Published in final edited form as:

Infancy. 2012 ; 17(1): 102–125. doi:10.1111/j.1532-7078.2011.0093.x.

Infant Eye-tracking in the Context of Goal-Directed Actions

Daniela Corbetta, Yu Guan, and Joshua L. Williams

The University of Tennessee, Knoxville

Abstract

This paper presents two methods that we applied to our research to record infant gaze in the context of goal-oriented actions using different eye-tracking devices: head-mounted and remote eye-tracking. For each type of eye-tracking system, we discuss their advantages and disadvantages, we describe the particular experimental setups we used to study infant looking and reaching, explain how we were able to use and synchronize these systems with other sources of data collection (video recordings and motion capture) in order to analyze gaze and movements directed toward 3D objects within a common time frame. Finally, for each method, we briefly present some results from our studies to illustrate the different levels of analyses that may be carried out using these different types of eye-tracking devices. These examples aim to highlight some of the novel questions that may be addressed using eye-tracking in the context of goal-directed actions.

Keywords

Infant eye-tracking; head-mounted and remote eye-tracking; reaching

In recent years, eye-tracking has become an increasingly popular tool amongst infant researchers. Significant advances, particularly in the area of automated eye-tracking, have made this technology much more user friendly and applicable to human infant populations than ever before (see Aslin & McMurray, 2004 for a history of eye movement techniques). Despite this fast and growing interest in infant eye-tracking, the use of this technology has been mostly applied to capturing infants' eye movements and gaze patterns when looking at objects or scenes depicted in two dimensions on a computer screen (Farzin, Rivera, & Whitney, 2010, 2011; Johnson, Amso, Slemmer, 2003; Johnson, Slemmer, & Amso, 2004; Gredebäck & von Hofsten, 2004; Quinn, Doran, Reiss, & Hoffman, 2009). Very few attempts have been developed to assess infant eye-tracking in the context of real, three dimensional (3D) objects or scenes, and even less efforts have been made to assess how infants' looking patterns relate to the selection and decision making processes involved in the planning and execution of their actions. Vision is important for detecting, identifying, and understanding what is in the surrounding world. The emergence of infant-friendly eye-tracking tools has offered an extraordinary extension to current investigative methods of infant cognition and has subsequently lead to significant advances in understanding the cues and perceptual processes infants use to detect, extract information, make inferences, and form predictions about events and scenes from two dimensional (2D) stimuli (i.e. Bertenthal, Longo, & Kenny, 2007; Falck-Ytter, Gredebäck, & von Hofsten, 2006; Farzin et al., 2010, 2011; Johnson et al., 2003, 2004). However, knowledge does not solely arise from looking at the world; the actions infants perform on the physical environment also play a fundamental role in the process of learning and development (Bojczyk & Corbetta, 2004;

Corbetta & Snapp-Childs, 2009; Smith, Yu, Pereira, 2011; Thelen & Smith, 2006; von Hofsten & Rosander, 1997; Yu, Smith, Shen et al., 2009). Thus, understanding how vision contributes to the formation of actions and how actions can in turn inform the process of looking and information-gathering represents a critical aspect of developmental science, particularly in delineating how our daily experiences with our environment contribute to shaping our behavior.

The study of how vision and action interact with one other to contribute to the development and refinement of behavior is not novel. For decades, infant researchers have designed clever experiments to manipulate and analyze the amount of visual information available to infants in order to assess its impact on infants behavioral responses (see Berthier & Carrico, 2010; Clifton, Muir, Ashmead, & Clarkson, 1993; Cowie, Atkinson, & Braddick, 2010; McCarty & Ashmead, 1999; von Hofsten, 1979; von Hofsten, Vishton, Spelke, et al., 1998, to cite a few). These studies and many others have provided valuable information, for example, on how vision contributes to the emergence and formation of new behaviors, on how vision is used in the control and guidance of actions, and on how perception and action coupling can lead to more refined, and more effective behavioral strategies. But using eye-tracking in the context of action can offer much more. Eye-tracking in the context of action can provide profound insights into the real-time dynamics of perception and action as infants interact with the environment by the seconds, minutes, and hours. Eye-tracking can provide detailed records of the process of visual exploration prior to, during and after the actions have been carried out. It can also inform about the particular cues that are being selected prior to and during actions, and it can address many aspects of how perception and action contribute to the cognitive processes of memory formation, decision making, action planning and movement correcting. Such rich and detailed information can directly inform theories of development, particularly by pointing at the fundamental coupling between cognitive, motor, and perceptual processes that are taking place as development unfolds.

The goal of this paper is to present methods of eye-tracking for use in the context of goal-directed actions using real, 3D objects with the hope to stimulate research in these much needed areas of study. While in adult research, the use of eye-tracking technology in conjunction with motor responses and patterns of action has already made huge strides in understanding the various contexts, cognitive and dynamic processes in which eye and hand “talk” to one another when solving tasks (see Flanagan & Johansson, 2003; Jonikaitis & Deubel, 2011; Horstmann & Hoffmann, 2005; Land, Mennie, & Rusted, 1999; Hayhoe & Ballard, 2005), much remains to be done in early development to comprehend even the most basic processes of how infants select visual information for action, how such information maps onto their changing action capabilities, and how vision and action together affect attention allocation and decision-making over the first year of life. In the sections that follow, we offer an overview of two automated eye-tracking methods that our laboratory has used to record eye-movements and gaze in the context of infants’ goal-directed actions: head-mounted eye-tracking and remote eye-tracking¹. For each eye-tracking method, we explain how the devices work, discuss their advantages and disadvantages, and provide information on how each method may be used in the context of goal-directed actions. We also provide information on how eye-tracking recordings can be combined and synchronized with other technological tools such as motion analysis and/or behavioral videos. Such tools are essential to capture, quantify, and interpret movement patterns in infants.

¹We will not discuss electro-oculography (EOG), another potentially viable method to record eye movements in conjunction to action patterns. To our knowledge, EOG with infants has only been used in conjunction with horizontal eye and head movements in the context of smooth visual pursuit (Aslin, 1981; von Hofsten & Rosander, 1997). EOG would present some real limitations if applied to more spatial dimensions (see Aslin & McMurray, 2004).

Head-mounted eye-trackers

These eye-tracking devices are called head-mounted because they require participants to either wear a cap or a band placed on the crown of their head, or wear goggles resting in front of their eyes. These devices are designed to be light-weight such that they are not too intrusive and do not impair the control of head movements. Furthermore, for use with infants and children, these devices come in different sizes such that they can fit the smaller skull dimensions of those particular populations. Eye-movements from these devices are tracked and overlaid on the visual scene by combining the recordings of two miniature cameras mounted on the head piece or goggle structure of the device. One camera, the scene camera, is facing forward and recording the scene present in front of the participant's visual field at all times. The second camera, the eye camera, records the corneal reflection of an infrared light usually projected on one eye. Through a calibration procedure (see below), both camera views are merged into one, allowing for the identification of the point of regard in the visual scene with reasonable accuracy.

Depending on the head-mounted model, the eye-tracking sampling rate can vary. Some infant head-mounted eye-trackers sample at a rate of 30Hz (that is, 30 images per seconds) such as the Positive Science head-mounted eye-tracker used by Franchak & Adolph (2011) and Yu, Smith, Fricker et al. (2011). Other systems like the ETL-500 (ISCAN, Inc.) that we used in our research (see figure 1a) samples at 60Hz. The higher the sampling rate, the better, because higher sampling rates will provide a more complete and more accurate output of the position/duration of the fixation points, amplitude/speed of the saccades, and will reduce sampling error (see also Gredebäck, Johnson, & von Hofsten, 2010 on this point). Despite differences in sampling rate, both systems provide comparable outputs: a video recording of the point of gaze on the scene, the time series of the gaze point position in pixels from the calibrated eye-camera, and the video times from the scene camera. With both systems, users have the option to code the visual patterns directly from the video outputs by transferring these outputs to a computer-based video coding system (i.e. Franchak & Adolph, 2011), or they can use the video outputs in conjunction with the gaze time series to perform more in depth spatio-temporal analyses of the visual patterns (i.e. Corbetta & Williams, 2011; Yu, Smith, Fricker, et al., 2011;).

As illustrated on figure 1a, the infant version of the ETL-500 head-mounted eye-tracker that we used in our laboratory had the two miniature cameras mounted on the visor of the cap. The eye camera was looking downward toward a lens located in front of the infants' left eye. Thus, with this model, eye-tracking was performed through the recording of the position of a tiny infrared light projected onto the cornea as reflected on that lens. Figures 1b and 1c provide examples of still frames of the video output provided by such a head-mounted eye-tracker once both camera views have been merged together. The intersection of the cross-hair over the scenes indicates where the infant is directing his/her regard on the scene. As shown in these examples, the infant is staring directly at the objects that the experimenter is presenting for reaching.

Calibration of the ETL-500 was performed through a 5 point procedure. The 5 points corresponded to the 4 corner and center points of a user defined, 2D vertical plane facing the child. Accurate calibration of the eye-tracker required the timely coordination between one experimenter facing the child and one experimenter running the interactive calibration software of the eye-tracker. Basically, the experimenter facing the child was presenting a small, visually attractive and sounding toy at one of the five predefined spatial positions. When the child was staring at the toy in that position, the experimenter running the calibration software was prompted to drag the intersecting point of a passive cross-hair appearing on the video output of the scene camera at the center of the toy position on the

video. This sequence was repeated 5 times for each toy position. Calibration accuracy and steadiness of the eye-tracking signal could be assessed immediately after entering the fifth calibration point, because the eye-tracker shifted automatically into the tracking mode and the cross-hair began moving dynamically on the video scene as a result of the infant active looking behavior. We checked calibration accuracy by enticing the child to visually track the toy being moved slowly in the predefined 2D area and verifying that the displacement of the cross-hair on the scene would remain on top of the moving target.

Advantages and disadvantages of the head-mounted devices

One obvious advantage of these systems is that the participants are not constrained to direct their visual attention to limited scenes in a single location. Participants can move their head freely to explore their surroundings while their eye movements are tracked continuously. Furthermore, some head-mounted eye-trackers will allow the participants to navigate in the 3D space or move from one room to another without being tied to a computer with connecting wires. Indeed, some of these head-mounted eye-tracker systems are wireless and power can be provided by standalone battery packs embedded in a jacket or a backpack worn by the participant (see for example Franchak & Adolph, 2011 with children). However, using head-mounted eye-tracking can also present a number of challenges.

Issues with infant participants—One challenge concerns using head mounted eye-tracking devices with human infant populations. Infants may not always be willing to wear these devices on their head, plus, setting the device on an infant's head such that it can fit snug enough to not move or slide around and provide accurate eye-recordings can be tricky. Also, infants may attempt to grab and remove the device from their head. One way we dealt with these issues in our research was to divert infants' attention by showing them an infant friendly video or entertaining them with a "busy box" while an experimenter was adjusting the eye-tracker on their head.

Infants also do not respond to instructions as older children or adults do. This can sometimes make the calibration of the head-mounted device difficult. Maintaining the infants' looking interest at the toy targets throughout the five point procedure, or maintaining infants' visual attention to one target position long enough to allow the experimenter at the computer to mark that gaze point can be tricky. To help with this issue, we made sure to have enough interesting toy targets available to maintain the infant attention throughout several repetitions of the calibration procedure.

Issues with the analysis and interpretation of the gaze data—Other limitations of head-mounted eye-trackers relate to analyzing and interpreting the data. Eye-tracking devices provide gaze accuracy within a 2D plane as defined by the calibrated area. Extending eye-tracking in the third dimension outside of the calibrated area can infringe upon the accuracy of the measurements. One possible solution to that problem involves performing multiple calibration procedures at different distances from the participant, but not all eye-tracking devices are amenable to that. Furthermore, when working with infants, given the already existing challenge of performing single calibrations with this population, performing multiple calibrations may not always be an available option. In all cases, when eye-tracking is used in 3D contexts to collect points of regard that are at different depths from the calibration area, researchers should always consider asking participants to look at specific distance points on the scene once the calibration is completed (i.e. some closer and some more distant) in order to determine the areas and depth range where gaze position on the scene remains reasonably accurate.

Another issue to keep in mind relates more specifically to the use of the gaze time series from the head-mounted eye-trackers (not just the video output). Time series allow for more detailed analyses of the eye patterns, however, because in head-mounted eye-trackers the head is free to move and the eyes are embedded in the head, the displacement of the cross-hair on the scene can be the product of different combinations of eye and head movements, and not just eye-movements. If the head is still, then we know that the movement of the cross-hair on the scene is the product of eye movements, but if the head moves, the cross-hair motion on the scene can be the result of either head movements, or a combination of both head and eye movements. In such cases, to really understand how participants use their eye and head to explore a scene, and to be able to parse out movements of the eyes (to identify a saccade for example) from the movements of the head, one would need to record head movements as well.

It is also important to control the testing environment to eliminate situations that could lead to ambiguities in the interpretation of the gaze signal, particularly when participants are looking at objects in the distance. Imagine that a participant is looking at a fenced backyard through a window with a head-mounted eye-tracker. In reality, the yard, the fence, and the window are at different distances from the observer, but on the 2D video output they appear as if they are on top of each other. If the observer looks at one area of the scene where all three objects converge and overlap, it becomes extremely difficult or even impossible to determine from the crosshair location on the video to where the participant is actually directing her visual attention, that is, whether she is looking at her reflection on the window, the fence just behind it, or the grass field in the background.

One last consideration concerning the interpretation of the gaze output from head-mounted eye-trackers, depends on the questions researchers want to address. Head-mounted eye-trackers provide reasonably good accuracy of *where* infants direct their gaze on the scene, but more refined eye movements as would be used to explore more detailed features of objects or a scene, are more difficult to extract. This is something we will address again later in this paper when we will discuss the use of remote eye-trackers. That said, aside from all of these challenges, using head-mounted eye-trackers with young populations and combining them with some measures of movement outcome to gain insights on how vision and action work together in early development is possible, as we illustrate here.

Using a head-mounted eye-tracker in the context of infant reaching

Our research laboratory has engaged in such an endeavor a few years ago in order to be able to capture the real-time dynamics of perception and action in the context of goal-directed actions and gain more insights into the visual-motor processes at work when young infants learn, plan, and execute reaching movements toward a seen target. For decades, vision has been considered critical for the emergence of infant reaching (Bushnell, 1985; Piaget, 1952; White, Castle, & Held, 1964), however, in recent years, the question as to whether infants rely on vision for learning to reach has become somewhat controversial. Indeed, a number of studies have demonstrated that young infants can reach in the dark without needing to see their hand (Clifton, Muir, Ashmead, & Clarkson, 1993; Clifton, Rochat, Litovsky, & Perris, 1991; Clifton, Rochat, Robin, & Berthier, 1994), and others have revealed that, initially, infants rely more on proprioception than vision to direct their arm to the target (Berthier & Carrico, 2010; Carrico & Berthier, 2008; Thelen et al., 1993). Thus, our interest in recording eye movements in relation to arm movements arose from wanting to reassess and better understand how perceptual-motor mapping takes place in the development of infant reaching. Some initial basic questions we wanted to address were as follows: Do infants look at the target before and during reaching? Do they initially just look at the object

location or do they also identify basic object features such as their shape or orientation? How do infants integrate this visual information into their movement when reaching?

Head-mounted eye-tracking appeared to be a natural solution to address these questions. First, with head-mounted eye-trackers, object presentations in the infant's reaching space will not obstruct the recording of the eye, since the eye-tracking camera is close to the eye of the observer (this is something to consider when using remote eye-trackers, as we will explain later, because objects located between the child and the remote eye-tracker can interfere with gaze recording). Second, many of these head-mounted eye-trackers, which were initially intended for adult studies, have been designed to communicate with other computerized pieces of equipment. This was an important factor in our choice as we wanted to be able to synchronize the infant gaze at the target with the infant arm movement to the target. The particular head-mounted eye-tracker we used in figure 1 could communicate with our motion analysis system through its software and could be triggered remotely such that when we pressed a key on the computer to begin data collection with our motion analysis system, the motion software sent a signal to the eye-tracker software to trigger data collection in that system as well. Furthermore, a frame counter appearing on our video recordings of the reaching behavior and also connected to our motion analysis, was similarly triggered remotely by our motion tracker when we began data collection. Thus, by triggering our motion analysis system, all three systems (motion analysis, eye-tracking, and video frame counter) began to collect data simultaneously, allowing us to compare and relate events and data points from all videos and time series sources within the same time frame. In addition, the sampling rates between sources of equipment (reaching video 30Hz, eye-tracker 60Hz, motion analysis 120Hz) were multiples of one another such that synchronization and data reconstruction between sources could be facilitated and done reliably (we will discuss in the next section how to synchronize different pieces of equipment when the equipment used is not designed to communicate with other sources of data collection).

Figures 2 and 3 each display an example of time series outputs from a 9-month-old infant in that study, plotted in MATLAB (The MathWorks, Inc.), with the eye-tracker data (top graph) and the motion analysis system data (3 bottom graphs) from the same trial plotted as a function of their common time scale. The symbols on the figures identify particular events in our task presentation and the infant's responses were coded from the eye-tracking video and the reaching behavioral video, respectively. Basically, trials always began with the target objects hidden behind a table, out of the infants' view. When data collection began, the experimenter facing the infants brought the target object into the infant's view and held it steadily and slightly out of the infant's reaching space for a brief duration (1 second) before moving it closer to the infant for reaching. The symbols on figures 2 and 3 indicate when these specific events occurred: the circle symbol indicates when the object presented by the experimenter came into the infant's view, the triangle indicates when the object was held steadily and slightly out of the infant's reaching space, and the two diamond symbols mark the onset and offset of the reaching arm, that is, from the moment the infant began moving the arm toward the target until the time the hand first contacted the target. Our rationale for presenting the task in such way was to ensure that infants would be given some opportunity to visually scan the object prior to reaching, something that infants would not necessarily do if the object is presented immediately within their reaching space.

The two trials displayed on figure 2 and 3 were picked because they are quite representative of the kind of object-related looking and reaching responses we obtained in this study. In figure 2, the object was a 13 cm long infant spoon and it was presented to the child horizontally. In figure 3, the object was a sphere measuring 5 cm in diameter. The elongated objects (i.e. spoons and rods presented vertically or horizontally) elicited more frequent

target-directed visual scanning prior to reaching and also sometimes during reaching, while presentation of the 5 cm spherical toys rarely resulted in such visual scanning. Differences in these object-specific looking responses can be seen clearly on the top graphs of figure 2 and 3. In figure 2, the solid line which corresponds to the horizontal gaze patterns, reveals a shift in gaze to the right (up on the graph) and then to the left (down on the graph) happening shortly after the object was positioned steadily in front of the child (after the triangle mark). This was an object scan performed along the length of the spoon that preceded the reaching action. At second 2, on that same graph, the object was moved into the infant's reaching space and following the reach onset (at the first diamond mark), we can see the infant was performing another horizontal scan on the object that occurred during the reaching action. The top graph of figure 3, displaying the looking pattern at the spherical toy, in contrast, reveals a gaze time series that looks mostly flat between seconds 1 and 3 in the trial, that is, throughout the object presentation and reaching time windows. In fact, it appeared that for most of the spherical toy presentations, infants used a single fixation, usually at the center of the object (see figure 1b), and maintained it for most of the pre- and during reaching action.

The movements that infants used to reach for these objects when linked to the looking patterns were also revealing. The three bottom graphs of figures 2 and 3 report the displacement, velocity, and wrist rotation of the reaching hand toward these 2 objects. These exemplar trials illustrate a number of features that are typical of 9-month-old infants who are capable of integrating visual information into the planning and organization of their movement: first, the duration of the reach (distance between the 2 diamond marks) was longer for the spoon target (figure 2) compared to the spherical target (figure 3); second, the deceleration phase of the movement, which was measured from the point of the highest velocity peak on the velocity profile, was much longer for the spoon (figure 2) compared with the spherical toy (figure 3) where the maximum velocity peak occurred approximately midway into the reach; and third, the arm rotation during the reach was much greater for the spoon than for the spherical toy. Again, recall that the spoon in this exemplar trial was presented horizontally. The arm rotational data for this object (figure 2) show that following reach onset, the hand first began adopting a 90 degree orientation, but then rotated back toward a 0 degree orientation to line up with the orientation of the spoon. Note that the onset of the hand rotation toward the 0 degree orientation coincided with the horizontal visual scan performed on the object during reaching, as if vision and detection of the object orientation, played a significant role in helping to map the hand orientation onto the object orientation, and also may have contributed to the longer movement deceleration phase to allow such mapping to occur. If we compare these data to the hand rotation for the spherical object in figure 3, we see a much smaller variation in hand rotation amplitude and a shorter deceleration phase, which makes sense since spherical targets are non directional and thus require less motor adjustment to be grasped. These exemplars are very much in line with the interpretation that 9-month-old infants are capable of visually detecting the physical features of objects (different types of scanning for different objects), and that they are capable of integrating these different features into the planning and execution of their movements. Interestingly, data from this study, which spanned the reaching behavior of infants aged 6 to 11 months, revealed a much more nuanced developmental picture in this process of perceptual-motor mapping, with some infants attending the physical features of objects prior to reaching and some not at all, or some attending the physical features of objects only during the reach. These looking patterns fundamentally affected the way infants organized their reaching movements and raised important issues about the attentional processes at stake when infants are acting on the world (Corbetta & Williams, 2011).

Remote eye-trackers

These types of eye-trackers are called remote because they are typically placed in front of the participants, facing them. They track eye movements directed at scenes that are spatially fixed and usually much more restricted than the head-mounted eye-trackers. One most common use of these devices is to record gaze to images or animations displayed on a computer screen generally located right above the eye-tracker, but these types of eye-trackers (if they are not embedded in the computer screen itself) can also be used to record gaze at more distant scenes. For example, gaze may be recorded toward a projection screen located 2–3 meters away from the participant (see Guan & Corbetta, 2010), or at 3D objects located either within or beyond reaching distance. These systems offer advantages to some of the challenges associated to the head-mounted eye-trackers described above, but as for every sophisticated piece of equipment, they also come with their own limitations.

Advantages and disadvantages of remote eye-tracking devices

Remote eye-trackers are particularly attractive to infant researchers because they do not always involve that the participant be instrumented with a head cap or device to record the eyes, nor do they require specific head stabilization. Indeed, since these remote eye-trackers track eye movements toward scenes that are predefined, fixed, and more confined spatially, gaze data are usually recorded when the head is still and oriented straight toward the stimulus. As a result, the time series are not as affected by head movements. One important drawback, however, compared to the head-mounted systems, is that if participants turn their head away from the eye-tracker and the targeted scene, or lean to one side so that only one eye signal is available, then the eye-tracker loses track of the participant's pupil which results in missing data. When the head returns to midline, some remote eye-trackers are quite good at recapturing the eyes immediately, others couple their eye-tracker to a small motion tracker placed on the participant's head (see Aslin and McMurray, 2004) or track a small sticker placed on the participant's forehead to facilitate eye recapture. The loss of eye-tracking data during head turn or other behaviors preventing eye capture can be a significant problem when working with infants. Indeed, infants are more distractible and more likely to look away from the target scene. This raises a critical issue for the analysis of these infant eye-tracking data, because not all infants will yield sufficient or comparable amounts of data. This implies that researchers set clear rules and cut-off thresholds to define how much missing data can be acceptable and how much actual data is needed to obtain valid and reliable results in a given testing situation (see also the eye-tracking guidelines set by Oakes (2010) for infant researchers).

Another advantage of remote eye-trackers is their rate of sampling can be higher. Many of the infant-friendly systems currently available on the market offer sampling rates that range from 60 to 300 Hz, and some companies even offer systems that can sample up to 2000Hz (although we are not aware of infant researchers using such systems). This adds significant details and precision in gaze recording compared to the infant head-mounted eye-trackers. Improvement in the quality of gaze recordings is also enhanced from more user-friendly calibration procedures that can yield more accurate gaze recordings. Furthermore, since there is no need to adjust the eye-tracker on the infant head, the experimenters can move into the calibration and data collection phases much more quickly, which is certainly an advantage when working with populations that have short attention spans. Finally, the software provided by some of the eye-tracking companies also provide video outputs and text exports of the gaze patterns on the scene/stimulus that are easier to code and interpret than plain crosshairs. For example, some software packages apply algorithms and filters to the original time series such that fixation points and saccades overlaid on the scene may be easily distinguished. Below we describe how we used our remote eye-tracking system to record eye-movement in the context of infant reaching for 3D objects.

Using a remote eye-tracker in the context of infant reaching

Two major motivations led us to transition from our head-mounted to a remote eye-tracker to study infant looking patterns in the context of reaching. First, our head-mounted eye-tracking study led to an attrition rate of 64% (out of 75 infants recruited for this project, we successfully collected useable data in 27 infants). Many infants would not tolerate wearing the hat, or we could not calibrate the system satisfactorily. Second, in that prior study, we could not determine with strong confidence where exactly infants were directing their gaze on the objects; we could distinguish reliably whether infants were paying attention to the objects and whether they globally scanned the orientation and shape of the objects, but more precise spatial analyses of the looking patterns were difficult to carry out. In this new study, we continued to be interested in how infants use vision to pick up information about the orientation/shape of objects to map it to their actions, but more specifically, we wanted to evaluate whether infants could make more precise reaching decisions based on their prior history of looking patterns on the object such as deciding *where* to grasp the objects exactly. For instance, if the task involved reaching for a cup or a tool-like object with a handle, we wanted to assess whether infants would visually identify all or only parts of these objects (i.e. cup versus handle), if they scrutinized some areas of the objects more than others (i.e. the handle), and if they used that visual information to decide where to direct their hand to pick up the objects, that is, whether they went on to grasp the object from the handle or not? Remote eye-trackers held the potential to allow us to perform such analyses. Furthermore, as in our prior study, we wanted to collect movement kinematics in order to link the trajectory and speed of the reaching movements to the observed visual patterns and grasping outcomes, but in order to reach those goals, we needed to solve a number of issues.

A first issue was to figure out how to present the objects to the infants such that they would not obstruct eye tracking, and a second issue was to find a way to create a calibrated 2D spatial frame of reference in order to make sure that accuracy of gaze on our 3D objects would be maximally preserved when the objects were presented to the infants. The remote eye-tracker we use is a Tobii x50 (Tobii Technology, Inc.), a stand-alone eye-tracker that is not embedded in a computer screen, but, it can be paired with a computer screen if we want to use it that way. Also, this eye-tracker is designed to record the eyes from below, but depending on the setup one could also consider using it to record the eyes from above by inverting it upside-down. We used the eye-tracker as designed for its original use, by recording the eyes from below, and defined a 2D perpendicular area directly above the eye-tracker as our object presentation area (see figure 4A & B). This area was accurately defined by an opening within in a large black wooden standing board that surrounded the eye-tracker and the small table supporting the eye-tracker. Consistent with our prior study, objects presented in that pre-defined area, were out of the direct reach of infants (see figure 4B). Note that a similar setup would also work if we decided to present the objects immediately within the infant reach without obstructing the eye-tracker, but pilot work confirmed that when objects were presented within immediate reach, infants do not systematically take time to scan the object prior to reaching. This could introduce huge inter-individual differences with regard to looking time and reaching to the objects between infants who would reach as soon as they see the object, and others who would take time scrutinizing the object first before reaching. By holding the objects out of reach, we had greater control on looking time at the object and we were more readily able to identify how and where infants directed their visual attention on the objects prior to reaching.

We performed eye-tracking calibration by fitting a flat screen computer monitor through the board opening, directly above the eye-tracker, where the objects were to be presented (see figure 4A). Calibration was performed by running the 5 point calibration procedure provided by the Tobii software (Clearview or Studio) where an attractive figurine moved and sounded

in concert successively at the 4 corners and center of the screen. Once calibration was achieved, we removed the computer monitor and placed a double layer of black curtains in front and behind the opening of the board to conceal it and provide a solid background that would blend with the surrounding board during object presentations. The front curtains, which were closed at the beginning of each trial, were mounted on a hidden rail and could be opened from behind the board by pulling strings, thereby providing a clean start to each object presentations. To capture the scene and record the object presentations from the infant view, we used a digital scene camera, placed behind the infant as shown on figure 4A. These digital video recordings of the object presentations were fed on-line to the eye-tracking software, which permitted the production of video outputs of the scene with the corresponding points of visual regard overlaid on it (see figures 5B and 6).

Our next issue to solve was to synchronize the eye-tracker with our motion analysis system and behavioral video recordings. Unlike our head mounted eye-tracker, this remote eye-tracker was not designed to communicate with other pieces of equipment; therefore, we devised our own custom-built synchronization system using inexpensive hardware. We proceeded by pairing two systems or sources of recording at a time. First, we made sure that a common time marker could be inserted simultaneously on both our video sources, that is, on the behavioral cameras recording the reaching behavior of the infants and the video output of the visual field from the scene camera connected to our eye-tracker. To do this, we equipped the lenses of each camera view (the reaching and scene cameras) with a small diode facing inward toward the lens (see figure 5A). These diodes were connected via long cords to a single command box and they would light on at the same time for a duration of one second at the press of a button on the box. These diodes may be seen at the edges of both the scene and reaching camera views on figure 5B & C. When collecting data, we pressed the button on the command box at the beginning of each trial. This allowed us to synchronize the two video sources according to a common time frame at a later time by looking specifically at the video frames when the diodes were briefly lit.

Our motion tracker, which as explained before, was connected to a frame counter appearing on the reaching video (figure 5C), would start running when the motion tracker would begin collecting data. Thus, for each trial, the kinematics from the motion tracker and their corresponding reaching videos could be aligned to one another by synchronizing the first frame of the counter onsets on the video with the beginning of the corresponding kinematic file (see figures 5C & D). All videos and time series sources could be imported into our coding station (The Observer XT, Noldus Inc.) and synched to one another to fully reconstitute, integrate, and observe the looking, reaching, and grasping behaviors of the infants as they occurred and succeeded in relation to one another on each trial (figures 5B to D provide a frame output from the Observer XT containing all video views and movement kinematics of one trial after they have been synchronized to one another).

This setup, which for now we have used to collect data in 9-month-old infants, allowed us to achieve our research goals. First, our attrition rate was reduced to 59%; out of 37 infants brought to the lab, we were able to obtain useable eye-tracking data from 15 of them. Attrition ranged from fussiness, to poor calibration, or lack of sufficient eye-tracking data due to infants not paying enough attention to the objects. More importantly, this remote eye-tracking system made it possible to identify accurately where infants directed their visual attention on the objects prior to reaching, and to relate this visual information to where infants directed their reaching patterns on the object shortly after. This is illustrated in figure 6 with two examples of gaze plots from one infant that spatially matched her reaching to her looking at the object. In figure 6A, the object presented was a vertical rod with a sphere at the top similar to a drumstick, and in figure 6B, the object was a similar rod, however, without sphere and presented horizontally (note that all our objects were presented vertically

and horizontally). Both trials show that this infant spent more time looking at one end of the object, either the sphere at the top of the drumstick or the right end of the plain rod. When the toy was brought into her reaching space, she directed her hand toward the area of the objects where she looked most to grasp the toy. Data from the 15 infants who yielded useable data for this study revealed that this kind of perceptual-motor match was produced by the majority of the infants, however, the observed rate of spatial matching between looking and reaching varied greatly between infants (Corbetta, Guan, & Williams, 2010). Some infants produced rates of spatial matching between looking and reaching that were as high as 73%, and others produced spatial perception-action matches that were as low as 23%. We are currently collecting data with younger and older infants to examine whether this rate of matching between looking and reaching increases or decreases over developmental time. Also, given the wide individual differences we observed in our 9-month-old population sample, we began collecting longitudinal data on the development of looking and reaching using the same procedure described above to gain a better understanding of how such perceptual-motor mapping develops over time and determine why infants differ so much in their rate of perception-action matching. Here, we provide very preliminary results in one infant for whom we completed weekly data collection from when she was 10 weeks old up to 49 weeks old. Figure 7 displays the rate of spatial matching between where she looked the most on the object and where she touched the object first when she made contact with it from reach onset at week 16 (3:2 months old) until week 49 (11:5 months old). These data show that the rate of matching between where she looked the most on the object and where she directed her hand to reach for it was very low initially. From week 20, the rate of look-reach match began to increase steadily until week 36 (8:1 month) where this rate attained a peak value of 88%. From that point on, the matching rate between looking and reaching declined again to values neighboring 50%. We can only speculate on the meaning of these results given that we only have data for one infant, however, it is interesting to note that the rate of matching between looking and reaching displayed a sustained increase during the early developmental period when infants are still learning to control their arm and consolidating their reaching behavior (Thelen et al., 1996; von Hofsten, 1979). In contrast, after 8 months of age, a period corresponding to more stable and more flexible reaching behavior, this match between looking and reaching becomes less predominant. It could be possible that by that later period, as infants are better at modulating their movement, they also become less dependent from the direct input of vision to direct their hand, but clearly, more data on more infants will be needed to confirm this possible explanation.

The greater gaze precision we obtained with the remote eye-tracker also allowed us to analyze the distribution of the looking patterns as a function of the objects used. To take the example of the 2 objects discussed above -- the drumstick and plain rod -- infants as a group, spent significantly more time looking at the sphere portion of the drumstick than the handle portion regardless of their orientation, however, no systematic group looking trend was observed for the plain rods. In fact, looking patterns on the plain rods tended to be more spread along the length of the rod, unlike the example presented on figure 6B. Overall, it seemed that if objects had distinct parts and some parts were larger or more salient, these parts were more likely to be visually explored (Corbetta et al., 2010).

Final considerations

We have presented two methods and types of eye-tracking devices that we have used to study how infants rely on visual information to plan and execute their actions when reaching for objects. Both the methods and eye-tracking systems discussed have their advantages and disadvantages. For infant researchers interested in tying infants' visual inputs with their action, identifying which device would be the most suited to address their research questions

will depend greatly on the task and research setup available. As discussed above, despite the systems limitations, especially when using them with infant populations, they are amenable to address questions of perception and action in development. From our experience, using a head-mounted eye-tracker with infants has been the most challenging, but there is a growing interest in the infant research community to make those systems more user-friendly and more readily available to other scientists. Indeed, such eye-tracking systems open the door to the study of infant perception in more natural, less constrained environments, and hence allow researchers to obtain a better understanding of what is present in the infants' view, where they look, and how they learn from their interaction with the world (Smith, Yu, & Pereira, 2011; Yu, Smith, Shen et al., 2009). For researchers wanting to work in more controlled environments, the use of stand-alone remote eye-trackers may offer the best flexibility. As we mentioned earlier, in our laboratory we have been able to use such eye-tracking devices to investigate infant looking patterns at 2D scenes located as far as 2–3 meters away from the infants (Guan & Corbetta, 2010), we have been able to use it in the context of reaching for 3D objects when the object was presented slightly out of reach (Corbetta et al., 2010; Williams et al., 2010), it can be used with objects presented within infants reach, and, such systems can also be used in the way most infant researchers prefer to utilize it, that is, with a computer screen atop of the eye-tracker to display still or animated scenes on the computer screen.

We have described how we synchronized different sources of data collection into one common time frame of reference for the purpose of relating information between videos recordings, and between kinematics and videos. The low cost hardware solution we used to do that can easily be employed in a wide range of data collection settings to synchronize various sources of cameras and inputs. Also, our use of the Observer XT, from Noldus, to import the different sources of information is a great way to visualize all behavioral aspects of the task as the infants are completing it (other companies also offer coding stations providing similar capabilities), however, researchers need to keep in mind that video-based coding stations such as those are designed to work with images at a rate of 30Hz, and thus will lose the more fine grained eye-tracking information that would be collected with a higher sampling system. If researchers use higher speed eye-trackers and want to preserve all the details in the gaze, they may want to consider importing the eye-tracking time series data into another data analyses software such as MATLAB, or performing some eye coding using the software provided by the eye-tracking manufacturing company, if available.

Finally, the most important lesson we have learned from using eye-tracking with infants in the context of goal-oriented actions is that everything matters. Hundreds of studies, including studies from our laboratory, have studied infant reaching using all kinds of colorful and attractive toys to keep infants interested in the task and entice them to reach. In our pilot studies with eye-tracking, we quickly realized that the patterned details, variations in texture, contrasts between colors on the objects, and the shapes of the objects could all drastically alter infants looking patterns at the objects, and ultimately affect their reaching patterns. For example, in one pilot study we presented varied spherical objects to the infants. Some were painted with one solid color; others had diamond shapes painted all over their surface. We observed that infants presented with the uniformly solid painted objects were more likely to look at the contours of the objects where the light contrast with the background appeared, while the infants presented with the diamond decorated spheres spent more time scrutinizing the diamonds on the spheres. This was an important detail to know as we were designing the objects for our reaching study because we wanted to ensure that infants would direct their attention mostly to the contours of the objects in order to assess how the shape and orientation of objects would affect the looking to reaching response. In the diamond decorated spheres, we could never infer with certitude from the infants' looking patterns whether infants encoded the overall shape of the objects when looking at the

diamond shapes or not. Clearly, object shape matters as it dictates not only object-directed visual exploration but also the decision-making process of where to grasp the object prior to reaching for it.

This contribution is far from covering every possible context in which eye-tracking may be used in the context of action, nonetheless, we hope to have at least provided sufficient information to help researchers make an informed decision as to which type of device to use if engaging in similar kinds of studies. It is our hope that infant researchers will learn from our initial attempts to either use, further extend, or develop new methods to study infant eye-tracking in the context of actions.

Acknowledgments

We thank Damian Fricker, Chen Yu, and Linda Smith from Indiana University for providing information about the Positive Science head-mounted eye-tracker. Portion of the research reported in this paper was supported by NICHD grant R03 HD043236 to D.C.

References

- Aslin, RN. Development of smooth pursuit in human infants. In: Fisher, DF.; Monty, RA.; Senders, JW., editors. *Eye movements: Cognition and visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 1981. p. 31-51.
- Aslin RN, McMurray B. Automated corneal-reflection eye tracking in infancy: methodological developments and applications to cognition. *Infancy*. 2004; 6:155–163.
- Bertenthal BI, Longo MR, Kenny S. Phenomenal Permanence and the Development of Predictive Tracking in Infancy. *Child Development*. 2007; 78:350–363. [PubMed: 17328710]
- Berthier NE, Carrico RL. Visual information and object size in infant reaching. *Infant Behavior and Development*. 2010; 33(4):555–566. [PubMed: 20708270]
- Bojczyk KE, Corbetta D. Object retrieval in the first year of life: Learning effects of task exposure and box transparency. *Developmental Psychology*. 2004; 40:54–66. [PubMed: 14700464]
- Bushnell E. The decline of visually-guided reaching during infancy. *Infant Behavior and Development*. 1985; 8:139–155.
- Carrico RL, Berthier NE. Vision and precision reaching in 15-month-old infants. *Infant Behavior & Development*. 2008; 31:62–70. [PubMed: 17706290]
- Clifton RK, Muir DW, Ashmead DH, Clarkson MG. Is visually guided reaching in early infancy a myth? *Child Development*. 1993; 64(4):1099–1110. [PubMed: 8404258]
- Clifton RK, Rochat P, Litovsky RY, Perris EE. Object representation guides infants' reaching in the dark. *Journal of Experimental Psychology: Human Perception and Performance*. 1991; 17(2):323–329. [PubMed: 1830078]
- Clifton RK, Rochat P, Robin DJ, Berthier NE. Multimodal perception in the control of infant reaching. *Journal of Experimental Psychology: Human Perception and Performance*. 1994; 20:876–886. [PubMed: 8083641]
- Corbetta, D.; Guan, Y.; Williams, JL. Do 9-months-old infants reach where they look?. In: Corbetta, D., editor. *New insights into motor development from eye-tracking studies*; Symposium paper presented at the 2010 Conference of the North American Society for the Psychology of Sports and Physical Activity; Tucson, Arizona. 2010 Jun. (Chair)
- Corbetta D, Snapp-Childs W. Seeing and touching: the role of sensory-motor experience on the development of infant reaching. *Infant Behavior and Development*. 2009; 32:44–58. [PubMed: 19081142]
- Corbetta, D.; Williams, JL. The impact of object scanning on the execution of reaching movement infants. 2011. Manuscript in preparation
- Cowie D, Atkinson J, Braddick O. Development of visual control in stepping down. *Experimental Brain Research*. 2010; 202:181–188.

- Guan, Y.; Corbetta, D. Eight months olds sensitivity to object size and depth cues in 2-D displays; Poster presented at the 17th International Conference on Infant Studies; Baltimore, Maryland. 2010 Mar.
- Falck-Ytter T, Gredebäck G, von Hofsten C. Infants predict other people's action goals. *Nature Neuroscience*. 2006; 9:878–879.
- Farzin F, Rivera SM, Whitney D. Spatial resolution of conscious visual perception in infants. *Psychological Science*. 2010; 21:1502–1509. [PubMed: 20817914]
- Farzin F, Rivera SM, Whitney D. Time crawls: The temporal resolution of infants' visual attention. *Psychological Science*. 2011; 22:1004–1010. [PubMed: 21712505]
- Flanagan JR, Johansson RS. Action plans used in action observation. *Nature*. 2003; 424:769–771. [PubMed: 12917683]
- Franchak JM, Adolph KE. Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. *Vision Research*. 2011; 50:2766–2774. [PubMed: 20932993]
- Gredebäck G, von Hofsten C. Infants' evolving representations of object motion during occlusion: A longitudinal study of 6- to 12-month-old infants. *Infancy*. 2004; 6:165–184.
- Gredebäck G, Johnson S, von Hofsten C. Eye tracking in infancy research. *Developmental Neuropsychology*. 2010; 35:1–19. [PubMed: 20390589]
- Hayhoe M, Ballard D. Eye movements in natural behavior. *Trends in Cognitive Science*. 2005; 9:188–194.
- Horstmann A, Hoffmann KP. Target selection in eye-hand coordination: do we reach to where we look or do we look to where we reach? *Experimental Brain Research*. 2005; 167(2):187–195.
- Johnson SP, Amso D, Slemmer JA. Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100:10568–10573. [PubMed: 12939406]
- Johnson SP, Slemmer JA, Amso D. Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*. 2004; 6:185–201.
- Jonikaitis D, Deubel H. Independent allocation of attention to eye and hand targets in coordinated eye–hand movements. *Psychological Science*. 2011; 22:339–347. [PubMed: 21270450]
- Land M, Mennie N, Rusted J. The roles of vision and eye movements in the control of activities of daily living. *Perception*. 1999; 28:1311–1328. [PubMed: 10755142]
- McCarty ME, Ashmead DH. Visual control of reaching and grasping in infants. *Developmental Psychology*. 1999; 35:620–631. [PubMed: 10380854]
- Oakes L. Infancy guidelines for publishing eye-tracking data. *Infancy*. 2010; 15:1–5.
- Piaget, J. The origins of intelligence in children. New York: International Universities Press; 1952.
- Quinn PC, Doran MM, Reiss JE, Hoffman JE. Time course of visual attention in infant categorization of cats versus dogs: Evidence for a head bias as revealed through eye tracking. *Child Development*. 2009; 80:151–161. [PubMed: 19236398]
- Smith LB, Yu C, Pereira AF. Not your mother's view: the dynamics of toddler visual experience. *Developmental Science*. 2011; 14:9–17. [PubMed: 21159083]
- Thelen E, Corbetta D, Kamm K, Spencer JP, Schneider K, Zernicke RF. The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development*. 1993; 64:1058–1098. [PubMed: 8404257]
- Thelen E, Corbetta D, Spencer JP. The development of reaching during the first year: The role of movement speed. *Journal of Experimental Psychology: Human Perception and Performance*. 1996; 22:1059–1076. [PubMed: 8865616]
- Thelen, E.; Smith, LB. Dynamic systems theories. In: Lerner, RM., editor. *Handbook of child psychology, Vol. 1: Theoretical models of human development*. 6. New York, NY: John Wiley & Sons, Inc; 2006. p. 258–312.
- von Hofsten C. Development of visually directed reaching: The approach phase. *Journal of Human Movement Studies*. 1979; 5:160–178.
- von Hofsten C, Rosander K. Development of smooth pursuit tracking in young infants. *Vision Research*. 1997; 37(13):1799–1810. [PubMed: 9274766]

- von Hofsten C, Vishton PM, Spelke ES, Feng Q, Rosander K. Predictive action in infancy: Tracking and reaching for moving objects. *Cognition*. 1998; 67:255–85. [PubMed: 9775511]
- White BL, Castle P, Held R. Observations on the development of visually-guided reaching. *Child Development*. 1964; 35:349–364. [PubMed: 14163542]
- Williams, JL.; Guan, Y.; Corbetta, D.; Valero Garcia, AV. “Tracking” the relationship between vision and reaching in 9-month-old infants. In: Vasconcelos, O.; Botelho, M.; Corredeira, R.; Barreiros, J.; Rodrigues, P., editors. *Estudos em Desenvolvimento Motor da Criança III* [Studies in motor development in children III]. Porto: University of Porto Press; 2010. p. 15-26.
- Yu C, Smith LB, Shen H, Pereira AF, Smith TG. Active information selection: Visual attention through the hands. *IEEE Transactions on Autonomous Mental Development*. 2009; 1:141–151. [PubMed: 21031153]
- Yu, C.; Smith, LB.; Fricker, D.; Xu, L.; Favata, A. What Joint Attention is made of -- A Dual Eye Tracking Study of Child-Parent Interaction. Paper presented at the biennial meeting of the Society for Research in Child Development; Montreal, Canada. 2011 Mar.

Figure 1.

(a) Infant head-mounted eye-tracker used in our reaching study; (b) and (c) still frames of the video output provided by this eye-tracker on two different trials from the same infant, with the crosshair indicating where the infant is directing his gaze at this particular frame on the scene (Corbetta & Williams, 2011).

Figure 2. Illustration of synchronized time series from the eye-tracker (top graph) and motion tracker (3 bottom graphs: displacement, velocity, and rotation of the reaching arm). This example shows a trial in which the infants scanned the object horizontally prior to reaching.

Figure 3.

Illustration of synchronized time series from the eye-tracker (top graph) and motion tracker (3 bottom graphs: displacement, velocity, and rotation of the reaching arm). This example shows a trial in which the infants did not scan the object horizontally prior to reaching.

Figure 4. Illustration of the setup used with the remote eye-tracker, (a) A computer screen is fitted in the board opening above the eye-tracker for calibration, (b) the computer screen has been removed and is replaced by two layers of black curtains for object presentation.

Figure 5.

(a) Picture of a video camera fitted with the diode attached to and facing the lens used for synchronization; (b, c, d) Information seen on the coding station once all recording sources have been imported. (b) View from the eye-tracker scene camera with eye tracking data, (c) side views of the infant from the two reaching cameras, (d) time series from the motion tracker.

Figure 6.

Example of gaze plots from a 9 months old infant illustrating where the child directed her visual attention on the object prior to reaching for it. In A, the infant reached to the top and in B, the infants reached to the right, the areas explored visually the most (from Corbetta et al., 2010).

Figure 7.

Rate of matching between where the infant looked on the object prior to reaching and where she directed her arm for reaching. These data are from one infant followed longitudinally from week 16 (reach onset) to week 49.