can be leveraged to train a mature control of visual attention. One classic demonstration of this idea is the A-not-B error (Piaget, 1954). After repeating the sequence of seeing one object being hidden at location A and retrieving the object at location A several times, infants were shown the object being hidden at location B. Despite seeing the object being hidden in the new location B, infants continued searching at location A. However, changes in the motor actions or giving infants more motor experiences led to improved performance (Bertenthal, Campos, & Barrett, 1984). For instance, changing the manual behavior that children need to perform to approach the object through changes in posture (sitting vs. standing) led children to search at the correct location much often (Smith, Thelen, Titzer, & McLin, 1999).

In conclusion, the current study supports the developmental systems view of visual attention: visual attention emerges from the interaction among multiple sensory modalities, which are dynamically coordinated during moment-by-moment perception and action events to support cognitive development. In particular, the current study showed that changes in manual behavior alter the patterns of toddlers' visual attention during toy play. Further, we provided evidence that the hand-eye coordination is the underlying mechanism: toddlers' hands and eyes were dynamically coupled, such that when hands were on an object, the eyes were also likely to be on the same object. These results have implications for the research and development of visual attention, as well as the possibility to leverage on manual action as a way for training the control of visual attention.

Acknowledgments

We thank Melissa Elston, Charlotte, Wozniak, Melissa Hall, Charlene Tay, Madison Philips and Seth Foster for data collection. This research was supported by National Institutes of Health Grants R01 HD074601 and R21 EY017843.

References

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. Behavioral and Brain Sciences, 20(04), 723-742.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. Developmental psychology, 36(2), 190.

Bambach, S., Crandall, D. J., Smith, L. B., & Yu, C. (2016). Active Viewing in Toddlers Facilitates Visual Object Learning: An Egocentric Vision Approach.

Bertenthal, B. I., Campos, J. J., & Barrett, K. C. (1984). Self-produced locomotion. In Continuities and discontinuities in development (pp. 175-210). Springer US.

Byrge, L., Sporns, O., & Smith, L. B. (2014). Developmental process emerges from extended brain–body–behavior networks. Trends in cognitive sciences, 18(8), 395-403.

Colombo, J. (2001). The development of visual attention in infancy. Annual Review of Psychology, 52, 337–367.

Edelman, G. M. (1987). Neural Darwinism: The theory of neuronal group selection. Basic Books.

Ellis, R., & Tucker, M. (2000). Micro- affordance: The potentiation of components of action by seen objects. British journal of psychology, 91(4), 451-471.

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-Mounted Eye Tracking: A New Method to Describe Infant Looking. Child development, 82(6), 1738-1750.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston, MA: Houghton Mifflin.

Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. Psychonomic Bulletin & Review, 15(3), 548-554.

Kannass, K. N., Oakes, L. M., & Shaddy, D. J. (2006). A longitudinal investigation of the development of attention and distractibility. Journal of Cognition and Development, 7(3), 381–409.

Koterba, E., Leezenbaum, N. B., & Iverson, J. M. (2014). Object exploration at 6 and 9 months in infants with and without risk for autism. Autism, 18(2), 97-105.

Lansink, J. M., Mintz, S., & Richards, J. E. (2000). The distribution of infant attention during object examination. Developmental Science, 3(2), 163–170.

Makin, T. R., Holmes, N. P., & Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. Journal of Neuroscience, 27(4), 731-740.

Pereira, A. F., James, K. H., Jones, S. S., & Smith, L. B. (2010). Early biases and developmental changes in self-generated object views. Journal of Vision, 10(11), 1–13.

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. Psychonomic Bulletin & Review, 21(1), 178–85.

Piaget, J. (1954). The construction of reality in the child. New York: Basic Books.

Quittner, A. L., Smith, L. B., Osberger, M. J., Mitchell, T. V, & Katz, D. B. (1994). The impact of audition on the development of visual attention. Psychological Science, 5(6), 347–353.

Reed, C. L., Grubb, J. D., & Steele, C. (2006). Hands up: attentional prioritization of space near the hand. Journal of Experimental Psychology: Human Perception and Performance, 32(1), 166.

Ruff, H. A., & Lawson, K. R. (1990). Development of sustained, focused attention in young children during free play. Developmental Psychology, 26(1), 85–93.

Smith, L., & Gasser, M. (2005). The development of embodied cognition: Six lessons from babies. Artificial life, 11(1-2), 13-29.

Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: the task dynamics of the A-not-B error. Psychological review, 106(2), 235.

Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: motor skill acquisition facilitates three-dimensional object completion. Developmental psychology, 46(1), 129.Mullette-Gillman, Cohen, & Groh, 2005.

Thelen, E., & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press.

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. PloS one, 8(11), e79659.

Yu, C., & Smith, L. B. (2014). Linking Joint Attention with Hand-Eye Coordination – A Sensorimotor Approach to Understanding Child-Parent Social Interaction, 2763–2768.

Yu, C., & Smith, L. B. (2016). The Social Origins of Sustained Attention in One-Year-Old Human Infants. Current Biology, 1–6.

Yu, C., & Smith, L. B. (2016). Multiple Sensory-Motor Pathways Lead to Coordinated Visual Attention. Cognitive Science, 1–27.

Yu, C., Smith.L.B., Shen, H., Pereira, A. F.,& Smith,T.(2009). Active information selection:Visual attention through the hands. Autonomous Mental Development, IEEE Transactions on, 1(2), 141–151.

Yuan, L., Uttal, D., & Franconeri, S. (2016). Are Categorical Spatial Relations Encoded by Shifting Visual Attention between Objects? PloS one, 11(10), e0163141.