Using Smooth Pursuit Calibration for Difficult-to-Calibrate Participants

Pieter Blignaut
University of the Free State
Bloemfontein, South Africa

Although the 45-dots calibration routine of a previous study (Blignaut, 2016) provided very good accuracy, it requires intense mental effort and the routine proved to be unsuccessful for young children who struggle to maintain concentration. The calibration procedures that are normally used for difficult-to-calibrate participants, such as autistic children and infants, do not suffice since they are not accurate enough and the reliability of research results might be jeopardised.

Smooth pursuit has been used before for calibration and is applied in this paper as an alternative routine for participants who are difficult to calibrate with conventional routines. Gaze data is captured at regular intervals and many calibration targets are generated while the eyes are following a moving target. The procedure could take anything between 30 s and 60 s to complete, but since an interesting target and/or a conscious task may be used, participants are assisted to maintain concentration.

It was proven that the accuracy that can be attained through calibration with a moving target along an even horizontal path is not significantly worse than the accuracy that can be attained with a standard method of watching dots appearing in random order. The routine was applied successfully for a group of children with ADD, ADHD and learning abilities.

This result is important as it provides for easier calibration – especially in the case of participants who struggle to keep their gaze focused and stable on a stationary target for long enough.

Introduction

Video-based eye tracking is based on the principle that near-infrared light shone onto the eyes is reflected off the different structures in the eye to create four Purkinje reflections (Crane & Steele, 1985). The standard way of calibrating such eye trackers is through presentation of a series of dots (or gaze targets) at known positions on the display and expect the participant to watch the dots until enough gaze data is sampled (Nyström, Andersson, Holmqvist, & van de Weijer, 2013). While expensive commercial systems utilise a model of the eye to compute the gaze direction (Hansen & Ji, 2010), self-assembled eye

trackers use mostly polynomial expressions to map the relative position of the pupil to the corneal reflections (the so-called pupil-glint vector) to gaze coordinates. A least squares estimation is used to minimise the distances between the observed points and the actual points in the calibration grid (Hoormann, Jainta, & Jaschinski, 2008).

Normally, five or nine dots are used. The more dots that are used, the better the accuracy of the system should be. Good accuracy is important when the stimuli is close to each other as in reading, where a researcher wants to determine the number of fixations on individual syllables. A procedure is described in a previous study (Blignaut, 2016) where 45 dots are displayed in a 9×5 grid. Twenty-three of the dots are used as calibration targets, while the complete set of dots is used to select the best possible regression polynomial. The dots are displayed in random order to prevent participants to pre-empt the position of the next dot and take their eyes away from a dot before the gaze was registered.

While the procedure described in Blignaut (2016) is accurate with a reported average offset of 0.32°, it requires intense and prolonged concentration and participants do not always understand that they have to keep their eyes fixated on a dot until the next one appears. Unsurprisingly, the routine proved to be unsuccessful for young children with Attention Deficit Disorder (ADD), Attention Deficit Hyperactivity Disorder (ADHD) and learning disabilities. An occupational therapist using the system complained that young the children with these conditions did not understand exactly what was expected of them and some of them could not maintain concentration for the entire period.

The challenge is, therefore, to capture gaze data at as many known locations as possible, with the least possible mental effort while maintaining attention on the target. In this paper, a smooth pursuit calibration routine is proposed with a target moving across the display at a constant speed. The target could also be an animated image of something of interest to a small child, such as a butterfly or an airplane. In order to further motivate the child participant to watch the target closely, it could change colour, shape or image at varying intervals and the child could be challenged to count the number of changes.

The need for calibration and existing calibration procedures are discussed in the following section. The difficulties that are experienced with the standard routines to calibrate certain groups of participants (collectively referred to as difficult-to-calibrate (DC) participants) are highlighted and previous attempts to solve the problem are discussed. Thereafter, the presentation of a moving target with a related task is offered as a solution to capture the attention of the DC participants for long enough so that the procedure can be completed.

The evaluation of smooth pursuit calibration (SPC) is done in two phases: First, the accuracy of the approach is validated based on comparison with a standard calibration procedure using able and cooperating participants. Second, the applicability of the approach is validated for a group of early primary school children with various forms of one or more cognitive disorders.

The paper concludes with a discussion of the results.

The Role of Calibration

The need for calibration

The output from eye-tracking devices varies with individual differences in the shape or size of the eyes, such as the corneal bulge and the relationship between the eye features (pupil and corneal reflections) and the foveal region on the retina. Ethnicity, viewing angle, head pose, colour, texture, light conditions, position of the iris within the eye socket and the state of the eye (open or closed) all influence the appearance of the eye (Hansen & Ji, 2010) and, therefore, the quality of eye-tracking data (Holmqvist et al., 2011). In particular, the individual shapes of participant eye balls, and the varying positions of cameras and illumination require all eye-trackers to be calibrated.

The procedure

Calibration refers to a procedure to gather data so that the coordinates of the pupil and one or more corneal reflections in the coordinate system of the eye-video can be converted to x- and y-coordinates that represent the participant's point of regard in the stimulus space. The procedure usually consists of asking the participant to look at a number of pre-defined points at known angular positions while storing samples of the measured quantity (Abe, Ohi, & Ohyama, 2007; Kliegl & Olson, 1981; Tobii, 2010). There is no consensus on exactly when to collect these samples, but Nyström et al. (2013) showed that participants know better than the operator or the system when they are looking at a target.

Mapping to point of regard

The transformation from eye-position to point of regard can be either model-based (geometric) or interpolation-based (Hansen & Ji, 2010). With model-based gaze estimation, a model of the eye is built from the observable eye features (pupil, corneal reflection, etc.) to compute the gaze direction. In this case, calibration is not used to determine the actual gaze position but rather to record the eye features from different angles. See Hansen and Ji (2010) for a comprehensive overview of possible transformations.

Interpolation might involve, for example, a linear regression between the known data set and the corresponding raw data, using a least squares estimation to minimize the distances between the observed points and the actual points (Hoormann et al., 2008). Other examples of 2-dimensional interpolation schemes can be found in McConkie (1981) as well as Kliegl and Olson (1981) while

a cascaded polynomial curve fit method is described in Sheena and Borah (1981).

Theoretically, the transformation should remove any systematic error, but the limited number of calibration points that are normally used limits the accuracy that can be achieved. Typical calibration schemes require 5 or 9 pre-defined points, and rarely use more than 20 points (Borah, 1998).

Auto-calibration

Huang, Kwok, Ngai, Chan, and Leong (2016) presented an auto-calibrating system that identifies and collects gaze data unobtrusively during user interaction events since there is a likely correlation between gaze and cursor and caret locations. The procedure presented by Huang et al. (2016) recalibrates continuously and becomes more and more accurate with additional use. They reported an average error of 2.56° which is not good but has the advantage that there is no need for an explicit calibration phase.

Swirski and Dodgson (2013) describe a procedure that fits a pupil motion model to a set of eye images. Information from multiple frames is combined to build a 3D eye model that is based on assumptions on how the motion of the pupil is constrained. No calibration is needed and since the procedure is based on pupil ellipse geometry alone, it is not necessary to illuminate the eyes to create a corneal reflex. At best, a mean error of 1.68° was reported.

In summary, while auto-calibration might solve the problem of calibrating for participants who struggle to maintain concentration, it is not good enough for studies where high accuracy is needed.

Previous Attempts to Track Difficult-to-Calibrate Participants

In the quest for a solution to calibrate young children who find it difficult to concentrate on a target for the duration of a calibration routine, one can learn from the experience of others who faced similar challenges, for example tracking infants, toddlers and children with autism.

Tracking infants and toddlers pose a challenge as it is hardly ever possible to get them to sit down long enough to focus on calibration targets. Aslin (2012) mentioned that small flashing (or shrinking) targets work well with infants, but argued that accuracy is unlikely to ever be better

than 1° because infants are unable to precisely and reliably fixate small stimuli. He further asserted that if 1° of accuracy is insufficient to answer a particular question, then an eye tracker should not be used for the research and alternative methods should be implemented.

Sasson and Elison (2012) indicated that eye tracking of young children with autism involves unique challenges that are not present when tracking normal-developing older children or adults. They used the normal calibration routines provided by the manufacturer to track the gaze data of their participants, but used large stimuli, spanning more than 5°. Although participants find such stimuli pleasing to look at, the researcher cannot be exactly sure where the participant looked at the time of data capture. This will almost certainly result in bad accuracy that will not be feasible for tasks where high accuracy is required, such as reading.

In a study by Pierce, Conant, Hazin, Stoner, and Desmond (2011), toddlers were seated on their parent's lap in front of a Tobii T120 eye tracker and a partition separated the operator from the toddler. To obtain calibration information, toddlers were shown images of an animated cat that appeared in 9 locations on the screen. Using a software facility that superimposes the point of regard on the test image in real time, the operator observed the infant's gaze position and head position on a secondary monitor, making note of obvious deviations from expected gaze positions. The entire process was repeated if the infant's eyes were no longer picked up. No mention was made of the achieved accuracy, but it is reasonable to expect that the accuracy could not be better than the size of the calibration stimulus (the animated cat).

Franchak, Kretch, Soska, and Adolph (2011) used a head-mounted eye tracker but displayed stimuli on a computer screen. A sounding target appeared at a single location within a 3×3 matrix on the monitor to induce eye movements. Calibration involved as few as 3 and as many as 9 points spread across visual space. Subjective judgement was used to determine whether fixations deviated from targets by more than about 2° and the procedure was repeated if necessary. Although the spatial accuracy is lower than that of typical desk-mounted systems, it was regarded as adequate for determining the target of fixations in natural settings. The entire process of preparing the equipment and calibrating the infant took about 15 minutes.

Corbetta, Guan, and Williams (2012) followed a similar procedure to calibrate an ETL-500 head-mounted eye tracker through timely coordination between one experimenter facing the child and another experimenter running the interactive calibration software of the eye tracker. 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 capture the gaze data. The researchers did not report the accuracy achieved, but one can once again assume that the accuracy could not be better than the size of the toy used as calibration stimulus.

In summary, it is clear that in an attempt to calibrate so-called difficult-to-calibrate participants, various researchers used sound and animation of larger objects as calibration targets. Furthermore, the number of calibration points is mostly limited and the accuracy that can be achieved is not expected to be better than 2°. It is also difficult to tell the actual accuracy that was obtained during a specific experimental set-up or participant recording. The need exists, therefore, for a calibration routine that is easy to execute and can be used for difficult-to-calibrate participants, yet accurate enough to provide reliable research results – especially if the experiment involves smaller or closely spaced targets.

Smooth Pursuit Calibration

Smooth pursuit eye movements

Smooth-pursuit eye movements are continuous, slow rotations of the eyes (Spering & Gegenfurtner, 2008) that are used to stabilise the image of a moving object of interest on the fovea, thus maintaining high acuity (Nagel, Sprenger, Steinlechner, Binkofski, & Lencer, 2012; Thier & Ilg, 2005). Conscious attention is needed to maintain accurate smooth pursuit (Hutton & Tegally, 2005; Madelain, Krauzlis, & Wallman, 2005).

Smooth pursuit gain is expressed as the ratio of smooth eye movement velocity to the velocity of a foveal target (Sharpe, 2008). If the gain is less than 1, gaze will fall behind the target to create a retinal slip that will have to be reduced by one or more "catch-up" saccades (Van Gelder, Lebedev, Liu, & Tsui, 1995). According to Meyer, Lasker, and Robinson (1985), normal subjects can follow a target with a gain of 90% up to a target velocity of 100 deg/s.

Smooth pursuit gain increases with age, especially for the first 3 months of an infant's life (Von Hofsten & Rosander, 1997; Richards & Holley, 1999). Accardo, Pensiero, Da Pozzo, and Perissutti (1995) found that velocity gain of children aged 7-12 is slightly lower than that of adults.

Smooth pursuit can also be affected by attention (Souto & Kerzel, 2011; Van Donkelaar & Drew, 2002). More specifically, performance with smooth pursuit tasks can be dramatically improved when subjects are asked to analyse some or other changing characteristic of the target, such as reading a changing letter or number on the target (Holzman, Levy, & Proctor, 1976) or pressing a button (Iacono & Lykken, 1979).

Smooth pursuit performance is also affected by stimulus background (Brenner, Smeets, & van den Berg, 2001; Kerzel, Souto, & Ziegler, 2008; Lindner, Schwarz, & Ilg, 2001), target position (J. Pola & Wyatt, 2001), target velocity (Kowler & McKee, 1987; Meyer et al., 1985), target visibility (Becker & Fuchs, 1985; Pola & Wyatt, 1997), target direction (Engel, Anderson, & Soechting, 2000) and predictability of target direction (Soechting, Rao, & Juveli, 2010).

Smooth pursuit impairment and dysfunction can be linked to mental illnesses such as schizophrenia (Holzman & Levy, 1977; Holzman et al., 1976; Levin et al., 1988; O'Driscoll & Callahan, 2008), autism (Takarae, Minshew, Luna, Krisky, & Sweeney, 2004), physical anhedonia and perceptual aberrations (Gooding, Miller & Kwapil, 2000; O'Driscoll, Lenzenweger, & Holzman, 1998; Simons & Katkin, 1985), Alzheimer's disease (Fletcher & Sharpe, 1988) and attention-deficit hyperactivity disorder (ADHD) (Fried et al., 2014).

Smooth pursuit calibration in general

The concept of calibrating while a participant follows a moving target has been exploited with success in the past. Pfeuffer, Vidal, Turner, Bulling, and Gellersen (2013) explains a procedure where gaze data for calibration is only sampled when the participant is attending to the target as indicated by high correlation between eye and target movement. They showed that pursuit calibration is tolerant to interruption and can be used to calibrate without participants being aware of the procedure.

Pfeuffer et al. (2013) used a Tobii TX300 eye tracker to test their calibration procedure and collected gaze data at 60 Hz. At a target speed of 5.8°/s, it took 20 seconds to

complete the target trajectory and an accuracy of just less than 0.6° were achieved. The results were compared with the 5-point calibration routine of Tobii which took 19 seconds to complete and delivered an accuracy of ≈0.7°.

Celebi, Kim, Wang, Wall, and Shic (2014) follows the approach of Pfeuffer et al. (2013) but argues that an Archimedean spiral would provide better spatial coverage of the stimulus plane with little redundancy. At a linear velocity of 6.4°/s, the calibration procedure took 27 seconds during which 1600 data points were collected. Upon testing 10 healthy adults on an Eyelink 1000 eye tracker running at at 500 Hz, their approach delivered an average accuracy of 0.84° compared to 1.39° with a standard 9-point calibration procedure (which took 23 seconds to complete).

Celebi et al. (2014) stated explicitly that the goal of their smooth pursuit approach towards calibration is to improve the calibration for toddlers and children with or without developmental disabilities although they did not test the approach with such participants.

Gredebäck, Johnson, and Von Hofsten (2010) describes a calibration routine that makes use of a moving object to lure infants' eyes to 2 or 5 calibration targets, but they reported accuracy according to the manufacturer's specifications of 0.5° - a value which has been computed with a conventional calibration routine under ideal circumstances and with cooperating adult participants.

Smooth pursuit calibration for DC participants

Participants who struggle to maintain concentration on tedious tasks can be cognitively stimulated by indicating or counting the number of transitions of the target from one stimulus to another (Holzman et al., 1976). In this study, participants were requested to follow a grey disk of 1.5° diameter on a white background (Figure 1) that contained a coloured dot (0.2°) in the centre. The dot changed colour in cycles of blue (2 seconds) and red (500 ms) and the participants were then asked to say the word "Red" aloud every time that the disk changed to red.

The target is initially displayed statically in the top left corner and the participant can be prepared as to the direction and nature of the motion that could be expected. The experimenter can initiate movement with a button as soon as the participant is ready. Three alternative trajectories were tested with the target moving along an even horizontal path, a wavy horizontal path and vertical (*cf* Figure 1).
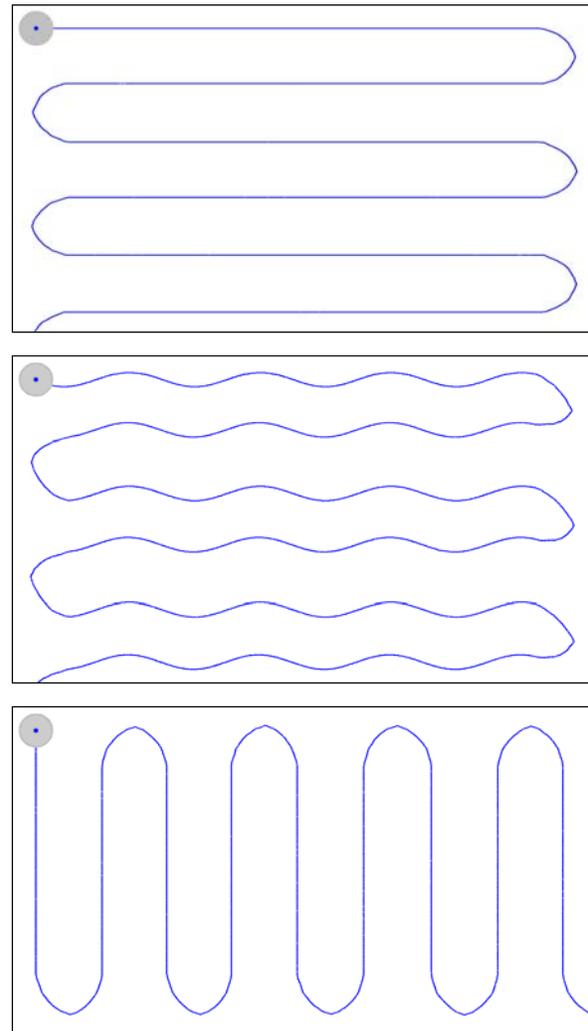
Figure 1. Trajectories of moving target

Depending on the speed of movement, the interval between windows and the trajectory, a large number of targets can be extracted from the continuous gaze data. In the procedure of Pfeuffer et al. (2013), gaze data is collected whenever a participant attends to the moving target. In our approach, gaze samples (or more specifically, pupil-glint vectors for each eye) are captured for very short windows (100 ms) at intervals of 500 ms. When data is not available at a specific interval, the point is ignored. This means that, at a framerate of 200 Hz, 20 samples were recorded within a 100 ms window. At a velocity of 6.65°/s (gaze distance 700 mm, 300 px/s on a 19.5", 1600×900 screen), the samples would span 0.665° in the direction of movement.

The radius of curvature was set so that the horizontal and vertical trajectories would cover 6 and 9 distinct Y and X coordinates respectively (*cf* Figure 1). This resulted in 76 targets (in 38 s) being captured for both the even and wavy horizontal movements and 66 targets (in 33 s) for the vertical movement.

Since the eyes move smoothly to follow the target, it can be expected that a convex hull around the sample points would be elongated along the direction of movement. The samples within every window were sorted according to the x and y dimensions of the pupil-glint vectors and only the intersection of the centre 80% of samples around the median in each dimension are retained.

In contrast with a standard 5-point or 9-point calibration procedure where all points are needed for the regression, the multitude of points that are available with this procedure allows the removal of points where participants blinked or where their attention was distracted. All windows for which the dispersion (`Max(maxX-minX, (maxY-minY))`) of contained samples are above 5°, are also removed. For each of the remaining windows of gaze data samples, the average location and the average pupil-glint vector are calculated.

From here on, the procedure as explained in Blignaut (2016) is followed. The gaze data windows represent calibration points at known locations and are used to determine a gaze mapping polynomial set per participant. Regression coefficients are recalculated in real-time – based on a subset of calibration points in the region of the current gaze. Real-time localized corrections are done that are based on calibration targets in the same region. See Blignaut (2016) for a detailed discussion of the procedure.

Accuracy of Smooth Pursuit Calibration

In this section, the accuracy of the approach is validated based on a comparison with a standard calibration procedure using healthy and cooperating adult participants. The applicability of the approach for difficult-to-calibrate participants will be addressed in the next section.

Equipment

For this study, an eye tracker with two infrared illuminators, 480 mm apart, and the UI-1550LE camera from IDS Imaging (https://en.ids-imaging.com) was assembled. All recordings were made at a framerate of 200 Hz.

Every frame that is captured by the eye camera was analysed and the centres of the pupils and the corneal reflections (glints) were identified. A regression-based approach was followed to map the pupil-glint vector to a point of regard in display coordinates. The regression coefficients are determined through a calibration process.

Method

Seventeen healthy and cooperating adult participants were recruited through convenience sampling and presented with four calibration routines, namely a moving target along an even horizontal path, a moving target along a wavy horizontal path and a target moving vertically (*cf* Figure 1). The 45-dots routine as proposed in a previous study (Blignaut, 2016) was also presented for comparison purposes. The procedure was executed only once for every participant.

After every routine, a 7×4 grid of dots was displayed to determine the accuracy of the procedure. As for the 45 dots, the 28 dots appeared in random order to prevent participants from pre-empting the position of the next dot and prematurely look away. The regression coefficients as determined in the preceding calibration routine was used to map the gaze data to screen coordinates. The accuracy for a specific participant was calculated as the average offset between the known locations and the reported gaze coordinates across the 28 dots. The performance of a calibration routine is expressed as the average accuracy over all participants.

Recording of calibration points

Figure 2 shows the calibration points that were recorded for a specific participant while the target was moving along an even horizontal path. The mapped gaze coordinates of the sample data are enclosed by convex hulls – green for the left eye and red for the right eye.

In the example presented in Figure 2, five of the calibration windows did not contain enough sample data – probably due to blinks. Table 1 shows that for routines that involve a moving target, on average between 2 and 4 calibration windows are lost in this way. Since there are more than enough other points to be used in the subsequent regression and because the lost points are seldom at successive locations, this does not pose a problem.
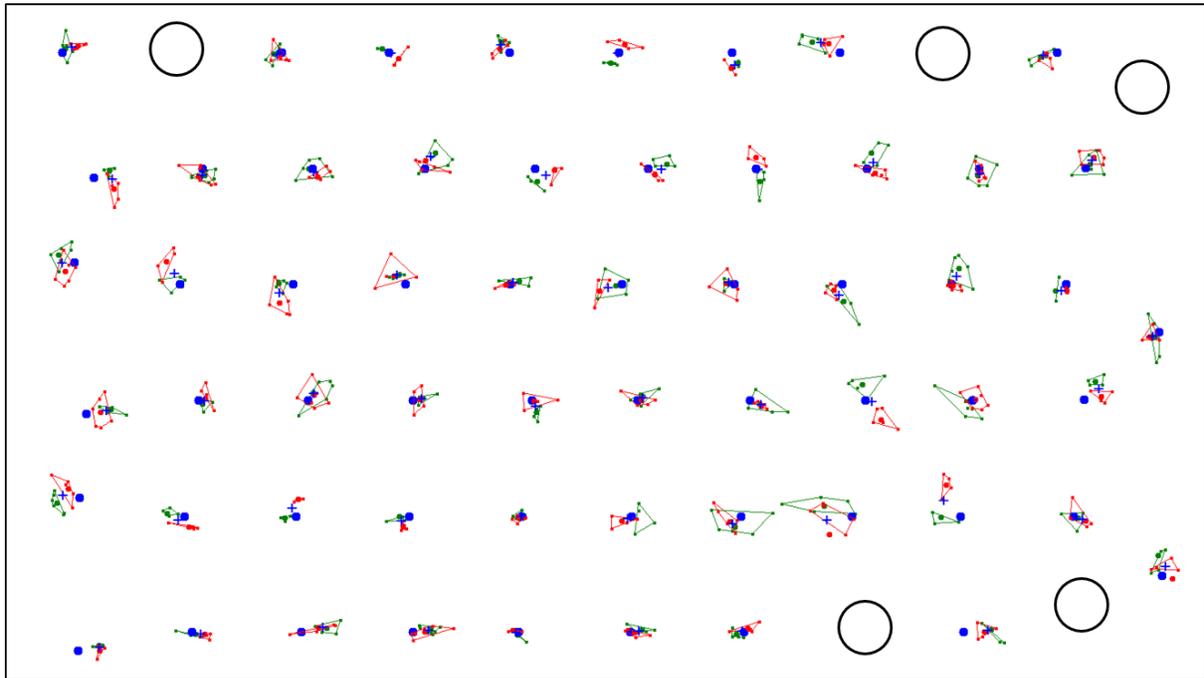
Figure 2. Calibration points with accompanying gaze data samples as captured with a horizontally moving target. Missing points marked with circles.

Initially, the set of calibration points was also used as validation targets and the offsets between the calibration points and the mapped gaze coordinates were calculated. All points with offsets larger than 1.0° were then removed from the set of calibration targets and the regression procedure was repeated. The remaining calibration points are shown in blue in Figure 2, while black dots indicate points that were excluded for real-time interpolation. Table 1 also shows the final number of points that was used to determine the polynomial coefficients through regression.

Table 1
Average number of points (across participants) with enough samples and with mapped gaze coordinates within 1° of the target per calibration routine (SD: Standard deviation)

Routine	Possible points	Points with enough samples			Points within 1° of target		
		Avg	SD	%	Avg	SD	%
45 dots	45	45	0	100	39.3	10.0	87.3
Even hor	76	72.5	6.1	95.4	69.1	8.2	90.9
Wavy hor	76	74.1	2.9	97.5	69.7	7.1	91.6
Vertical	66	63.9	3.3	96.9	60.9	6.0	92.3

Validation results

Figure 3 shows the validation points for a calibration that was done for a moving target along an even horizontal path. The average of all samples within a window is shown for the left and right eyes. A + indicates the average position between the left and right eyes.

The example in Figure 3 was specifically selected to illustrate the occurrence of outliers. These outliers may occur if the participant loses concentration or is distracted by external stimuli. Sometimes (some of) the samples are captured during a blink, in which case the samples for the left and right eyes appear to be disconnected. Validation points were excluded from the calculation of average offset if the offset was larger than 3° or if the mapped gaze coordinates for the two eyes were more than 3° apart. Table 2 shows the average number of validation points that was included for each of the calibration routines. These thresholds were set large enough not to exclude valid data but small enough to ensure that unwanted gaze behaviour is excluded.
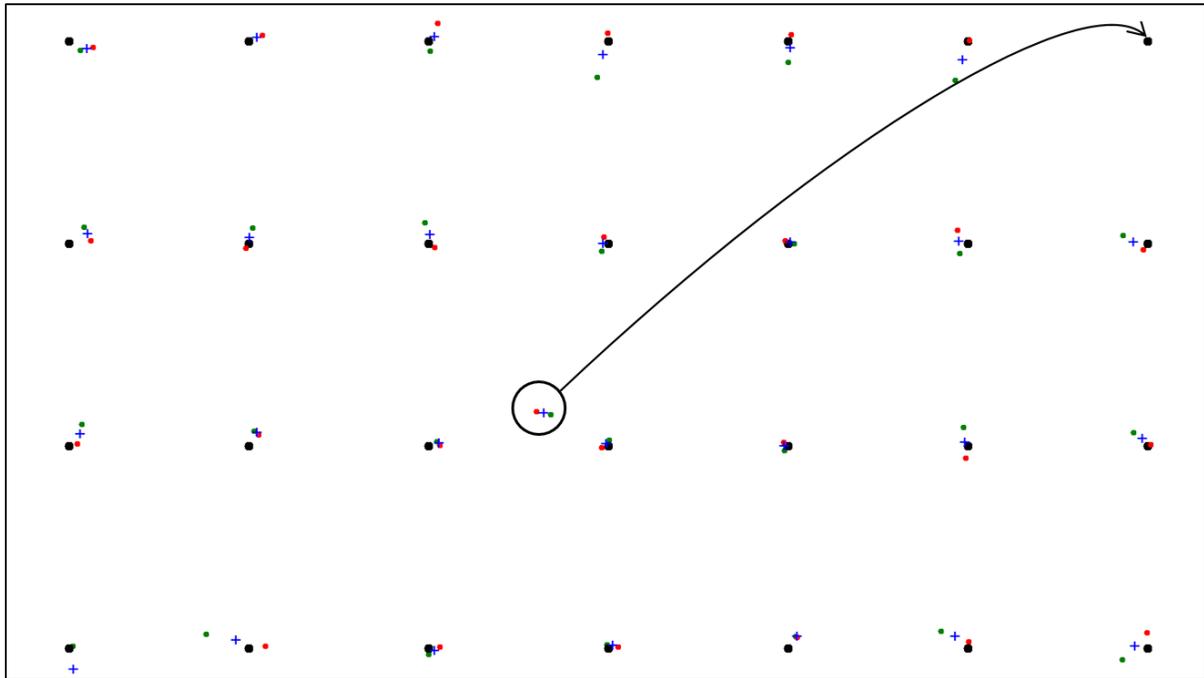
Figure 3. Validation points with left (green) and right (red) eye averages of samples per point. The average between eyes is indicated with a blue +. A point where the participant was distracted, is also shown.

Table 2 also shows the average error across the 28 validation points and 17 participants per calibration routine. The important column that needs to be interpreted to compare the four calibration routines is boldfaced. Figure 4 provides a visualisation of the same results. The vertical bars denote the 95% confidence intervals of the means.

Table 2

Average number of validation points that was included and the average error (over participants and validation targets) for each of the calibration routines. (SD: Standard deviation, SEM: Standard error of the mean)

	Number of points				Error (degrees)				
Routine	Min	Max	Avg	SD	Min	Max	**Avg**	SD	SEM
45 dots	14	28	25.7	4.4	0.31	0.65	**0.47**	0.10	0.024
Even hor	24	28	27.1	1.2	0.41	0.68	**0.53**	0.08	0.019
Wavy hor	19	28	26.6	2.3	0.39	1.03	**0.61**	0.17	0.041
Vertical	23	28	26.6	1.6	0.45	1.18	**0.68**	0.18	0.043
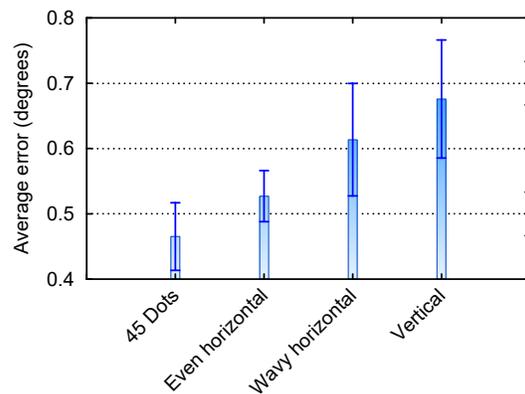
Figure 4. Average error over participants and validation targets for four calibration routines. The vertical bars denote the 95% confidence intervals of the means.

A repeated measures analysis of variance (each participant calibrated with four different routines) showed that the calibration routine has a significant ($\alpha = .001$) effect on the magnitude of the error ($F(3,48) = 9.74$, $p < .000$). Table

3 shows the results of Tukey's test for the honestly significant difference between pairs of means. The differences between the means for the 45-dots and vertical movement as well as the difference between 45-dots and horizontal movement along a wavy path were significant ($\alpha = .01$). Although the accuracy for the moving target along an even horizontal path was worse than that of the 45-dots routine, it was not significantly so ($\alpha = .05$).

It can, therefore, be concluded that a moving target along an even horizontal path has the potential to be used as alternative calibration routine. This will be tested in the next section with a sample of difficult-to-calibrate participants.

Table 3
p-Values for the significance of the difference in error between pairs of means

	Even hor	Wavy hor	Vertical
45 dots	0.465	0.005	0.000
Even hor		0.183	0.005
Wavy hor			0.457

Applicability of Smooth Pursuit Calibration for Difficult-to-Calibrate Participants

Equipment

The same self-assembled eye tracker was used as in the previous section.

Method

A school for learners with special education needs were visited and all learners from Grade 1 to Grade 3 (ages 6 – 11) for which permission of the parents were obtained, were tested. The school accommodates learners who are cerebrally palsied, physically and/or learning disabled. The school has specially qualified remedial teachers as well as a multi-disciplinary support structure that includes psychologists, social workers, occupational therapists, speech therapists, physiotherapists as well as a professional nurse. The school follows the normal mainstream syllabi but there are no more than 10 learners in a class to enable teachers to provide specialised and individual attention.

Several lessons were learned in the process of capturing data. Initially, learners were requested to follow the moving target without any further instruction. It soon became evident that they struggle to maintain focus on the target for the duration of the trajectory. The target was then programmed to change colour in cycles of blue (2 seconds) and red (500 ms) and learners were instructed to call out the word "Red" whenever the target changes to red. For the procedures where dots were involved, every dot appeared in a different colour and learners were instructed to call out the colour for the dot every time.

Although the system allows moderate head movements, many learners had excessive sideways and back and forth head movements – some of which were involuntary. A chinrest was then used to maintain head position, but it caused instability of the eyes every time that the learners vocalised their response on a colour change of the target. Finally, the learners were instructed to push their foreheads against a barrier that was set such that a fixed gaze distance of 700 mm was maintained.

It was also realised that the sets of 45 dots for calibration and 28 dots for validation was too exhausting and therefore these were limited to 23 calibration targets (in rows of 5, 4, 5, 4, 5 targets each) and 15 validations targets in a grid of 5×3.

Eventually, 24 participants were tested with the final configuration of target movement, headrest and calibration sets. The number of learners per grade and condition is shown in Table 4. Note that some learners had more than one condition.

Table 4
Number of learners per grade and condition
(ADD: Attention deficit disorder; ADHD: Attention deficit hyperactivity disorder; ASP: Asperger syndrome; DSL: Dyslexia; EPSY: Epilepsy; LD: Learning disability)

Grade	n	Age	ADD	ADHD	ASP	DSL	EPSY	LD
1	4	6.75	1	3		1	1	1
2	13	8.62	5	2	2			6
3	7	9.86		4				4

The learners were presented with a moving target along an even horizontal path (SP) (*cf* Figure 1) and the 23-dots routine. After every routine, the 5×3 grid of dots was displayed to determine the accuracy of the procedure. As was the case for the validation of accuracy with healthy adults, the dots appeared in random order to prevent learners from pre-empting the position of the next dot and prematurely look away. The performance of both the 23-dot and SP calibration routines was expressed as the average accuracy over all participants as determined through the 15 dots validation routine.

Validation results

As for the validation of accuracy with healthy adults, validation points were excluded from the calculation of average offset if the offset was larger than 3° or if the mapped gaze coordinates for the two eyes were more than 3° apart. Table 5 shows the average number of validation points that were included for each one of the calibration routines. A repeated measures analysis of variance for the effect of calibration routine (23-dots vs SP) on the number of valid points showed that the SP routine leads to more reliable data as there are significantly less points that have to be discarded ($F(1,47) = 47.7$, $p < .001$).

Table 5 also shows the average error across the 15 validation points and 24 participants per calibration routine. A repeated measures analysis of variance for the effect of calibration routine (23-dots vs SP) on the accuracy of tracking showed that the SP routine is significantly better than a dots-based routine ($F(1,47) = 12.57$, $p < .000$) for difficult-to-calibrate participants.

Table 5
Average number of validation points that was included and the average error (over participants and validation targets) for each of the calibration routines. (SD: Standard deviation, SEM: Standard error of the mean)

	Number of points				Error (degrees)				
Routine	Min	Max	Avg	SD	Min	Max	**Avg**	SD	SEM
28 dots	2	15	10.6	3.34	0.60	2.49	**1.15**	0.39	0.080
Even hor	9	15	13.7	1.67	0.53	1.77	**0.94**	1.15	0.235

Summary

The conventional way of calibrating remote video-based eye trackers is through presentation of a series of gaze targets at known positions while participants are expected to watch the targets. For regression-based mapping of eye features to gaze coordinates, more gaze targets normally mean better (more accurate) calibration. Unfortunately, more gaze targets also require more mental effort from participants. Through informal observations, it was realised that, although the 45 dots-routine of a previous study (Blignaut, 2016) provided very good accuracy, it expected too much mental effort for participants who struggle to maintain concentration.

Depending on the type of experiment, better accuracy might be expected than can be achieved with calibration free or auto-calibrating systems. The calibration procedures that are normally used for infants, toddlers and autistic children do also not suffice since they are not accurate enough and the reliability of research results might be jeopardised.

In this paper, the use of smooth pursuit with a target moving across the display at a constant speed, is proposed. This approach is motivated by the fact that attention to a moving target can be maintained more easily – especially if accompanied by a concurrent and related task such as analysis of some or other changing characteristic of the target (Holzman et al., 1976; Iacono & Lykken, 1979).

While the participant is following the target, gaze data is captured at regular intervals and many calibration targets are saved that can be used in subsequent regression and interpolation. Because of the abundance of points, the procedure allows the exclusion of points of dubious quality. Depending on the speed of movement and the trajectory, the procedure could take anything between 30 s and 60 s to complete.

Validation of the proposed routine was done in two phases: The accuracy of the routine was validated by comparing its performance with that of a standard calibration procedure for healthy and cooperating adults. Thereafter, the applicability of the approach for participants who are normally difficult to calibrate, is validated by applying it for a group of early primary school children with various forms of one or more cognitive disorders.

It was proven through a repeated measures, within-participants, analysis of variance that the accuracy that can be attained through calibration with a moving target along an even horizontal path is not significantly worse than the accuracy that can be attained with a standard method of watching dots. Accuracy of around the 0.5° mark were obtained for both routines for a group of seventeen adults which is comparable with the 0.6° attained by Pfeuffer et al. (2013) and better than the 0.84° attained by Celebi et al. (2014).

For a group of young children with various forms of cognitive disorders such as ADD, ADHD and learning disabilities, smooth pursuit calibration proved to be superior to the standard routine. For this group, an average accuracy of below 1° could be achieved with SP while it was not the case with a standard routine of 28 dots. This is a significant improvement on the 1.5°-2.5° errors that can be attained by calibration-free or auto-calibrating routines such as those of Huang et al. (2016) and Swirski and Dodgson (2013).

Future Research

Since smooth pursuit ability develops until the age of adolescence (Accardo et al., 1995), one can expect that older children will benefit even more from the smooth pursuit approach. This needs to be investigated.

Furthermore, the smooth pursuit approach was tested above for children with ADD, ADHD and learning disabilities. No children with autism were tested and it remains to be seen of the approach will work for such conditions since it is known that smooth pursuit is impaired in autism and similar conditions (Pierce et al., 2011; Takarae et al., 2004).

References

Abe, K., Ohi, S., & Ohyama, M. (2007). An eye-gaze input system using information on eye movement history. In C. Stephanidis (Ed.), *Universal Access in HCI*, Part II. HCII2007, LNCS 4555, 721-729. Berlin, Heidelberg: Springer Berlin Heidelberg.

Accardo, A. P., Pensiero, S., Da Pozzo, S., & Perissutti, P. (1995). Characteristics of horizontal smooth pursuit eye movements to sinusoidal stimulation in children of primary school age. *Vision Research, 35*(4), 539-548.

Aslin, R.N. (2012). Infant eyes: A window on cognitive development. *Infancy, 17*(1), 126-140. doi:10.1111/j.1532-7078.2011.00097.x

Becker, W., & Fuchs, A.F. (1985). Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Experimental Brain Research, 57*(3), 562-575. doi:10.1007/BF00237843.

Blignaut, P.J. (2016). Idiosyncratic feature-based gaze mapping. *Journal of Eye Movement Research, 9*(3), 1-17. doi:10.16910/jemr.9.3.2.

Borah, J. (1998). Technology and application of head based control. *RTO Lecture series on Alternative Control Technologies,* 7-8 October 1998, Brétigny, France and 14-15 October 1998, Ohio, USA.

Brenner, E., Smeets, J.B., & van den Berg, A.V. (2001). Smooth eye movements and spatial localisation. *Vision Research, 41*(17), 2253-2259.

Celebi, F.M., Kim, E.S., Wang, Q., Wall, C.A., & Shic, F. (2014). A smooth pursuit calibration technique. Eye Tracking Research and Applications (ETRA), 26-28 March 2014, Safety Harbor, Florida.

Corbetta, D., Guan, Y., & Williams, J.L. (2012). Infant eye-tracking in the context of goal-directed actions. *Infancy, 17*(1), 102-125. doi:10.1111/j.1532-7078.2011.0093.x

Crane, H.D., & Steele, C.M. (1985). Generation-V dual-Purkinje-image eyetracker. *Applied Optics, 24*(4), 527-537. doi:10.1364/AO.24.000527

Engel, K.C., Anderson, J.H., & Soechting, J. F. (2000). Similarity in the response of smooth pursuit and manual tracking to a change in the direction of target motion. *Journal of Neurophysiol, 84*(3), 1149-1156.

Fletcher, W.A., & Sharpe, J.A. (1988). Smooth pursuit dysfunction in Alzheimer's disease. *Neurology, 38*(2), 272-277.doi:http://dx.doi.org/10.1212/WNL.38.2.272.

Franchak, J.M., Kretch, K.S., Soska, K.C., & Adolph, K.E. (2011). Head-mounted eye tracking: A new method to describe infant looking. *Child Development, 82*(6), 1738-1750. doi:10.1111/j.1467-8624.2011.01670.x

Fried, M., Tsitsiashvili, E., Bonneh, Y.S., Sterkin, A., Wygnanski-Jaffe, T., Epstein, T., & Polat, U. (2014). ADHD subjects fail to suppress eye blinks and microsaccades while anticipating visual stimuli but recover with medication. *Vision Research, 101*, 62-72. doi:10.1016/j.visres.2014.05.004

Gooding, D.C., Miller, M.D., & Kwapil, T.R. (2000). Smooth pursuit eye tracking and visual fixation in psychosis-prone individuals. *Psychiatry Research, 93*(1), 41-54.

Gredebäck, G., Johnson, S., & Von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology, 35*(1), 1-19. doi:10.1080/87565640903325758

Hansen, D.W., & Ji, Q. (2010). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(3), 478-500. doi:10.1109/TPAMI.2009.30

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 University Press, London.

Holzman, P.S., Levy, D.L., & Proctor, L.R. (1976). Smooth pursuit eye movements, attention, and Schizophrenia. *Archives General Psychiatry, 33*(12), 1415-1420. doi:10.1001/archpsyc.1976.01770120019001

Holzman, P.S., & Levy, D.L. (1977). Smooth pursuit eye movements and functional psychoses: A review. *Schizophrenia Bulletin, 3*(1), 15-27. http://dx.doi.org/10.1093/schbul/3.1.15.

Hoormann, J., Jainta, S., & Jaschinski, W. (2008). The effect of calibration errors on the accuracy of the eye movement recordings. *Journal of Eye Movement Research, 1*(2):3, 1-7. doi:10.16910/jemr.1.2.3.

Huang, M.X., Kwok, T.C.K., Ngai, G., Chan, S.C.F., & Leong, H.V. (2016). *Building a personalized, auto-calibrating eye tracker from user interactions*. CHI 2016, San Jose, California, 7-12 May 2016. http://dx.doi.org/10.1145/2858036.2858404.

Hutton, S.B., & Tegally, D. (2005). The effects of dividing attention on smooth pursuit eye tracking. *Experimental Brain Research, 163*(3), 306-313. 10.1007/s00221-004-2171-z.

Iacono, W.G., & Lykken, D.T. (1979). Electro-oculographic recording and scoring of smooth pursuit and saccadic eye tracking: A parametric study using monozygotic twins. *Psychophysiology, 16*(2), 94-107. doi:10.1111/j.1469-8986.1979.tb01451.x.

Kerzel, D., Souto, D., & Ziegler, N.E. (2008). Effects of attention shifts to stationary objects during steady-state smooth pursuit eye movements. *Vision Research, 48*(7), 958-969. doi:10.1016/j.visres.2008.01.015.

Kliegl, R., & Olson, R.K. (1981). Reduction and calibration of eye monitor data. *Behavior Research Methods and Instrumentation, 13*(2), 107-111. doi:10.3758/bf03207917.

Kowler, E., & McKee, S. P. (1987). Sensitivity of smooth eye movement to small differences in target velocity. *Vision Research, 27*(6), 993-1015. https://doi.org/10.1016/0042-6989(87)90014-9.

Levin, S., Luebke, A., Zee, D.S., Hain, T.C., Robinson, D.A., & Holzman, P.S. (1988). Smooth pursuit eye movements in schizophrenics: Quantitative measurements with the search-coil technique. *Journal of Psychiatric Research, 22*(3), 195-206. https://doi.org/10.1016/0022-3956(88)90005-2.

Lindner, A., Schwarz, U., & Ilg, U.J. (2001). Cancellation of self-induced retinal image motion during smooth pursuit eye movements. *Vision Research, 41*(13), 1685-1694. https://doi.org/10.1016/S0042-6989(01)00050-5.

Madelain, L., Krauzlis, R.J., & Wallman, J. (2005). Spatial deployment of attention influences both saccadic and pursuit tracking. *Vision Research, 45*(20), 2685-2703. doi:10.1016/j.visres.2005.05.009.

McConkie, G.W. (1981). Evaluating and reporting data quality in eye movement research. *Behavior Research Methods & Instrumentation, 13*(2), 97-106. doi:10.3758/bf03207916

Meyer, C.H., Lasker, A.G., & Robinson, D.A. (1985). The upper limit of human smooth pursuit velocity. *Vision Research, 25*(4), 561-563. https://doi.org/10.1016/0042-6989(85)90160-9.

Nagel, M., Sprenger, A., Steinlechner, S., Binkofski, F., & Lencer, R. (2012). Altered velocity processing in schizophrenia during pursuit eye tracking. *PLoS One, 7*(6), e38494. doi:10.1371/journal.pone.0038494.

Nyström, M., Andersson, R., Holmqvist, K., & van de Weijer, J. (2013). The influence of calibration method and eye physiology on eyetracking data quality. *Behavior Research Methods, 45*(1), 272-288. doi:10.3758/s13428-012-0247-4

O'Driscoll, G.A., & Callahan, B.L. (2008). Smooth pursuit in Schizophrenia: A meta-analytic review of research since 1993. *Brain and Cognition, 68*(3), 359-370. doi:10.1016/j.bandc.2008.08.023

O'Driscoll, G.A., Lenzenweger, M.F., & Holzman, P.S. (1998). Antisaccades and smooth pursuit eye tracking and schizotypy. *Archives of General Psychiatry, 55*(9), 837-843.

Pfeuffer, K., Vidal, M., Turner, J., Bulling, A., & Gellersen, H. (2013). Pursuit calibration: Making gaze calibration less tedious and more flexible. P*roceedings of the 26th annual ACM symposium on user interface software and technology*, 261-270. http://dx.doi.org/10.1145/2501988.2501998.

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry, 68*(1), 101-109. doi:10.1001/archgenpsychiatry.2010.113

Pola, J., & Wyatt, H.J. (1997). Offset dynamics of human smooth pursuit eye movements: Effects of target presence and subject attention. *Vision Research, 37*(18), 2579-2595. doi:http://dx.doi.org/10.1016/S0042-6989(97)00058-8

Pola, J., & Wyatt, H.J. (2001). The role of target position in smooth pursuit deceleration and termination. *Vision Research, 41*(5), 655-669.

Richards, J.E., & Holley, F.B. (1999). Infant attention and the development of smooth pursuit tracking. *Developmental Psychology, 35*(3), 856-867.

Sasson, N.J., & Elison, J.T. (2012). Eye tracking young children with autism. *Journal of Visualized Experiments* (61), e3675. doi:10.3791/3675.

Sharpe, J.A. (2008). Neurophysiology and neuroanatomy of smooth pursuit: lesion studies. *Brain and Cognition, 68*(3), 241-254. doi:10.1016/j.bandc.2008.08.015

Sheena, D., & Borah, J. (1981). Compensation for some second order effects to improve eye position measurements. In D.F. Fisher, R.A. Monty & J.W. Senders (Eds.), *Eye Movements: Cognition and Visual Percep-tion*, Hillsdale: Lawrence Erlbaum Associates.

Simons, R.F., & Katkin, W. (1985). Smooth pursuit eye movements in subjects reporting physical anhedonia and perceptual aberrations. *Psychiatry Research, 14*(4), 275-289. https://doi.org/10.1016/0165-1781(85)90096-4.

Soechting, J.F., Rao, H.M., & Juveli, J.Z. (2010). Incorporating prediction in models for two-dimensional smooth pursuit. *PLoS One, 5*(9), e12574. doi:10.1371/journal.pone.0012574.

Souto, D., & Kerzel, D. (2011). Attentional constraints on target selection for smooth pursuit eye movements. *Vision Research, 51*(1), 13-20. doi:10.1016/j.visres.2010.09.017.

Spering, M., & Gegenfurtner, K.R. (2008). Contextual effects on motion perception and smooth pursuit eye movements. *Brain Research, 1225*, 76-85. doi:10.1016/j.brainres.2008.04.061

Swirski, L., & Dodgson, N. (2013). *A fully automatic, temporal approach to single camera, glint-free 3D eye model fitting.* European Conference of Eye Movements (ECEM), Lund, Sweden.

Takarae, Y., Minshew, N.J., Luna, B., Krisky, C.M., & Sweeney, J. A. (2004). Pursuit eye movement deficits in autism. *Brain, 127*(12), 2584-2594. doi:10.1093/brain/awh307.

Thier, P., & Ilg, U.J. (2005). The neural basis of smooth-pursuit eye movements. *Current Opinion in Neurobiology, 15*(6), 645-652. doi:10.1016/j.conb.2005.10.013.

Tobii Technology. (2010). *Product description, Tobii T/X series eye trackers, rev. 2.1.* June 2010. Tobii Technology AB.

Van Donkelaar, P., & Drew, A.S. (2002). The allocation of attention during smooth pursuit eye movements. *Progress in Brain Research, 140*, 267-277. doi:http://dx.doi.org/10.1016/S0079-6123(02)40056-8

Van Gelder, P., Lebedev, S., Liu, P.M., & Tsui, W.H. (1995). Anticipatory saccades in smooth pursuit: Task effects and pursuit vector after saccades. *Vision Research, 35*(5), 667-678. doi:http://dx.doi.org/10.1016/0042-6989(94)00161-E

Von Hofsten, C.V., & Rosander, K. (1997). Development of smooth pursuit tracking in young infants. *Vision Research, 37*(13), 1799-1810. http://dx.doi.org/10.1016/S0042-6989(96)00332-X.