Eye tracking

Tim J. Smith* & Irati R. Saez De Urabain

E-mail: tj.smith@bbk.ac.uk
iurabain@gmail.com

Web site: http://www.bbk.ac.uk/psychology/our-staff/academic/tim-smith/tim-smith

Word count (2300-3125): 3030 words

*Corresponding authors contact:
Department of Psychological Sciences
Birkbeck, University of London
Malet Street
LONDON WC1E 7HX

General comments:

* Please ensure consistent use of APA style throughout text, Further reading and reference list.

* Please check that references cited in the text are also included in the reference list.

* Terms in bold in the text are for inclusion in the glossary. You may have others you want to include. The entries for the glossary can be as long as you want, and include figures and tables if necessary.

* Make sure publications cited under Further reading are not also cited in the reference list.

* Please ensure that section and sub-section headings are not indented. The same applies to the first paragraph in each (sub-) section.

* In the text, you will find a few comments/queries. Please check them out and see if you can account for them.

Introduction

Tracking is the measurement of a viewer's eye movements relative to a visual array, whether the array is a real-world scene, a tabletop, or stimuli presented on a computer screen. Methods for recording eye movements have been around for over 100 years (Wade & Tatler, 2005), but up until the last twenty years these techniques have been highly invasive and uncomfortable for the user, wearer? making them impractical for use in developmental research. Researchers who were interested in monitoring where infants looked had to resort to hand coding of video footage or using electrooculography (EOG), the electrical signals produced as the eyeball rotates detected by electrodes placed on the face. Both techniques were spatially (in terms of *where* gaze was directed) and temporally (*when* gaze shifts occurred) imprecise≤ and could only reliably be used with horizontal eye-movements.

Fortunately, recent advances in computer and imaging technologies has meant that it is now possible to track infant eye movements using a system of remote high-speed video cameras and infrared (IR) illumination. These video-based combined pupil/corneal reflection tracking systems exploit the fact that infra-red light shone on the human eye behaves in two distinct ways: 1. if the IR illuminator is offset from the optical axis of the camera, the light will enter the pupil and not be reflected back, creating a dark pupil, and simultaneously 2. some IR light will be reflected back off the outside of the eye (the cornea), creating a glint known as corneal reflection. As the eye rotates, the pupil moves with the eye, but the glint always remains in roughly the same position relative to the IR light source due to the fact that the eye ball is assumed to be a perfect sphere. By identifying the displacement of the pupil center relative to the glint, we are able to identify the precise vector of the eye's movement in two dimensions. These vectors can then be calibrated relative to a two-dimensional (2-D) plane, such as a computer screen, by asking the participant to look at a series of points on the screen (typically 3, 5 or 9). The computer uses these points to build a model of the eye's movements and to infer where the viewer's eyes are pointing.

This entry will now discuss various motivations, issues and considerations that a researcher should be aware of when designing an eye tracking study with infants and young children.

Why record eye movements?

We may experience the visual world as continuous and coherent, but this percept is constructed from a series of discrete periods of visual sampling when our eyes are relatively static (known as fixations) separated by rapid eye-movements to new locations in the scene (known as saccades) during which vision is suppressed. These fixations are necessary to direct the focus of interest to the fovea, the small region of high acuity in the eye and the location at which visual activity is centered. Visual acuity of stimuli degrades greatly as the distance from fovea increases. If the object of interest is stationary, our eyes will stabilize the reflected image of the object on the fovea to maximize visual acuity. If the object is moving relative to our viewpoint, we will pursue the object by rotating our eyes so the image remains as close to the fovea as possible (known as a smooth pursuit eye-movement). In adults, each fixation lasts on average 330ms (when focused on a static visual scene; Rayner, 1998), and varies in duration with the complexity of visual stimuli and viewing task (Henderson, 2003).

Infant eye-tracking has shown that fixation durations are typically longer in infants than adults (590ms for 10-month-olds viewing static images), and also vary with stimulus features (Wass & Smith, 2014). The location and duration of each fixation is controlled by a mixture of endogenous factors such as the relevance of fixated information to the viewing task, individual preference, memory, or monitoring of current visual processing, as well as exogenous factors, such as the visual salience of a target location. In newborns and young infants, eye-movement control is mainly driven by exogenous factors, and it is not until 3 to 4 months that endogenous factors start playing a more important role in gaze allocation.

Oculomotor measures can also be used to investigate individual differences in development. Peak look duration (duration of the longest unbroken look to a particular area of the screen, which may involve several fixations and saccades within this area) has been shown to be a reliable marker of attentional control during infancy (Colombo & Mitchell, 1990). Likewise, individual differences in mean fixation duration have been shown to be highly stable overtime in adults (Castelhano & Henderson, 2008) and in infants (Wass & Smith, 2014) and may be indicative of later cognitive and behavioral differences. For instance, Papageorgiou and colleagues (2014) found that individual differences in mean fixation duration during the first year of life positively predicted parent-report measures of effortful control and negatively predicted hyperactivity and inattention three years later. Such oculomotor markers of cognitive development may prove highly useful in predicting future developmental atypicalities such as Attention Deficit and Hyperactivity Disorder (ADHD) or Autism Spectrum Disorder (ASD).

Eye-trackers also provide developmental researchers with a great tool for administering reaction time studies (such as the popular **gap-overlap task** in which saccades are triggered to peripheral targets under various viewing conditions (Johnson, Posner, & Rothbart, 1991)) at ages before fine-motor skill has fully matured. It also allows experimenters to modify stimulus presentation time and location depending on the participant's gaze. At the time of writing this entry such reactive experiments require programming in a compatible presentation platform such as Matlab, Python, E-Prime or SR Research's Experiment Builder, making them difficult to implement for some researchers.

Using such techniques, developmental eye-tracking studies have provided new insights into various aspects of development including but not limited to the emergence of cognitive, social and emotional processing in infancy, language development, perceptual learning, memory, and face processing.

Hardware

Currently, there are a large variety of eye-tracker models from different manufacturers able to satisfy a researcher's needs. For instance, depending on the physical relationship between the eye-tracker and the user, wearer? one can classify these devices into two categories: head-mounted eye-trackers (e.g., Positive Science eye-trackers, Tobii or SMI Glasses) or remote (e.g., Tobii TX300, EyeLink 1000, SMI RED). Head-mounted eye-trackers are becoming increasingly popular in developmental research as they allow the investigation of gaze behavior during relatively unconstrained natural actions such as free play or walking via eye or head? cameras mounted either on small glasses or a cap. However, as the video footage from each participant's head-mounted camera differs, where their gaze is directed within this visual scene also differs and requires hand-coding during analysis. This greatly increases the labor involved compared to the more popular remote screen-based eye-trackers.

Remote trackers are typically fixed in location relative to a presentation screen and the participant is either free to move their head with a region in front of the eye tracker (known as the 'head-box'), or their head is fixed in place using a chin or forehead rest. In general, head-fixed remote trackers will provide better spatial and temporal accuracy than head-free eye trackers as they can assume that the model of eye movements constructed during the calibration procedure is accurate throughout recording and does not have to be updated each time the head moves. However, as infants and young children cannot be instructed to keep their head in a fixed position, developmental researchers have to use head-free trackers. This currently means that the high-precision eye-tracking available for adult researchers is not currently available to developmental researchers, and the data gathered may often be noisier and less reliable than adult data (see section Data quality and parsing).

The number of samples per second is an important property of eye-trackers that is measured in hertz (Hz). For instance, a 300 Hz eye-tracker will provide gaze data 300 times per second (3.3 ms between each sample). Low-sampling rate systems (50-120 Hz) are useful for identifying where a participant fixates, changing stimuli based on fixation location or in some saccadic reaction time studies. Higher sampling rate systems (>500Hz) are essential for investigating the dynamics of the eyes during a saccade or fixational eye movements such as microsaccades. At the opposite end of the cost scale, commercial eye-tracker gaming peripherals have recently become available at a tenth of the cost of science-grade eye trackers which, while having considerable lower spatial (>1 degree) and temporal accuracy (~60Hz), may be suitable for some developmental studies interested in roughly where a child is looking.

Another critical feature of eye trackers is the latency and temporal precision. Eye-tracker latency is the delay that happens from when the eye is recorded until the recording computer detects its signal. Even more important is the stimulus synchronization latency, which arises in the interaction between the software being used for receiving the eye tracker's signal and the one for presenting the stimuli on the experimental display. It is essential to keep these latencies as low as possible, especially for studies that require precise synchronization to external devices such as EEG, NIRS or fMRI, or when using gaze-contingencies (Holmqvist *et al.*, 2011). The temporal precision, on the other hand, refers to the standard deviation of the eye tracker's latency, or in other words, the time interval between successive samples.

Set-up and calibration

The initial setup of an infant eye tracking is critical for ensuring useable data throughout the experimental session. Positioning the infant is an important step. If the child's head is tilted too far forward or backward, the tracker may lose sight of their eyes as they saccade across the screen or fidget. Good practice is to ensure the cameras on the tracker are pointing directly at the child's head, the tilt of their head matches that of the screen surface and their eyes are level with the upper half of the screen. If using a car/high-seat, the heads of younger babies must be supported and older toddlers should be securely held in the seat to limit torso movement. Seating the child in their parent's lap can sometimes be preferable for both parent and child, but by doing so, the researcher must ensure that the parent holds the child in a secure and stable position in front of the tracker and that the eyes detected by the tracker are not the parent's (instructing the parent to close their eyes during setup is an easy way to ensure this).

Before the experiment starts, each participant needs to perform a calibration. Its main purpose is to adapt the parameters for the calculation of gaze direction to the participant's eye and the particularities of the current testing session, such as the luminance of the room or the participant's distance and angle with respect to the eye-tracker. Differences between eyeballs (size of or distance between?) can be large in infants, whose eyes are still undergoing development.

In a typical calibration procedure the participant needs to look at a number of pre-defined calibration points that appear subsequently on the experimental screen (typically 3, 5 or 9, see Figure 2). What do the words in bold mean? Do you mean see Fig. 1?) Ideally, calibration points should be small and animated, often shrinking to a point to ensure gaze samples are as closely clustered as possible (Holmqvist *et al.*, 2011). It is also good practice to include audio cues coinciding with the point onset when calibrating infants to attract the baby's attention back to the screen.

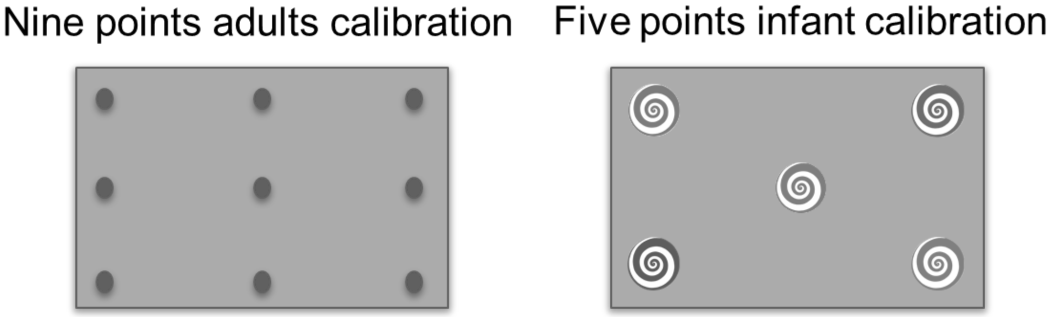

Figure 1: Left=a typical 9 points adult calibration; Right=a typical animated infant 5 point calibration.

The first step for performing a typical calibration procedure is to detect the participant's eyes with the eye-tracker. Evidently, for this to happen the participant needs to be looking at the screen, which may not always be the case when testing non-compliant populations such as infants. It is essential for researchers to find a way to keep participants looking and happy during this process, which can occasionally take longer than expected (e.g., if the participant's eyes are not easily detected by the eye-tracker). For instance, when testing infants a baby-friendly video can be presented in the background while this procedure is performed (see Figure 2).

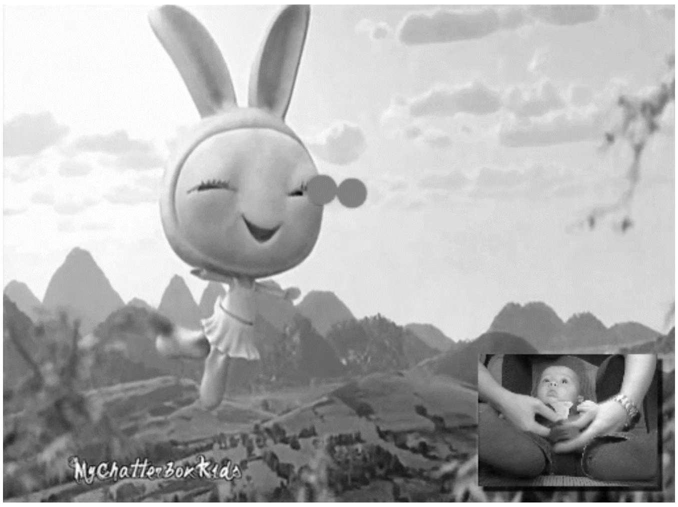

Figure 2: Before the calibration procedure starts, the participant's eyes need to be detected. When testing infants a baby-friendly video can be presented on the background while infants are placed in front of the eye tracker and their eyes are detected. The screen location of the child’s gaze is shown here as two grey dots (left and right eye). The image above shows a typical video used in several studies at the Centre for Brain and Cognitive Development, (CBCD, Birkbeck, University of London). The embedded webcam video of the child is important for the researcher to diagnose why the gaze may have disappeared from the screen (e.g. a hand occluding the eyes or looks away) or data quality has decreased (e.g. fidgeting) and allow them to pause the study and correct the problem.

Once the eyes are detected the calibration points can be presented. Many calibration procedures display a diagram at the end of the procedure showing the calibration points together with the gaze data recorded during the proceeding. While this can be informative, it may not be sufficient to evaluate the calibration results accurately. Some researchers evaluate the calibration procedure by asking the participants to look at certain points in the screen and evaluating the offset between the points and the actual gaze. On-line monitoring of the child’s live gaze during an experiment is also essential to ensure that data are being gathered and that the experimenter can pause the experiment, and attract the child’s attention back to the screen using a noisy and visually attractive attention-getter or conduct a recalibration.

Data quality and parsing

The quality of the raw data generated by an eye tracker will vary as a result of the eye-tracker model and manufacturer, the eye physiology, the calibration procedure, the position of the participant relative to the eye-tracker, the degree of head motion, or even ethnicity and iris color (Holmqvist *et al.*, 2011; Saez de Urabain, Johnson, & Smith, 2014).

Low data quality can affect both spatial and temporal accuracy of gaze measurements. Spatial accuracy or offset is defined as the difference in space between the detected gaze and the real gaze, and needs to be considered when analyzing areas of interest (AOIs; Holmqvist *et al.*, 2011). On the other hand, spatial precision refers to the consistency in detecting and calculating gaze points. Together with data loss, precision can seriously affect event detection (e.g., fixation durations).

Recording the eye movements of infants or other non-compliant populations can entail a number of issues that seriously affect data quality. For example, they do not respond to experimental instructions, they may not sit still, or their eye physiologies are highly variable (Saez de Urabain *et al*., 2014). Poor calibration is also a more pronounced problem with infants than adults as it is often unclear if infants are following the calibration points (Oakes, 2012). To minimize this problem and increase the infants' interest in the points, the researcher can regularly change the calibration stimuli (e.g., varying it in colors and/or shapes) and display calibration points together with attractive sounds. Other problems are that the calibration points are presented too fast for the infants to be able to follow them. This is particularly obvious in subjects younger than 4 months, as the neural structures implicated in oculomotor control are still underdeveloped affecting their disengagement abilities (Johnson, Posner, & Rothbart, 1991; Johnson, 2011). This is manifest through long fixation durations also known as 'sticky fixations' or 'obligatory attention', and may prevent them from moving their eyes readily from one point to the next. It is an issue that can be solved by having a live display of the infant's gaze on a second screen and allowing the researcher to decide when the next calibration point should appear (Saez de Urabain *et al*., 2014).

The last problem related to the calibration procedure is finding the right distance and viewing angle between the eye tracker and the participant. During the first 3 to 4 months in particular, infants are generally not able to accommodate as a function of target distance, such that they can only see objects in focus within a certain distance (Salapatek, Bechtold, & Bushnell, 1976). Thus, for subjects younger than 3-4 months, the distance between the infant and the presentation screen needs to be considerably shorter (around 30-40 cm) than the distance that is recommended for most eye-tracking systems (around 60 cm; Holmqvist *et al*., 2011; Saez de Urabain *et al*., 2014), effecting data quality.

Infants' eyelids can be particularly watery, especially during the first few months of life, and this can considerably interfere with the glint detection process. Bright pupil techniques (placing the IR illuminators on or around the optic axis of the camera to make the light reflect back off the retina) are considered to be more accurate than dark pupil ones when dealing with certain eye physiologies like light iris color or watery eyelids (Gredebäck, Johnson, & von Hofsten, 2009). But aren't watery eyes a sign of an infection such as pink eye?

To date, even though some procedures have been developed for calibrating infant eye movements (Gredebäck *et al*., 2009), there are still no standards for performing and evaluating a calibration procedure (Oakes, 2012; see also Frank, Vul, & Johnson, 2009), and thus it is up to the researcher to decide whether a calibration procedure was successful or not. It is also important that researchers understand how the analysis software they are using generates summary statistics (e.g., AOI dwell times, mean fixation durations and saccade latencies) from the raw sample data recorded by the tracker. Given the potential for increased spatial and temporal noise in infant raw gaze data, parsing algorithms that use hard (what does 'hard' mean here? Something like 'stringent'?) thresholds for classifying fixations (e.g., spatial dispersal) or saccades (e.g., a minimum velocity or acceleration) will be artificially triggered by flickery data and produce spurious oculomotor events that are not representative of the actual visual cognition going on (Saez de Urabain *et al*., 2014; Wass, Smith, &

Johnson, 2013). Several open-source software solutions now exist that allow smoothing, interpolation, and hand cleaning of noisy eye-tracking data, which have been shown to produce reliable oculomotor measures (Frank *et al.*, 2009; Saez de Urabain *et al.*, 2014; Wass *et al.*, 2013). These techniques are also being incorporated in commercial platforms, and hardware manufacturers are beginning to tackle the problem and allow high-speed, high-spatially accurate infant eye tracking.

Conclusions

Eye tracking is an increasingly popular tool for investigating development. The non-invasive nature of eye tracking makes it attractive for developmental scientists whilst providing high volume observations of behavior and real-time cognition. However, the apparent ease with which commercial eye-tracking tools can be used by relative novices means that many researchers employ eye tracking without fully engaging with the assumptions the software is making about their data. Greater understanding of all stages involved in gathering and analysing eye-tracking data will produce more interesting and replicable insights into development.

Eye tracking holds great promise for providing fine-grained insights into developmental cognitive neuroscience of both typical and atypical developing populations, especially when combined with other measures such as those used in psychophysiology (e.g., heart rate, galvanic skin response), EEG, NIRS, fMRI, or naturalistic observations of behaviour with head-mounted tracking.

See also:

Head-mounted eye tracking; Observational methods; Attention; Biological motion perception; Cognitive development during infancy; Face perception and recognition; Perceptual-motor calibration and space perception; Vision; Perception and action; Attentional deficit hyperactivity syndrome (ADHD); Autism

Further reading

Please ensure APA referencing style here, in the text and the References. Also, do not include doi's.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking. A comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.

Saez de Urabain, I. R., Johnson, M. H., & Smith, T. J. (2014). GraFIX: A semiautomatic approach for parsing low- and high-quality eye-tracking data. *Behavior Research Methods, 47*, 53-72.