

THEMATIC COLLECTION: INTRODUCTION

Automated Corneal-Reflection Eye Tracking in Infancy: Methodological Developments and Applications to Cognition

Richard N. Aslin and Bob McMurray
*Department of Brain and Cognitive Sciences
University of Rochester*

Since the mid-1800s, experimental psychologists have been using eye movements and gaze direction to make inferences about perception and cognition in adults (Müller, 1826, cited in Boring, 1942). In the past 175 years, these oculomotor measures have been refined (see Kowler, 1990) and used to address similar questions in infants (see Aslin, 1985, 1987; Bronson, 1982; Haith, 1980; Maurer, 1975). The general rationale for relying on these visual behaviors is that where one is looking is closely tied to what one is seeing. This is not to deny the fact that we can detect visual stimuli in the peripheral visual field, but rather that there is a bias to attend to and process information primarily when it is located in the central portion of the retina. Thus, although the direction of gaze is not perfectly correlated with the uptake of visual information (e.g., as in a blank stare or a covert shift of attention), there is a strong presumption that the direction of gaze can provide important information about visual stimuli even in newborn infants (Haith, 1966; Salapatek, 1968; Salapatek & Kessen, 1966).

HISTORY OF EYE MOVEMENT TECHNIQUES

Techniques for measuring eye movements and the direction of gaze have evolved over the past century, although the most rudimentary—simple observation of the

eyes—remains the most common technique in use today. The primary limitation of such observations is that even trained online coders can only judge reliably the direction of an infant's gaze when it falls within one of three or four regions of a stimulus field, and determining vertical gaze position is much more difficult than horizontal gaze position. Such global observations are the mainstay of single-stimulus habituation paradigms and two-stimulus preferential looking paradigms. In addition, although observers can judge the presence or absence of particular types of eye movements (e.g., optokinetic nystagmus, saccades, smooth pursuit; see Aslin & Salapatek, 1975; McGinnis, 1930; Teller & Lindsey, 1993), the quantitative characteristics of these eye movements can only be specified with a more precise measurement technique. Although offline coding of films or videotapes of the eyes allows for more detailed temporal resolution through the use of single frames and slow-motion playback, it does not improve the spatial resolution of coding or the ability to make a quantitative assessment of different types of eye movements. Moreover, although reliability ratings of trained coders are often quite high, the difficulty of determining eye position from global video recordings renders the coding and analysis of these data extremely time consuming.

Two measurement techniques developed for use in studies of adult eye movements and direction of gaze were subsequently applied to infants in the early 1960s: electro-oculography (EOG) and corneal-reflection photography (see Maurer, 1975). EOG is based on the gradient of metabolic activity between the front and back of the eyeball. When passive electrodes are placed on the face of an infant, flanking one (or both) eyeballs, rotation of the eye with respect to these fixed electrodes creates a change in electrical potential that is approximately linear with the change in direction of gaze (up to $\pm 30^\circ$).

The fundamental problems with EOG (see Finocchio, Preston, & Fuchs, 1990) are that (a) a pair of electrodes is needed for each axis of rotation (two pairs for each eye often cannot be attached to the infant), and (b) metabolic activity varies over time (causing the EOG signal to drift in the absence of any change in gaze). The first problem has typically been resolved by limiting the question of interest to horizontal eye movements (especially because the placement of electrodes to assess vertical eye movements often leads to artifacts from blinks). This single-axis (horizontal) solution, however, does not enable EOG to be used in studies of eye movements or gaze shifts in two dimensions.

The second problem (drift) can be overcome in adults by conducting repeated calibrations of the EOG signal. This solution is often impractical in infants because of the limited time during which they maintain interest in visual displays without becoming fussy (especially with electrodes attached to their face) and because of the difficulties in calibrating prelinguistic participants. Moreover, because the EOG signal is relative to the position of the electrodes (which are fixed to the head), it raises another problem: Any change in head position with respect to the stimulus field will alter the interpretation of the EOG signal. For example,

if a participant's gaze is fixed on a stationary target and the head is moved, the EOG signal will change despite the fact that gaze remains on the target. Thus, either the head must be stabilized or some other device must be used to measure the position of the head, thereby compensating for the gaze error induced by the movement of the head.

Given the limitations of EOG, the second technique that emerged in the early 1960s was corneal-reflection photography (see Haith, 1969, 1980). The basis for this technique is remarkably simple. If a single eye is directed to fixate a small point of light, then that light creates a reflection off the front surface (cornea) of the eyeball. For a camera located just behind this light source and directed at the eyeball of the fixating participant, the corneal reflection will appear to be located in the center of the pupil. As the participant moves fixation to the right, left, up, or down with respect to the light, the corneal reflection will be displaced relative to the center of the pupil. That is, there is a lawful (monotonic) relation between the relative position of the corneal reflection with respect to the center of the pupil and the direction of gaze, and this relation holds for both horizontal and vertical shifts of gaze.

Initially, the corneal-reflection technique used more than one light to create several corneal reflections (Salapatek & Kessen, 1966), thereby reducing the parallax error (i.e., nonlinearity) that occurs as the eyeball rotates away from the direction of the camera. Moreover, because these lights were salient to the infants and may have distracted them from attending to the experimental stimuli of interest, infrared filters were placed over the lights to render them invisible to the infant, and infrared-sensitive film was used in the movie camera so that the corneal reflections were visible on each film frame (typically taken at 4/sec). As a result, a series of film frames provided images of the relation between the lights reflecting off the cornea and the center of the pupil. Because the location of the lights with respect to any visual stimulus being presented to the infant could be measured precisely, it was possible by simple geometric reconstruction to obtain a detailed measure of changes in gaze with a temporal resolution of 4 Hz.

With the development of infrared-sensitive video cameras in the late 1960s, it was rather simple (in principle) to move from film to videotape, thereby improving temporal resolution from 4 to 30 Hz (Haith, 1969). However, there were intrinsic problems with the corneal-reflection technique that video technology could not overcome. First, the inference that the center of the pupil corresponded with the axis of gaze was incorrect because the fovea (the visual axis) is displaced slightly with respect to a line (the optic axis) that passes through the center of the pupil (Salapatek, Haith, Maurer, & Kessen, 1972). Thus, as with EOG, the output of the eye-position measurement technique (EOG signal or distance between a corneal reflection and the center of the pupil) has no objective status, but rather must be mapped onto the stimulus field by having the infant fixate several known (and preferably small) calibration targets (see discussion in the following section). Average

estimates of this error suggest that it is much larger than in adults and that there is considerable variation among infants (Maurer, 1975). As a result of this variation, calibration is critical.

Another problem with the corneal-reflection technique was its reliance on judging the center of the pupil. Under normal illumination conditions, the pupil is black (i.e., it is the absence of light created by a hole in the iris). For participants with brown or dark-blue eye color, the border between the pupil and the iris is difficult to define from the film (or video) frame, thereby adding error to the measurement. One solution, which every flash photographer has experienced, was to place the light that creates the corneal reflection close to the axis of the lens of the camera. This creates a *red reflex*, which is light being reflected back from the surface of the retina and filling the pupil. The so-called bright pupil technique uses this phenomenon to create images of the eyeball that consist of a dark iris, a light pupil (created by this reflected light), and an even brighter corneal reflection (because the cornea is a much better reflective surface than the retina).

As the corneal-reflection technique migrated to video and more powerful digital circuitry became available in the early 1970s, it became possible to perform an electronic analysis of each video frame, extracting the coordinates of the center of the pupil and the coordinates of the corneal reflection and compensating for parallax errors. This analysis computed the direction of gaze with respect to the location in space of the light source that creates the corneal reflection. Video circuitry by this time was fast enough to perform these computations in real time at a rate of 60 Hz (e.g., Aslin, 1981). More recent developments have increased this rate to 120 Hz and even 250 Hz.

Although these automated corneal-reflection systems had a temporal resolution that made it possible to extract quantitative parameters of eye movements as well as gaze direction, they continued to suffer from the most serious limitation of the corneal-reflection technique. To optimize the spatial resolution of the system, the video camera had a close-up lens that filled the video frame with the pupil. As a result of this very small video frame, any movement of the infant resulted in the pupil moving outside the video frame, thereby losing the opportunity to capture any usable data.

In the past several years, automated corneal-reflection systems have solved this out-of-frame problem using three innovations. First, the video camera is mounted on a motor-driven base that can quickly move the field of view of the camera to compensate for small (and slow) head movements. The second innovation was the development of relatively simple algorithms to drive these motors using information from the video image of the eye to maintain the pupil in the center of the camera's field of view. Although these innovations proved superior to a fixed camera, they did not compensate for rapid head movements. (Note that head movements that place the eye completely outside the possible field of view of the camera are, of course, not relevant here because no single-camera system can deal effectively with such large head movements, and such head movements are the infant's attempt to look at nonexperimental stimuli such as the mother.)

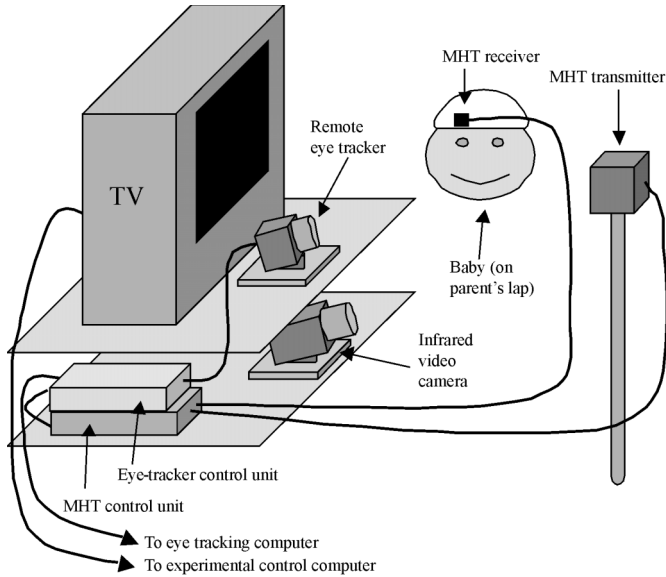


FIGURE 1 Components of the automated eye- and head-tracking system used in our lab.

The final innovation was the use of a position-sensing system that monitors head movements and feeds that information to the motor-driven video camera, thereby maintaining the image of the eye in the field of view of the video camera (except when the eye is completely out of range of the camera). Moreover, unlike optical methods, when the infant turns away from the camera, the position-sensing system can reacquire the image of the eye automatically as the eye returns to the possible range of the camera's field of view. The position-sensing system consists of a magnetic field within which the infant is placed and a small sensor worn on the infant's head. This system sends information to circuitry that (a) knows the relation between the sensor's location on the infant's head and the location of the infant's eye, and (b) feeds that head position signal to the motor-driven base of the video camera in real time. As a result, even for an infant who makes substantial head movements, the automated corneal-reflection system with head tracking can maintain a steady stream of eye position data from most infants (see Figure 1 for a schematic diagram of the head- and eye-tracking equipment used in our lab).

CALIBRATION OF EYE-TRACKING SYSTEMS

The beauty of the original corneal-reflection photography technique was that it promised to deliver an objective method for measuring gaze position. Unfortunately, it could not deliver on this promise because of the uncertain relation between the center of the pupil and the direction of gaze. Thus, every eye-tracking

system must rely on calibration data to map the signals provided by the system (e.g., EOG voltages or video frame coordinates) onto the stimulus field. Because every system suffers from nonlinearities (e.g., a compression of signals as gaze position deviates from the axis of the camera), the calibration data provide a means of linearizing the system, thereby providing a metric for any gaze position within the stimulus field.

In adults, calibration is a relatively simple matter of asking each participant to fixate (and hold) on several small targets placed at known locations in the stimulus field. However, the problem of calibrating infants is twofold. First, they are less likely than adults to fixate (and hold) on small visual targets because such targets are relatively uninteresting and infants cannot be instructed to perform this demanding calibration task. Second, the more time that is expended on calibration, the less likely the infant will remain in a cooperative state for providing the real data that are the goal of the eye-tracking experiment.

Given these problems, compromises have been made in studies of infants. First, the usual 9-point calibration schemes used with adults are typically reduced to 5- or 2-point schemes. Second, the calibration target that elicits fixation begins as a large, moving (or “looming”) stimulus that attracts the infant’s gaze and then shrinks and becomes stationary at each calibration location in the stimulus field. Most infants will exhibit interest in the large calibration target, follow its movement, and maintain fixation for 1 to 2 sec as it shrinks and remains stationary. Thus, a 2-point calibration routine can be completed in less than 1 min so that the experiment can begin promptly, and this 2-point scheme can be repeated during the experiment to ensure that calibration has been maintained.

DATA ANALYSIS

The advantages of an automated corneal-reflection system with head tracking are obvious: (a) access to both horizontal and vertical eye position, (b) a relatively high sample rate of 60 Hz, (c) rapid calibration, (d) minimal loss of data from off-camera head movements, and (e) minimal coding time. Of course, this last advantage is also a potential problem because of the voluminous stream of data. Within this data stream there are undoubtedly some segments where the algorithms that compute eye position have made errors. How can these errors be detected and eliminated?

One strategy is the use of a time-stamped video record of the experimental session. Most commercial eye trackers output two sources of video data (in addition to the digital data stream). The “scene output” is typically a video record of the visual scene in front of the infant (either from a fixed video camera, or from the television or monitor the infant is viewing). The “eye output” shows the raw video data used by the eye tracker to determine pupil and corneal reflection (often with these features marked in some way). In addition, our lab (and others)

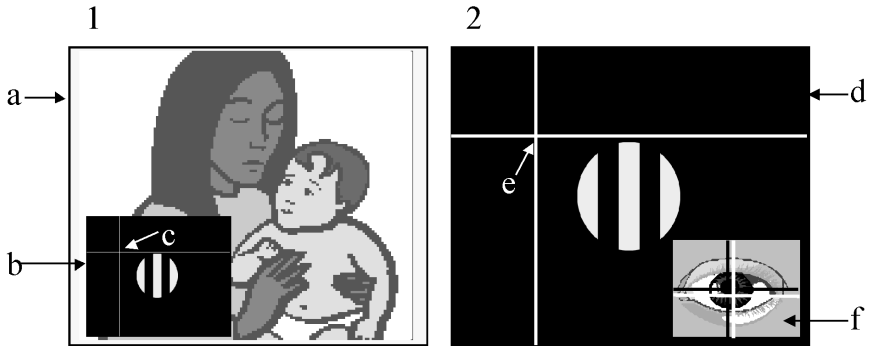


FIGURE 2 Panel 1: A screen used for coding infants' eye movements from the video record that includes (a) output from a wide-angle, fixed video camera, and (b) the output of the eye tracker's scene camera, which shows the image the infant is watching with superimposed crosshairs (c) indicating gaze position. Panel 2: Alternatively, data can be coded from a screen showing the scene camera (d) and gaze position (e) with the eye tracker's eye camera (f), which shows a close-up of the infant's eye with the pupil and corneal reflection marked by crosshairs.

employs a second video camera (wide angle) that records the entire experimental session. These three video sources can be combined using picture-in-picture technology onto a single tape, or they can be recorded separately on synchronized videocassette or digital recorders (see Figure 2 for two examples). In this way, we always have a record of what the infant was doing during the experiment when the data stream shows anomalies. For some of these segments of output from the eye tracker, the detailed settings (e.g., the threshold for what video features define the pupil) may have been suboptimal even though the global video image is perfectly scorable. In these cases, the backup video enables an offline scorer to code the direction of the infant's gaze (with admittedly less accuracy). Alternatively, the output of the eye camera can be used to determine whether track loss was the result of a blink, the loss of the eye (from the field of view of the camera), or suboptimal settings. Moreover, as the number of regions of interest in the stimulus display grows, holistic coding becomes less reliable, whereas the close-up eye image may still remain useful.

PREVIEW OF THEMATIC COLLECTION ARTICLES

The relative ease and power of today's automatic corneal reflection eye trackers has enabled researchers to use eye movements as a powerful measure of infant cognitive and perceptual abilities. In each of the four articles that comprise this thematic collection, an automated corneal-reflection system was used to track infant

eye movements and direction of gaze to either stationary or moving visual targets with implications for a diverse set of research issues.

Gredebäck and von Hofsten (2004/*this issue*) used predictive changes in gaze to assess 6- to 12-month-old infants' ability to represent the movement of an object behind an occluder. Johnson, Slemmer, and Amso (2004/*this issue*) recorded detailed fixations while 3-month-olds viewed rod-and-box displays in an object unity task. McMurray and Aslin (2004/*this issue*) used a two-choice spatial localization task to study the formation of categories and the generalization to novel stimuli, with anticipatory eye movements as the dependent measure of categorization. Finally, Hunnius and Geuze (2004/*this issue*) studied the changes in scanning of faces and nonface control stimuli between 1 and 6 months of age. In each of these studies of infant perception and cognition, eye movements and gaze direction were used as measures of processing. Without an automated corneal-reflection system, the questions under investigation in these studies could not have been addressed. Commentaries on these four target articles are provided by Haith (2004/*this issue*), one of the pioneers in the use of corneal reflection eye trackers with young infants, and by Hayhoe (2004/*this issue*), who employs head-mounted eye tracking and virtual reality displays to study adult perception.

ACKNOWLEDGMENT

Bob McMurray is now at the Department of Psychology, University of Iowa.

REFERENCES

- Aslin, R. N. (1981). Development of smooth pursuit in human infants. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 31–51). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Aslin, R. N. (1985). Oculomotor measures of visual development. In G. Gottlieb & N. Krasnegor (Eds.), *Measurement of audition and vision during the first year of postnatal life: A methodological overview* (pp. 391–417). Norwood, NJ: Ablex.
- Aslin, R. N. (1987). Motor aspects of visual development in infancy. In P. Salapatek & L. B. Cohen (Eds.), *Handbook of infant perception* (pp. 43–113). New York: Academic.
- Aslin, R. N., & Salapatek, P. (1975). Saccadic localization of visual targets by the very young human infant. *Perception & Psychophysics*, *17*, 293–302.
- Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century.
- Bronson, G. W. (1982). The scanning patterns of human infants: Implications for visual learning. In L. P. Lipsitt (Ed.), *Monographs on infancy* (pp. 1–102). Norwood, NJ: Ablex.
- Finocchio, D. V., Preston, K. L., & Fuchs, A. F. (1990). Obtaining a quantitative measure of eye movements in human infants: A method of calibrating the electrooculogram. *Vision Research*, *30*, 1119–1128.

- Gredebäck, G., & von Hofsten, C. (2004). Infants' evolving representations of object motion during occlusion: A longitudinal study of 6- to 12-month-old infants. *Infancy*, 6, 165–184.
- Haith, M. M. (1966). The response of the human newborn to visual movement. *Journal of Experimental Child Psychology*, 3, 235–243.
- Haith, M. M. (1969). Infrared television recording and measurement of ocular behavior in the human infant. *American Psychologist*, 24, 279–283.
- Haith, M. M. (1980). *Rules that babies look by: The organization of newborn visual activity*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Haith, M. M. (2004). Progress and standardization in eye movement work with human infants. *Infancy*, 6, 257–265.
- Hayhoe, M. M. (2004). Advances in relating eye movements and cognition. *Infancy*, 6, 267–274.
- Hunnis, S., & Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy*, 6, 231–255.
- Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, 6, 185–201.
- Kowler, E. (Ed.). (1990). *Eye movements and their role in visual and cognitive processes: Reviews of oculomotor research* (Vol. 4). Amsterdam: Elsevier.
- Maurer, D. (1975). Infant visual perception: Methods of study. In L. B. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition: Vol. 1. Basic visual processes* (pp. 1–76). New York: Academic.
- McGinnis, J. M. (1930). Eye movements and optic nystagmus in early infancy. *Genetic Psychology Monographs*, 8, 321–430.
- McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, 6, 203–229.
- Salapatek, P. (1968). Visual scanning of geometric figures by the human newborn. *Journal of Comparative and Physiological Psychology*, 66, 247–258.
- Salapatek, P., Haith, M. M., Maurer, D., & Kessen, W. (1972). Error in the corneal-reflection technique: A note on Slater and Findlay. *Journal of Experimental Child Psychology*, 14, 493–497.
- Salapatek, P., & Kessen, W. (1966). Visual scanning of triangles by the human newborn. *Journal of Experimental Child Psychology*, 3, 155–167.
- Teller, D. Y., & Lindsey, D. T. (1993). Motion nulling techniques and infant color vision. In C. E. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 47–73). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.