

Navigating Through the Experienced Environment: Insights From Mobile Eye Tracking

Koraly Pérez-Edgar¹ , Leigha A. MacNeill^{1,2}, and Xiaoxue Fu^{1,3,4}

¹Department of Psychology, The Pennsylvania State University; ²Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill; ³Center for Biobehavioral Health, Nationwide Children's Hospital, Columbus, Ohio; and ⁴Department of Pediatrics, The Ohio State University

Abstract

Researchers are acutely interested in how people engage in social interactions and navigate their environment. However, in striving for experimental or laboratory control, we often instead present individuals with representations of social and environmental constructs and infer how they would behave in more dynamic and contingent interactions. Mobile eye tracking (MET) is one approach to connecting the laboratory to the experienced environment. MET superimposes gaze patterns captured through head- or eyeglass-mounted cameras pointed at the eyes onto a separate camera that captures the visual field. As a result, MET allows researchers to examine the world from the point of view of the individual in action. This review touches on the methods and questions that can be asked with this approach, illustrating how MET can provide new insight into social, behavioral, and cognitive processes from infancy through old age.

Keywords

mobile eye tracking, stationary eye tracking, visual attention, development, learning

Many researchers are acutely interested in how people engage in social interactions and navigate their environment. To help experimentally control experiences in the laboratory and standardize individual events for analysis, researchers often take contingent three-dimensional people, objects, and events and create flattened two-dimensional stimuli. In other words, researchers typically generate data that really reflect how people observe social information and process the environment (Redcay & Schilbach, 2019). Researchers have a limited sense of how the mechanisms and processes they are interested in play out in the actual circumstances of people's everyday lives. Two-dimensional stimuli do not do a good job of capturing second-person (Redcay & Schilbach, 2019) or person-centered (Fu & Pérez-Edgar, 2019) attention mechanisms. Individuals do not typically respond to a steady stream of controlled, predetermined stimuli. Rather, they create for themselves an “experienced environment” (Pérez-Edgar & Fox, 2018, p. 360) by selectively and idiosyncratically engaging with

only a subset of the world around them. In our work, we wish to capture how individuals navigate and take in their environment from their own point of view.

Central to people's personalized experience is attention, which acts as a domain-neutral mechanism for development, learning, and daily functioning. Attention links experiences across time and contexts, setting the boundaries for stimuli that will be attended to, processed, interpreted, and acted on (Amso & Scerif, 2015). Thus, it is a gateway for social, behavioral, cognitive, and neural processes. Here, our focus is on visual attention, which reflects an important source of environmental information for most individuals. For research purposes, the goal is often to capture the focus of attention via eye gaze. We interpret gaze direction—where

Corresponding Author:

Koraly Pérez-Edgar, The Pennsylvania State University, Department of Psychology, 270 Moore Building, University Park, PA 16802
 E-mail: kxp24@psu.edu

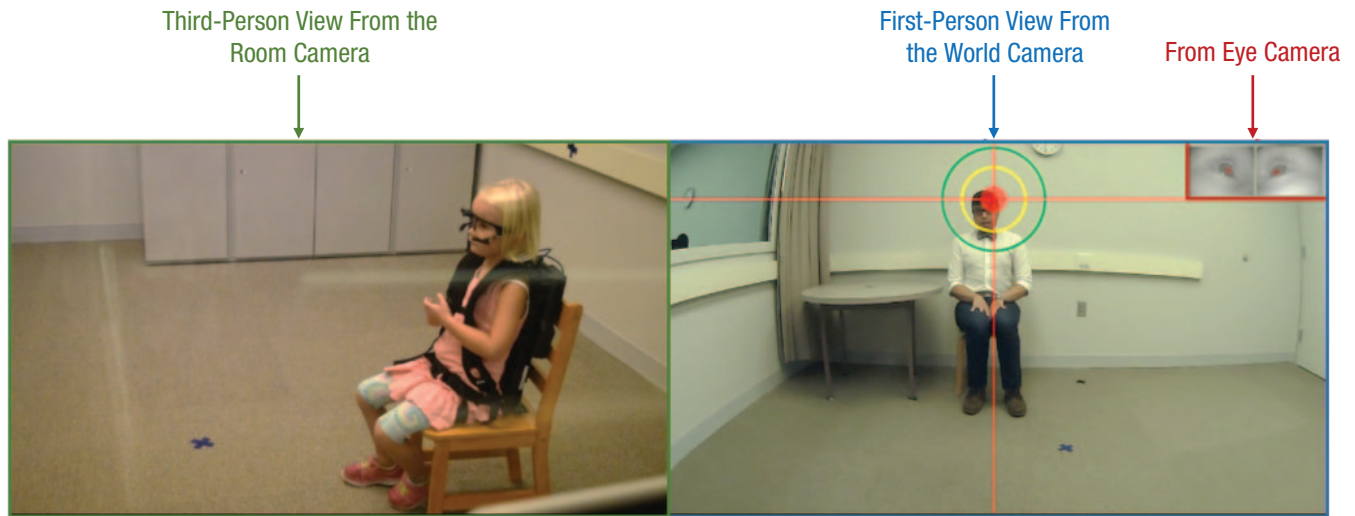


Fig. 1. Recordings from a mobile-eye-tracking paradigm. In this example, the child (left) is wearing eyeglasses equipped with three cameras, one of which looks out into the world while the other two point to each eye. The circles superimposed on the research assistant (right) indicate the child's point of fixation at a given moment (the red circle indicates the probable fixation point, and the yellow and green circles indicate the margin of error). The room recording (left) and the mobile eye-tracking recording (right) are synchronized off-line after the study visit. Researchers can code the child's eye gazes and the adult research assistant's behavior from the eye-tracking recording, as well as the child's bodily behavior from the room recording. Hence, the paradigm enables researchers to examine the child's visual attention processes embedded in active social interaction.

you are looking—as the focus of your visual attention (Colombo, 2001). Visual attention, in turn, is thought to reflect the focus of your current active processing.

From the first days of life, as our visual acuity steadily improves, we use vision to take in and make sense of the world. As we develop more complex and sophisticated motor skills, we continue to use vision to guide our actions (Franchak, 2020). Visual attention can help create the informational and experiential pipeline for development and learning. Visual exploration, often coupled with behavioral exploration, is opportunistic. Items appear in the world unexpectedly or take on new meanings as our goals shift. This exploratory process may become more purposeful and guided with age. A similar evolution takes place when we learn a new skill and transition from novices to experts. However, we are still captured by the salient and the unexpected well into adulthood and well after mastering a specific set of skills.

Currently, researchers typically rely on the perspective of a researcher observing from a distance with a camera or extract indirect information through behavioral responses elicited with computer-based tasks. We can come a bit closer to seeing the world through the eyes of another person by using mobile eye tracking (MET). MET uses a cap-mounted camera (for infants; Franchak, Kretch, Soska, & Adolph, 2011) or cameras embedded inside eyeglasses (Fu & Pérez-Edgar, 2019). With the eyeglasses, one camera looks out into the world while two cameras point to each eye. By superimposing the signals from all three cameras, we can track where and when individuals gaze at specific aspects of their environment

(Fig. 1). Then, from gaze, we can loop back to infer processes of attention. Although MET has been used for decades to capture attention processes in adults (Isaacowitz, Livingstone, Harris, & Marcotte, 2015), its use from the earliest months of life is a relatively new phenomenon (Franchak, 2017).

Attention Is a Domain-General Mechanism

Attention is a complex, multicomponent processing system. Researchers often divide attention into three distinct but interwoven components (Posner & Rothbart, 2007): orienting, vigilance, and executive attention. These three core areas of functioning allow people moving through their busy environments to notice an important event (vigilance), shift attention to the event (orienting), and then decide whether they need to act (executive). Researchers can get a sense of each component by examining patterns of visual attention over time and comparing them with the surrounding pattern of events and behaviors over time. For example, given the choice between climbing a structure to get a toy or getting a parent to help, we can observe a child's visual attention before, during, and after the choice is enacted to infer decision-making processes. Over time, we can see how people gain greater flexibility in responding to the world and a wider range of options when needing to respond and regulate (Pérez-Edgar, Taber-Thomas, Auday, & Morales, 2014). With children in particular, we can see how infants and toddlers are “captured” by the environment, whereas older children

can recruit visual attention to meet a challenge as needed (Shulman et al., 2009).

Why Do We Need MET?

The scope of the current review does not allow us to discuss in depth the wide numbers of methods and measures that can be used to capture and assess attention (for a more thorough discussion, see Fu & Pérez-Edgar, 2019). These methods, particularly when used together in the same study with the same participants, can help us better triangulate how attention emerges, is deployed, and then influences broad patterns of thought and behavior (LoBue et al., 2020). Of course, every method has its limitations. For example, if we videotape people to capture their visual attention, we can observe seminaturalistic movements. However, we need to constrain the environment and set up our cameras in such a way that it will be obvious where the person is looking. Capturing these data can require laborious coding procedures, only to generate fairly coarse approximations. If the scene is ambiguous or if the person moves out of the correct camera angle, the information can be lost entirely.

Researchers can gain control and precision by using stationary eye tracking (SET) to capture visual attention. In SET, an infrared signal, often mounted on a computer, detects patterns of eye gaze as people view stimuli on the screen. The eye-gaze detection is much more precise than what can be tracked via video, and it does not require an overt behavioral response from the participant. For example, with a camera, we can tell that a person may be looking at a face. With SET, we can now track eye gaze as it moves from the eyes, to the mouth, to the ears, and back again (Oakes, 2012). However, we are still limited to stimuli that can be presented on a screen, and the person must sit still.

Fundamentally, SET cannot provide a sense of the person's self-generated visual experience. When a single or limited set of stimuli is presented on a computer screen, we can capture attention in the absence of person-driven alternatives. We are less able to see individual differences in selection and engagement. Thus, our tasks deliberately and systematically take away choice (Ladouce, Donaldson, Dudchenko, & Ietswaart, 2017).

MET provides the opportunity to capture the real-time dynamic relation among attention, self-regulation, and behavior (Fu & Pérez-Edgar, 2019). As researchers, our goal is often to observe, predict, and when needed, modify behavior—broadly defined across motoric, cognitive, and socioemotional domains. These behaviors, in turn, are embedded in a world that is reciprocal, where there is an exchange between people and things and one event cannot perfectly predict what will happen next. MET provides an additional tool to improving

ecological validity in research while taking a novel person-centered approach. It allows us to depict the ebb and flow of attention patterns in real time as the surrounding events and the child's own behavior unfold. In addition, MET can be used to capture patterns of attention in the first months of life. In this way, MET broadens the age window for capturing visual attention processes from birth to old age and across levels of ability and disability.

How Do We Use MET?

The first step in deploying a relatively novel technique is to define the parameters needed to generate reliable and valid data. In contrast to SET tasks, the depth and distance of objects and events that the participants are looking at can vary from moment to moment (Franchak, 2017). Calibration accuracy can be compromised when the objects are either closer or farther than the distance between the participants and the targets used for initial calibration. Hence, it is likely that the initial calibration accuracy will not apply for the entire recording. To address this issue, it is important to assess potential changes in accuracy by performing multiple validation procedures during which a target is presented at different viewing distances. The calibration process was illustrated by Slone et al. (2018), and an example of the validation procedure is presented at <https://osf.io/qawc4/>.

MET also allows us to make more fine-grained studies of dynamic processes, observing microlongitudinal trajectories within a task. This work also increases our ability to use innovative analytic approaches that capture dynamic relations (Hollenstein, 2013) and individual time series (Corbetta, Guan, & Williams, 2012). New approaches are needed because MET data quickly become complicated. As noted, MET captures attention selection and attention in motion. Thus, we cannot necessarily predict what people will attend to and when and how often they will attend to a particular target. As a result, we have to first characterize the person-by-person idiosyncratic behaviors and attention patterns we see. This can be done by coding frame by frame across a visit or by coding for specific events (e.g., child looks at a stranger's face) and then coding for the behaviors and attention patterns that are evident before and after the event.

For example, state space grids (Hollenstein, 2013; Lewis, Lamey, & Douglas, 1999) can illustrate how two sets of variables, each with multiple components, can “move within a figurative space” (MacNeill, 2019, p. 67). Plotting the movement across the grid allows researchers to observe and quantify the interplay between constructs or behaviors of interest over the course of a task or interaction (Fig. 2). These grids can also help extract transitional patterns that suggest sequences of behavior moving across states.

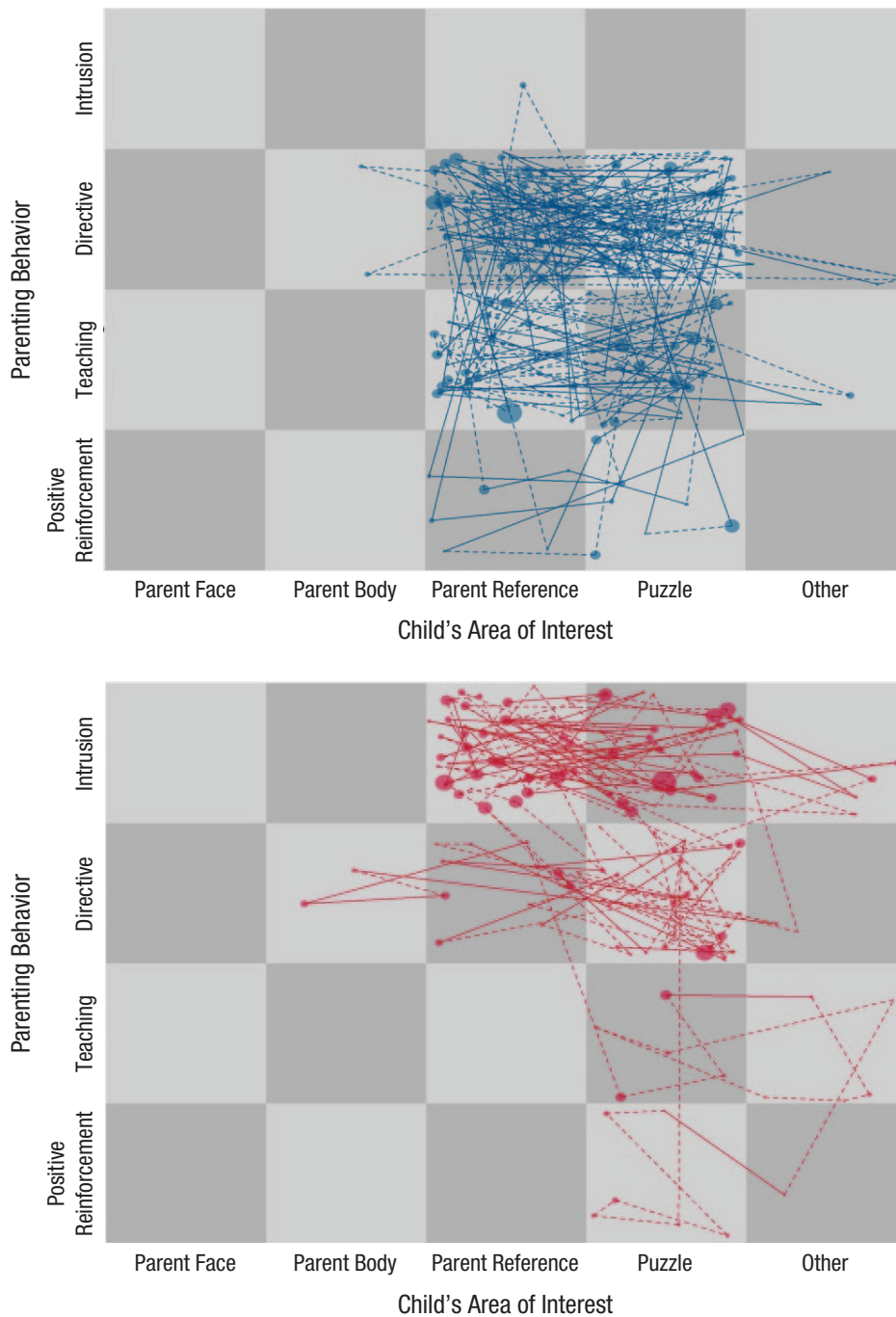


Fig. 2. Illustration of intraindividual variability of dyadic behavior using two state space grids, each showing a parent–child dyad during the course of a puzzle task. The child’s areas of interest (derived from mobile eye tracking) are shown as a function of parenting behaviors. The size of each node reflects the amount of time spent at that node, and transitions between states are illustrated with the lines that connect the nodes. Dotted lines indicate that there is at least one missing event between the two nodes the lines connect; solid lines indicate that there are no missing events. “Parent reference” refers to wherever the parent was pointing or referencing during the task.

What Are We Beginning to Learn With MET?

To illustrate one use of MET, MacNeill (2019) focused on the contours of parent–child interactions to examine

how parental traits and behaviors may be transmitted and shared in the moment to the child. A strong body of work has shown that oversolicitous and overprotective parenting can increase anxiety among children

(Hastings & Rubin, 1999), particularly if they are at heightened risk because of temperamental behavioral inhibition (Hastings, Rubin, Smith, & Wagner, 2019). MacNeill had parents and 5- to 7-year-old children complete a difficult puzzle task (<https://osf.io/qawc4/>). Areas of interest derived from MET were plotted on one axis, whereas the other axis captured parenting behaviors noted within the course of the parent-child puzzle task. Figure 2 shows the patterns of dyadic behavior for two unique parent-child dyads, capturing time spent at each square or state, representing the intersection of an area of interest with a specific behavior. The size of the node reflects the amount of time, and transitions between these states are illustrated with the lines that connect the nodes. Using these grids, we can identify patterns of dyadic exchanges that occur over the course of the task or interaction, potentially revealing dyadic processes that may influence development. In this study, MacNeill examined attractors, or states that pull the dyadic system from other states under particular conditions (Thelen & Smith, 2006). The data indicated that children's behavioral inhibition was positively associated with parent-focused and controlling-parenting attractor strength (i.e., longer amounts of time at each visit to states in which the child looked at the parent and the parent engaged in directive or intrusive behavior) but only when the parent also reported higher levels of anxiety.

Data derived from MET can lead researchers to question basic assumptions concerning the ways in which people navigate the environment and interact with others. For example, both children and adults spend far less time looking at each other's faces than we would predict on the basis of our everyday intuitions and studies that rely on SET (Franchak et al., 2011; Jung, Zimmerman, & Pérez-Edgar, 2018; MacNeill, 2019). These patterns clearly show that visual attention—where one is looking—cannot capture the full breadth of attentional processes that are central to everyday behavior. Anyone who has ever turned left to go to work, forgetting that you were supposed to turn right that day and first go to the bank, knows that we can see and process visual input without actually internalizing the meaning behind that input.

For example, Yu and Smith (2013) found that an infant's visual gaze to hands can be a better indicator of engagement than looking at a partner's face, depending on the task. At the other end of the age spectrum, Isaacowitz and colleagues (2015) found no difference in bias to positive valence if selecting from the environment with MET, although SET studies had consistently shown a positivity bias for older adults. And, in between, Fu, Nelson, Borge, Buss, and Pérez-Edgar (2019) had 5- to 7-year-old children complete a standard attention-bias

task with reaction time and SET measures. The same children also engaged in a standardized interaction with an adult stranger (modified from Buss, 2011). An association with behavioral inhibition was evident only in the MET task. In addition, a relation between SET task performance and MET behavior was significant only with increasing levels of behavioral inhibition.

(Interim) Conclusions

Clearly, people's daily environments are more complex than the protocols we construct in the research laboratory. Although we gain experimental control, we lose a connection to the world we wish to better understand. Social interactions rely on practical knowledge and experience with past social encounters. Thus, individuals will engage in social and communicative signaling, above and beyond the specific task or message, in a way not evident in noninteractive contexts, such as viewing a movie or pressing buttons in response to social stimuli.

The growing MET literature can help to bridge this gap. As equipment becomes lighter and more resistant to calibration loss, we can gather more data, over more time, as individuals roam more broadly. Thus, we can continue to apply MET to new contexts, including people with spider phobias confronting their fear (Lange, Tierney, Reinhardt-Rutland, & Vivekananda-Schmidt, 2004), adolescents taking in feedback during a stressful speech (Woody et al., 2019), or a young child exploring a children's museum (Jung et al., 2018). In this way, we can better capture the ultimate prize—understanding the mechanisms and processes that help us become who we are, embedded in our specific contexts, through our own eyes. This is the idiosyncratic, and embodied, experienced environment (Pérez-Edgar & Fox, 2018).

Recommended Reading

- Franchak, J. M. (2017). (See References). A review that outlines the ways in which researchers can use mobile eye tracking to capture developmental processes.
- Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). (See References). One of the first studies to apply mobile eye tracking to motor processes and movement in infants.
- Fu, X., Nelson, E. E., Borge, M., Buss, K. A., & Pérez-Edgar, K. (2019). (See References). Empirical study comparing attention patterns across reaction-time-based measured, stationary eye tracking, and mobile eye tracking.
- Fu, X., & Pérez-Edgar, K. (2019). (See References). A thorough theoretical and empirical analysis of methodologies available for assessing patterns of attention, particularly in children.
- Hollenstein, T. (2013). (See References). Illustrates and explains the use of state space grids to capture dynamic and interactive processes.

Transparency

Action Editor: Randall W. Engle

Editor: Randall W. Engle

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

Preparation of this article was supported by National Institute of Mental Health Grant R21 MH111980 (to K. Pérez-Edgar), National Institute of Child Health and Human Development T32 HD007376, Carolina Consortium on Human Development Training Program (to L. A. MacNeill), and internal funding from the Research Institute at Nationwide Children's Hospital (X. Fu).

ORCID iD

Koraly Pérez-Edgar  <https://orcid.org/0000-0003-4051-9563>

References

- Amso, D., & Scerif, G. (2015). The attentive brain: Insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience*, *16*, 606–619.
- Buss, K. A. (2011). Which fearful toddlers should we worry about? Context, fear regulation, and anxiety risk. *Developmental Psychology*, *47*, 804–819.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367.
- Corbetta, D., Guan, Y., & Williams, J. L. (2012). Infant eye-tracking in the context of goal-directed actions. *Infancy*, *17*, 102–125.
- Franchak, J. M. (2017). *Head-mounted eye tracking*. In B. Hopkins, E. Geangu, & S. Linkenauger (Eds.), *The Cambridge Encyclopedia of Child Development* (2nd ed., pp. 113–116). Cambridge, England: Cambridge University Press.
- Franchak, J. M. (2020). The ecology of infants' perceptual-motor exploration. *Current Opinion in Psychology*, *32*, 110–114.
- 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*, 1738–1750.
- Fu, X., Nelson, E. E., Borge, M., Buss, K. A., & Pérez-Edgar, K. (2019). Stationary and ambulatory attention patterns are differentially associated with early temperamental risk for socioemotional problems: Preliminary evidence from a multimodal eye-tracking investigation. *Development and Psychopathology*, *31*, 971–988.
- Fu, X., & Pérez-Edgar, K. (2019). Threat-related attention bias in socioemotional development: A critical review and methodological considerations. *Developmental Review*, *51*, 31–57.
- Hastings, P. D., & Rubin, K. H. (1999). Predicting mothers' beliefs about preschool-aged children's social behavior: Evidence for maternal attitudes moderating child effects. *Child Development*, *70*, 722–741.
- Hastings, P. D., Rubin, K. H., Smith, K. A., & Wagner, N. J. (2019). Parenting behaviorally inhibited and socially withdrawn children. In M. H. Bornstein (Ed.), *Handbook of parenting* (3rd ed., pp. 467–495). Mahwah, NJ: Erlbaum.
- Hollenstein, T. (2013). *State space grids: Depicting dynamics across development*. Cham, Switzerland: Springer.
- Isaacowitz, D. M., Livingstone, K. M., Harris, J. A., & Marcotte, S. L. (2015). Mobile eye tracking reveals little evidence for age differences in attentional selection for mood regulation. *Emotion*, *15*, 151–161.
- Jung, Y. J., Zimmerman, H. T., & Pérez-Edgar, K. (2018). A methodological case study with mobile eye-tracking of child interaction in a science museum. *TechTrends*, *62*, 509–517.
- Ladouce, S., Donaldson, D. I., Dudchenko, P. A., & Ietswaart, M. (2017). Understanding minds in real-world environments: Toward a mobile cognition approach. *Frontiers in Human Neuroscience*, *10*, Article 694. doi:10.3389/fnhum.2016.00694
- Lange, W., Tierney, K., Reinhardt-Rutland, A., & Vivekananda-Schmidt, P. (2004). Viewing behaviour of spider phobics and non-phobics in the presence of threat and safety stimuli. *British Journal of Clinical Psychology*, *43*, 235–243.
- Lewis, M. D., Lamey, A. V., & Douglas, L. (1999). A new dynamic systems method for the analysis of early socioemotional development. *Developmental Science*, *2*, 457–475.
- LoBue, V., Reider, L. B., Kim, E., Burriss, J. L., Oleas, D. S., Buss, K. A., . . . Field, A. P. (2020). The importance of using multiple outcome measures in infant research. *Infancy*. Advance online publication. doi:10.1111/inf.12339
- MacNeill, L. (2019). *Attention processes in context: A multi-method assessment of how parenting and temperament contribute to the development of attention* (Unpublished doctoral dissertation). The Pennsylvania State University, University Park.
- Oakes, L. M. (2012). Advances in eye tracking in infancy research. *Infancy*, *17*, 1–8.
- Pérez-Edgar, K., & Fox, N. A. (2018). Next steps: Behavioral inhibition as a model system. In K. Pérez-Edgar & N. A. Fox (Eds.), *Behavioral inhibition: Integrating theory, research, and clinical perspectives* (pp. 357–372). Cham, Switzerland: Springer.
- Pérez-Edgar, K., Taber-Thomas, B., Auday, E., & Morales, S. (2014). Temperament and attention as core mechanisms in the early emergence of anxiety. In K. Lagattuta (Ed.), *Children and emotion: New insights into developmental affective science* (Vol. 26, pp. 42–56). Basel, Switzerland: Karger Publishing.
- Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology*, *58*, 1–23. doi:10.1146/annurev.psych.58.110405.085516
- Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews Neuroscience*, *20*, 495–505.
- Shulman, G. L., Astafiev, S. V., Franke, D., Pope, D. L. W., Snyder, A. Z., McAvoy, M. P., & Corbetta, M. (2009). Interaction of stimulus-driven reorienting and expectation

- in ventral and dorsal frontoparietal and basal ganglia-cortical networks. *The Journal of Neuroscience*, *29*, 4392–4407. doi:10.1523/JNEUROSCI.5609-08.2009
- Slone, L. K., Abney, D. H., Borjon, J. I., Chen, C.-H., Franchak, J. M., Percy, D., . . . Yu, C. (2018). Gaze in action: Head-mounted eye tracking of children's dynamic visual attention during naturalistic behavior. *JoVE*, *141*, Article e58496. doi:10.3791/58496
- Thelen, E., & Smith, L. B. (2006). Dynamic systems theories. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (pp. 258–313). Hoboken, NJ: John Wiley & Sons.
- Woody, M. L., Rosen, D., Allen, K. B., Price, R. B., Hutchinson, E., Amole, M. C., & Silk, J. S. (2019). Looking for the negative: Depressive symptoms in adolescent girls are associated with sustained attention to a potentially critical judge during in vivo social evaluation. *Journal of Experimental Child Psychology*, *179*, 90–102.
- 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), Article e79659. doi:10.1371/journal.pone.0079659