Comparison of Three Gaze Position Calibration Techniques in First Purkinje-Image Based Eye Trackers

Michael Ntodie MSc, OD, [1] Shrikant R. Bharadwaj PhD, Bsopt.,[2,3] Swaathi Balaji Bsopt.,[2] Kathryn J. Saunders PhD, FCOptom,[1] Julie-Anne Little PhD, MCOptom[1]

[1]Optometry and Vision Science Research Group, Biomedical Sciences Research Institute, Ulster University, Coleraine, UK. [2] Prof. Brien Holden Eye Research Centre, Hyderabad Eye Research Foundation, L V Prasad Eye Institute, Hyderabad, India. [3] Bausch & Lomb School of Optometry, Hyderabad Eye Research Foundation, L V Prasad Eye Institute, Hyderabad, India.

*Corresponding address: G159, School of Biomedical Sciences, University of Ulster, Cromore Road, Coleraine Co., Londonderry, Northern Ireland, BT52 1SA, United Kingdom.

Corresponding author: Julie-Anne Little; ja.little@ulster.ac.uk

Commercial relationship: None

Number of figures: 6

Number of tables: 4

Word count in text: 5810

Financial interest: None

Significance: This study highlights potential differences that can arise in gaze position estimates from 1st Purkinje image-based eye trackers based on how individual Hirschberg ratios are calculated.

Purpose: To evaluate the accuracy and repeatability of eccentric-viewing, prism-based and theoretical techniques that are routinely used to calibrate Hirschberg Ratio (HR) in 1st Purkinje image-based eye trackers.

Methods: Hirschberg ratios of 28 UK participants (18-40 years) were obtained using the PlusOptix PowerRef3 photorefractor and eye tracker. In the gold-standard eccentric viewing technique, participants viewed eccentric targets (±12°, 4° steps) at 2m. In the prism-based technique, 4-16ΔD base-out and base-in prisms were placed in 4ΔD steps before an eye occluded with an infrared filter; the fellow eye fixated a target at 1m. Each participant's HR was calculated as the slope of the linear regression of the shift in Purkinje image, relative to the pupil centre, for each target eccentricity or induced prism power. Theoretical HR was calculated from the participant's corneal curvature and anterior chamber depth measures. Data collection was repeated on another visit using all three techniques to assess repeatability. Data were also obtained from an Indian cohort (n=30,18–40 years) using similar protocols.

Results: HR ranged from 10.61-14.63°/mm (median: 11.90°/mm) in the eccentric viewing technique. The prism-based and theoretical techniques demonstrated inaccuracies of 12% and 4% relative to the eccentric viewing technique. The 95% limits of agreement of intra-subject variability ranged from ±2.00°/mm, ±0.40°/mm, and ±0.30°/mm for the prism-based, eccentric viewing and theoretical techniques (P>.05). Intraclass correlation coefficients (95%CI) were 0.99 (0.98-1.00) for eccentric, 0.99 (0.99-1.00) for theoretical and 0.88 (0.74-0.94) for prism-based techniques. Similar results were found for the Indian cohort.

Conclusion: The prism-based and theoretical techniques both demonstrated relative inaccuracies in measures of Hirschberg ratio, compared to the eccentric viewing technique. The prism-based technique exhibited the poorest repeatability.

Keywords: Anterior chamber depth, corneal curvature, eye tracker, gaze position, Hirschberg ratio, Purkinje-image, repeatability, variability

Video-based eye trackers are increasingly used in research laboratories and clinical settings to obtain quantitative estimates of gaze position.[1-3] This is a valuable measure in studies investigating the biology of eye movements or those that use movements as an objective biomarker for assessing cognition[4, 5] or tracking disease progression.[6] These trackers use specific landmarks (e.g. 1st Purkinje image, limbus, pupil margin and blood vessel configuration) to rapidly and non-invasively track gaze position.[7, 8] Of the variety of eye trackers commercially available, those using the relative location of the 1st Purkinje image and entrance pupil center (i.e. virtual image of the anatomical pupil as seen through the cornea and anterior chamber[9]) as landmarks are of specific interest to the present study. A proportionality constant or calibration factor – the Hirschberg ratio is used in these trackers to directly convert the millimetre separation between the Purkinje image and pupil center into angular units of gaze position in degrees or prism dioptres (ΔD).[10, 11] Accuracy of gaze position estimates obtained by these trackers, relative to a gold-standard, depends on the accuracy of this calibration factor and on the angle kappa (i.e. the angle between the pupillary axis and visual axis).[12,3] Previous studies have reported a large inter-subject variability in Hirschberg ratio between 7 to 16°/mm or (12 to 28 ΔD/mm),[3, 9, 12, 13] and consequently, using a population-average Hirschberg ratio as the calibration factor (e.g. 11.8°/mm as used by the PlusOptix PowerRef3 device used in this study[12]) could result in inaccurate estimation of eye position. Using an individual's Hirschberg ratio as the calibration factor therefore enhances the accuracy of gaze position estimates determined by these eye trackers.[12]

Three techniques have been described in the literature to measure the individual Hirschberg ratio.[12-14] In the eccentric viewing technique, Hirschberg ratio is measured by asking participants to fixate on targets placed before them at known angular eccentricities and measuring the separation between the Purkinje image and pupil center for each of these positions using the eye tracker's high-resolution camera.[12] The reciprocal of the slope of a linear regression fit of the measured separation between Purkinje image and pupil center against the corresponding target eccentricity gives the Hirschberg ratio in degrees or prism dioptres per millimetre.[7, 11, 15] In the theoretical technique, Hirschberg ratio is derived from biometric measures of the eye's corneal curvature and anterior chamber depth using basic geometric formulae derived by Brodie.[7, 12] Finally, the prism-based technique involves the use of prism of known base powers to change the separation between Purkinje image and pupil center and derives Hirschberg ratio by calculating the reciprocal of the slope of a linear regression fit of the measured separation between these landmarks against a range of induced prism power.[14, 16] The fundamental characteristics, and practical advantages and disadvantages of all three techniques are presented in Table 1.

35 ---

36 Table 1 about here

37 ---

38

39   The utility of a given technique for measuring Hirschberg ratio in a clinical or research setting

40 depends on how successfully the technique can be implemented, how accurate the estimates of eye

41 positions are using this technique and how variable are the eye position estimates over repeat

42 measurements using this technique for the subject group. While some studies have investigated the

43 accuracy and repeatability of the eccentric viewing and the theoretical techniques and the agreement

44 between them,[3, 11, 12, 17] there is no such information available for the prism-based technique. There

45 are also no published data describing how the three techniques compare with each other. Therefore,

46 the overall aim of this study was to determine the accuracy, repeatability and agreement of Hirschberg

47 ratios obtained using each one of these three techniques.

48

49 **Materials and methods**

50   Adult subjects aged 18 – 40 years were recruited from staff and students of Ulster University

51 (UU), Coleraine, Northern Ireland, UK (n=28). All subjects that participated in the study had a

52 presenting distance visual acuity of 20/20 or better and were either emmetropic (-0.25D to +1.00D

53 spherical equivalent refractive error) or myopic (-0.50 to -6.00D spherical equivalent refractive error)

54 for the fixated eye (Table 2). Subjects were required to have refractive errors limited to <1.00D

55 anisometropia, and no more than 0.75DC of astigmatism. Subjects were corrected using soft contact

56 lenses as measurement of lens calibration was also being undertaken as a part of a different study

57 (Table 2). These soft contact lenses conformed to the shape of the cornea and produced minimal

58 movement during measurement. None of the subjects presented with any ocular pathology or

59 binocular vision anomaly. Subjects with amblyopia, pupil diameters <3.00mm (below the operational

60 range of the device) in a dimly lighted room, who could not maintain steady fixation or had an

61 excessive blink rate were excluded.

62

63   Data from a parallel study with small variations in experimental protocol and participant

64 cohort were collected from the Brien Holden Institute of Optometry and Vision Sciences, LV Prasad

65 Eye Institute (LVPEI), Hyderabad, India to determine the generality of the study outcomes in the UU

66 cohort. Table 2 summarises the participant characteristics in both cohorts. The study was conducted

67 in accordance with the tenets of the Declaration of Helsinki and commenced after approval by the

68 local Research Ethics Committee at UU and the Institutional Review Board at LVPEI.

69   Gaze position data were obtained using the PowerRef 3™ (Nuremberg, Germany) at a sampling

70   frequency of 50Hz for the prism-based calibration and the eccentric viewing calibration techniques.

71   The device is designed to capture dynamic measurements of the eye's refractive power at the pupil

72   plane using eccentric photorefraction technique.[18] To do so, the device is fitted with an array of

73   infrared light emitting diodes adjacent to the camera aperture, and this arrangement produces a crisp

74   1st Purkinje image, which can be compared to the position of the pupil center in the vertical and

75   horizontal meridian to determine respective gaze positions. Additional technical information about

76   the photorefractor used in this study can be found at https://plusoptix.com/images/support-

77   downloads/powerref3-specifications-usa.pdf. The use of different versions of eccentric

78   photorefractors for eye movement research has been previously reported,[12, 14, 16, 19] and more research

79   is being conducted with these instruments. Therefore, this study used the PowerRef 3™ as a model

80   instrument of 1stPurkinje image-based eye trackers for investigating the repeatability of three gaze

81   position calibration techniques, using eccentric viewing technique as the reference gold-standard. The

82   set-up for these techniques is described below for the UU data collection. The LVPEI data collection is

83   broadly similar, but Table 2 lists the variations between the two study sites.

84

85   ---

86   Figure 1 about here

87   ---

88

89   <u>Eccentric viewing technique</u>

90   In the eccentric viewing technique, subjects fixated with their left eye on a series of Maltese

91   cross targets arranged linearly at 2m while the right eye was occluded. The entire target array, from

92   left to right, produced eccentricity of ±12°, with each target separated from the adjacent by 4°. These

93   target eccentricities were all within the ±25°range described by Brodie, where the millimetre

94   separation between the Purkinje image and pupil center varies linearly with angular eccentricity.[13]

95   Beyond this range, the separation between these two landmarks varies in a sinusoidal manner with

96   angular eccentricity.[13]

97

98   <u>Prism-based technique</u>

99   In the prism-based calibration technique, the dominant eye fixated on a Maltese cross target

100 at 1m while the fellow non-dominant eye was occluded using an Optcast long pass infrared

101 transmitting filter (Edmund Optics™, NT43-954). This filter blocks virtually all the visible wavelengths

102 while allowing infrared light from the camera to pass through it. Prisms from 0ΔD to16ΔD in 4ΔD steps

5

103 were held before the occluded eye in a base-in and base-out sequence, at a vertex distance of 10–14

104 mm for at least four seconds each. This experimental set-up helped to eliminate any compensatory

105 vergence eye movements from the fellow eye that might contaminate the eye position calibration

106 results. Prism powers were converted into degrees using the formula: where degrees = arctan (prism

107 dioptres/100). The prism powers used here thus corresponded to target eccentricities of ±9.09°, in

108 steps of 2.29°.

109

110 --

111 Table 2 about here

112 --

113

114   Gaze position calibration measurements in both prism-based and eccentric viewing

115 techniques were made in a dimly lit room. Data analysis was performed using custom written software

116 in Matlab®. Raw PowerRef 3™ data were processed, removing blinks and extraneous data outside of

117 the instrument working range. Eye movement data were plotted against time and scrutinised for a

118 section of stable gaze, and two seconds worth of data (~100 samples) from each prism power or target

119 position was selected and averaged.[14, 16] These averaged gaze positions were plotted against the

120 corresponding prism power and target eccentricity in prism-based and eccentric viewing techniques,

121 respectively. Linear regression analysis was performed to obtain the eye position calibration slope.

122 The slope of this linear regression equation provided an estimate of the subject's Hirschberg ratio. The

123 calibration slope obtained from these two techniques is a unitless quantity describing the change in

124 eye position recorded by the PowerRef 3™ for a unit change in target eccentricity or prism power. The

125 actual Hirschberg ratio of the individual will be equal to that used by the PowerRef 3™ divided by the

126 calibration slope of that individual obtained using these techniques. In other words, an eye position

127 calibration slope that is equal to unity indicates an Hirschberg ratio of 11.8°/mm (i.e. equal to the

128 population average value used by the machine). Eye position calibration slopes greater than unity

129 correspond to Hirschberg ratios <11.8°/mm (smaller than population average value) while calibration

130 slopes smaller than unity correspond to Hirschberg ratios >11.8°/mm (larger than population average

131 value).

132

133 <u>Theoretical technique</u>

134   The theoretical calibration technique was based on the geometric optics model described by

135 Brodie, which posits that the Hirschberg ratio of the individual varies with their anterior corneal

136 curvature and the anterior chamber depth.[7] The procedure of obtaining Hirschberg ratio using the

137   theoretical technique is explained in detail by Jagini et al.[12] Briefly, the average of three measurements

138   of horizontal corneal curvature and anterior chamber depth of each subject was obtained using the

139   Zeiss IOL Master™. These values were used to theoretically predict the Hirschberg ratio using the

140   regression equation previously described by Jagini et al.[12] All theoretical calculations were performed

141   using custom software in Matlab[®]. Data from the left eye are presented for the theoretical technique

142   to allow for comparison with the other techniques.

143

144   Repeatability of Hirschberg ratio obtained with the three techniques was assessed by repeating each

145   technique for a second time within a week of the first measurement.

146

147   **Data analysis**

148   Representative eye position raw data and the linear regression fits that were derived from

149   these data for the eccentric viewing and the prism-based techniques are shown in Figure 1. There

150   were excellent linear regression fits for both prism-based and eccentric viewing calibration techniques

151   with r^2 value ≥0.90 in all subjects (consistent with previous work), and this was a criterion for inclusion

152   of data.[12] The Kolmogorov-Smirnov test indicated that the corneal curvature, anterior chamber depth

153   and Hirschberg ratio calculated from the slopes were not normally distributed and, therefore, non-

154   parametric statistics described the data. One-way ANOVA was used to assess the mean difference

155   between the three techniques and the post-hoc Scheffe test was used to determine differences

156   between the individual groups while the paired t-test was used to assess mean differences between

157   baseline and repeat measures. The intraclass correlation coefficient (ICC, two-way mixed effects) was

158   used to assess absolute agreement between measures. Results of intraclass correlation test were

159   considered excellent, good, moderate, and poor if >0.90, 0.75 to 0.90, 0.50 to 0.75, and <0.50,

160   respectively.[20] Linear regression analysis was used to determine range effects in Hirschberg ratio.

161   Significance was determined as P<.05.

162

163   **Results**

164   Data collection was successful in all 28 subjects recruited at UU and in all 30 subjects recruited

165   at LVPEI. The data from the UU cohort will be described first followed by a description of the data in

166   the LVPEI cohort. Given the differences in experimental protocols employed in the two study

167   locations, albeit minor, the two datasets are considered as stand-alone entities and no explicit

168   comparison is made between the two datasets. Only general judgments about the accuracy and

169   repeatability of the Hirschberg ratio measurements from the three protocols are made from the data

170   obtained from the two study locations.

Data from the UU cohort

172   The baseline regression slopes ranged from 0.81 to 1.11 [median (25th-75thinterquartile range;

173   IQR) =0.99 (0.91–1.03)] for the eccentric viewing technique, and 0.70 to 1.03 [0.89 (0.79-0.93)] for the

174   prism-based technique (Table 3). The Hirschberg ratios calculated from these slopes ranged from

175   10.61 to 14.63°/mm [11.90°/mm (11.44–12.97°/mm)] for the eccentric viewing technique and 11.47

176   to 16.93 °/mm [13.30°/mm (12.74–15.06°/mm)] for the prism-based technique Table 3 and Figure 2A).

177   Baseline corneal curvatures and anterior chamber depth results are presented in Figure 2B. These

178   translated into theoretically derived Hirschberg ratios ranging from 9.84 to 13.44°/mm [11.43°/mm

179   (10.55–11.96°/mm)] (Table 3).

180

181   ---

182   Table 3 and Figure 2 about here

183   ---

184

185   Results of one-way ANOVA test showed an overall statistically significant difference (signed

186   difference) between the mean Hirschberg ratio of all three techniques ($F_{(2,81)}$ = 31.24, P<.0001). Post-

187   hoc test using Scheffe method showed statistically significant differences between the mean

188   Hirschberg ratio of the prism-based and eccentric viewing techniques (-1.63 °/mm, P<.001), between

189   the prism-based and theoretical techniques (-2.54 °/mm, P<.001) and between the eccentric viewing

190   and theoretical techniques (-0.91 °/mm, P=.024), Table 4 (signed difference). However, to determine

191   if the difference between the techniques varied as a function of calculated Hirschberg ratios, the mean

192   absolute difference between the Hirschberg ratios was computed and presented with the mean

193   signed difference (Table 4). The Bland-Altman type plots of the absolute difference between

194   techniques indicated that there was systematic bias in the Hirschberg ratios obtained from one

195   technique, relative to the other (Figures 3A – C). Linear regression analyses were performed on the

196   absolute difference to determine the effect of Hirschberg ratio size on the mean difference. Results

197   of this analysis indicated that the slopes were significantly different from zero between the prism-

198   based and eccentric viewing techniques (Linear regression equation: Y= -4.18 + 0.46X (X= inter-

199   technique average Hirschberg ratio): $F_{(1,26)}$=7.52, P=.01), and between the prism-based and theoretical

200   techniques (Linear regression equation: Y= -5.51 + 0.64X: $F_{(1,26)}$ =6.98, P=.01, (Figures 3A and B show

201   this effect). However, there was no such Hirschberg ratio range effects on the mean difference

202   between the eccentric and theoretical techniques (Linear regression equation: Y= -0.80 + 0.15X: $F_{(1,26)}$

203   = 0.34, P=.57), (Figure 3C).

204   To determine the intra-subject variability of the three Hirschberg ratio techniques, each

205   technique was repeated for a second time within a week of the baseline measurement reported

206   (Figure 4A – C). The mean difference (95% LOA) in Hirschberg ratio between the first and second

207   measurements were 0.05°/mm (95% LOA: -0.30 to 0.40°/mm) for the eccentric viewing technique

208   (Figure 4A) and 0.09°/mm (95% LOA: -1.91 to 2.08°/mm) for the prism-based technique (Figure 4B).

209   Median repeat corneal curvature and anterior chamber depth measures for calculating Hirschberg

210   ratio using the theoretical technique was 7.91 (7.80 – 8.29mm) and 3.54 (3.28 – 3.74), respectively.

211   These translated into median Hirschberg ratio of 11.45°/mm (10.45 – 11.93°/mm). The mean intra-

212   subject variability of the theoretically derived Hirschberg ratio was therefore 0.04°/mm [paired t-test

213   (95% LOA: -0.20 to 0.28°/mm)] (Figure 4C). The mean difference between the first and repeat

214   measures of Hirschberg ratio for all three techniques was not significantly different (all P>.05).

215   Moreover, results of intraclass correlation test revealed excellent repeatability in the eccentric

216   viewing and theoretical techniques [0.99(95% CI: 0.98-0.997), P<.001, and 0.99 (95% CI: 0.99-0.998),

217   P<.001, for the eccentric viewing and theoretical techniques respectively]. In the prism-based

218   technique, there was a good agreement between the first and repeat measures 0.88(95% CI: 0.74-

219   0.944), P<.001) (Table 3).

220

221   --

222   Figures 3 and 4 about here

223   --

224

225   <u>Data from the LVPEI cohort</u>

226   The baseline regression slopes ranged from 0.79 to 1.24 [median (25th-75thIQR) = 1.04 (0.98–

227   1.09)] and 0.76 to 1.12 [0.97 (0.91–1.03)] in the eccentric viewing and prism-based techniques,

228   respectively (Table 3). The Hirschberg ratios derived from these regression slopes are shown in Table

229   3 and Figure 5A. The baseline corneal curvatures and anterior chamber depth results are presented in

230   Figure 5B and the theoretically derived Hirschberg ratios from these values are shown in Table 3.

231   Results of one-way ANOVA test showed an overall statistically significant difference (signed

232   difference) between the mean Hirschberg ratio of all three techniques ($F_{(2,87)}$ = 4.25, P=.01). Post-hoc

233   test using Scheffe method showed significant differences between the mean Hirschberg ratio of the

234   prism-based and eccentric viewing techniques (-0.77 °/mm, P=.03). However, there was no statistically

235   significant difference between the prism-based and theoretical techniques (-0.67 °/mm, P=.06) and

236   between the eccentric viewing and theoretical techniques (0.10 °/mm, P=.93). As before, results of

237   the absolute difference are presented in Bland Altman plots in Figure 6A-C, see also Table 4. There

238   were no Hirschberg ratio range effects on the mean absolute difference between the prism-based and

239   eccentric viewing techniques (Linear regression equation: Y= -2.13 + 0.27X (X= inter-technique average

240   Hirschberg ratio): $F_{(1,28)}$ =2.79, P=.11), and between the prism-based and theoretical techniques (Linear

241   regression equation: Y= -1.52 + 0.20X: $F_{(1,28)}$ = 3.71, P=.06) (Figure 6A&B). However, range effect was

242   observed on the mean difference between the eccentric viewing and theoretical techniques (Linear

243   regression equation: Y= -2.14 + 0.23X: $F_{(1,28)}$ = 4.62, P=.04) (Figure 6C).

244

245   Repeatability of the Hirschberg ratio estimate was available only for the prism-based

246   technique in the LVPEI cohort and this data showed a mean difference (95% LOA) in Hirschberg ratioof

247   0.08°/mm (95% LOA: -1.7 to 1.9°/mm) (P=.63) (Figure 6D), and the intraclass correlation test of

248   absolute agreement between the first and second measures showed good agreement between the

249   two [0.83(95%CI: 0.64-0.92), P<.001] (see Table 3).

250

251   ---

252   Figure 6 about here

253   ---

254

255   **Discussion**

256   The essence of calibrating an individual's Hirschberg ratio in 1st Purkinje image-based eye

257   trackers is to reduce the errors in gaze position estimates that may arise while using the population-

258   average Hirschberg ratio.[3, 9, 12] This problem is of real concern to commonly used eye trackers given

259   the large inter-subject variability in Hirschberg ratio that has been reported in the literature.[3, 9, 11, 12]

260   Following the decision to use the subject's own Hirschberg ratio to calibrate the eye tracker for

261   improved accuracy, a second challenge is to determine which calibration technique is to be adopted

262   for this purpose. The performances of three such techniques that have been used previously in the

263   literature –eccentric viewing, prism-based and theoretical – were tested in the present study. To

264   determine the accuracy of a given calibration technique, the values obtained by this technique need

265   to be compared against a "gold-standard" measure. For the present analysis, the eccentric viewing

266   technique is considered as the "gold-standard" technique simply because of its traditional use for

267   calibrating the Hirschberg ratio in most 1st PI-based eye trackers.[3, 21] This technique has also become

268   a "legacy technique" from which the population-average Hirschberg ratio has been derived in previous

269   studies.[3, 21] Moreover, this technique uses angles anchored in space, and it is based on actual eye

270   rotations, thus requires few assumptions to be made for deriving the Hirschberg ratio.

Compared to the eccentric viewing technique, the prism-based and theoretical techniques both demonstrated relative inaccuracies of 12% and 4% respectively in the UU cohort when the median values were compared [see Table 3 for median values: (100- (13.30÷11.90) ×100) = 12%), and (100-(11.43÷11.90) ×100) = 4%)]. In the LVPEI cohort, similar inaccuracies of 7% and 3% were recorded in the prism-based and theoretical techniques respectively. At individual level, these inaccuracies ranged from 6% of underestimation of the Hirschberg ratio to 33% of overestimation in the theoretical technique, and 20% underestimation to 37% overestimation in the prism-based technique in the UU cohort; and 19% underestimation to 13% overestimation, and 23% underestimation to 29% overestimation in the theoretical and prism-based techniques respectively in the LVPEI cohorts.

The present study also demonstrated the over-estimation of Hirschberg ratio by the prism-based technique, relative to both the eccentric-viewing and theoretical techniques (Figure 3A and B). This was particularly significant in the UU cohort, although there were range effects in the bias towards the prism-based technique and with the difference between techniques appearing to increase with an increase in the size of the Hirschberg ratio (Figure 3A and B). However, there was no such range effect between the eccentric viewing and theoretical techniques, and the mean difference between the two was closer to zero compared to the mean differences when the prism-based technique is considered.

In addition to accuracy, the calibration technique's repeatability also needs to be assessed to determine its usefulness in estimating the individual Hirschberg ratio. The theoretical technique demonstrated the least intra-subject variability as the Hirschberg ratio obtained with this technique was repeatable to within ±0.30°/mm 95% LOA in a subject, less than the 0.50°/mm 95% LOA previously reported by Jagini et al.[12] This may be attributed to the more consistent, repeatable measures of corneal curvature and anterior chamber depth (~0.01mm for both measures in this study, less than the 0.08mm reported previously[12, 22, 23]) available in the present study. The Hirschberg ratio obtained using the eccentric viewing technique was repeatable to within ±0.40°/mm 95% LOA in individual subjects; slightly less repeatable than the theoretical technique, but more repeatable than the prism-based technique. Furthermore, there was improved intra-subject repeatability in the eccentric viewing technique in this study than previously reported (1.5 to 3.0 degrees/mm).[21, 24] The use of different fixation targets in the present study compared to previous work could explain the differences in the intra-subject repeatability reported as fixation target characteristics are known to affect the stability of eye movements.[25] Perhaps the use of Maltese cross fixation target in this study minimized micro-eye movements, thereby contributing to enhanced repeatability. The highest intra-subject variability in Hirschberg ratio, with the lowest intraclass correlation coefficient was observed in the prism-based technique. The Hirschberg ratio measured in the prism-based technique was repeatable

305 to within ±2.0°/mm at UU and ±1.9°/mm at LVPEI. There are no previously published data with which

306 to compare these findings, but when compared with the other two techniques in this study, the

307 variability exhibited by the prism-based technique is high (Figure 4, panel B and Figure 6, panel D).

308 Moreover, the lowest intraclass correlation coefficient was recorded in the prism-based technique

309 demonstrating least agreement between the baseline and repeat measures. It is possible that the high

310 variability in Hirschberg ratio exhibited in the prism-based technique is inherent when using prisms

311 for calibration. Variability in Purkinje image displacements which can arise from minimal variance in

312 orientation and/or placement of the prisms before the infrared-occluded eye during repeat

313 measurements will influence the results.[26] Furthermore, variability in a subject's phoria adaptation at

314 different measurement times,[27] could lead to the high variability observed with the prism-based

315 technique. Finally, potential conflicts in fixation between the target presented to the non-occluded

316 eye and the image of the infrared LED's in the occluded eye could lead to additional variability in this

317 technique.

318

319   Another way to quantify the precision of a technique, is to compare the intra-subject and

320 inter-subject variability of the technique. If the magnitude of intra-subject variability equals the inter-

321 subject variability produced by the technique, then its usefulness for calibration could be questioned.

322 In the case of theoretical technique, the intra-subject variability was 13% relative to the inter-subject

323 variability [see Table 3 UU section, for a -0.20 to 0.28°/mm of intra-subject variability, expressed as

324 percentage of its inter-subject variability (9.84 to 13.44°/mm): (0.48÷3.6 × 100) =13%]. Similarly, the

325 eccentric viewing technique exhibited 17% variability of the inter-subject value. However, the prism-

326 based technique exhibited 73% variability in both cohorts relative to the inter-subject variability [e.g.

327 for a -1.91 to 2.08°/mm in the UU cohort, expressed as percentage of its inter-subject variability (11.47

328 to 16.93°/mm): (3.99÷ 5.46 × 100) = 73%]. From these data, it is evident that the prism-based

329 technique exhibited the greatest variability relative to the other two.

330

331   In the present study, data from two laboratories that evaluated three gaze position calibration

332 techniques using similar (but not identical) protocols are presented together in one publication. The

333 differences in protocols include: 1) the assessment of agreement between the eccentric viewing and

334 prism-based techniques in the UU data is limited by the fact that the non-dominant eye was used in

335 the prism-based technique, whereas the left eye was used in eccentric viewing technique (Table 2).

336 However, subjects were all binocularly normal, and wore optical correction during assessments, which

337 would have minimized any possible effects. Additionally, there was a high intraclass correlation

338 between the two ocular biometric measures of the two eyes [0.99(95% CI: 0.98-0.99), P<.001 and

339 0.99(0.97-0.99), P<.001] for the corneal curvature and anterior chamber depth, respectively. With this

340 very high level of agreement, the Hirschberg ratio of a subject would not have been significantly

341 different regardless of which eye was used in the prism-based technique; 2) subjects in the LVPEI

342 cohort includes myopes who were not optically corrected during the data collection process while

343 those at UU were corrected with soft contact lenses (Table 2). While this could affect accurate fixation

344 of targets in theory, it has been reported that blur from moderate levels of uncorrected refractive

345 errors has minimal impact on fixation accuracy, and rather that fixation target characteristics are

346 critical in determining accuracy and stability[28, 29]; 3) the two study centers used different technologies

347 to obtain measures of corneal curvature and anterior chamber depth in subjects (Table 2). Although

348 the individual instruments have been shown to produce accurate and repeatable measures of the two

349 ocular biometric parameters,[22, 30-32] we are unable report on the agreement between these

350 instruments, even though they produce overlapping results. However, despite these differences in

351 protocols, there was a general similarity in the results in that the corneal curvature and anterior

352 chamber depth results were similar in the two cohorts, the prism-based technique demonstrated the

353 least repeatability and the theoretical technique demonstrated the best repeatability amongst the

354 three protocols tested. These indicate that results obtained in this study are not limited to the specific

355 protocol being followed but it reflects a more general trend of one calibration protocol being of lesser

356 utility than others in obtaining accurate and repeatable estimates of gaze position using 1stPurkinje

357 image-based eye trackers.

358

359   In conclusion, the study demonstrates the existence of inter-and intra-subject variance of the

360 Hirschberg ratio in all three methods employed to convert the millimetre separation between the 1st

361 Purkinje image and pupil center into angular units. Consequently, using the population-average

362 Hirschberg ratio may lead to inaccurate estimates of gaze position as shown in the present study. The

363 prism-based and theoretical techniques both demonstrated relative inaccuracies to the eccentric

364 viewing technique. However, the prism-based technique showed the poorest repeatability. In

365 comparison, the theoretical and eccentric viewing techniques demonstrated better repeatability of

366 Hirschberg ratio.

367

368 **Acknowledgements**

375   References

376 1. Bedell HE, Stevenson SB. Eye movement testing in clinical examination. Vis Res 2013;90:32-7.

377 2. Young LR, Sheena D. Survey of eye-movement recording methods. Behav Res Methods
378 1975;7(6):397-429.

379 3. Schaeffel F. Kappa and hirschberg ratio measured with an automated video gaze tracker. Optom
380 Vis Sci 2002;79(5):329-34.

381 4. Raney GE, Campbell SJ, Bovee JC. Using eye movements to evaluate the cognitive processes
382 involved in text comprehension. J Vis Exp 2014(83):e50780.

383 5. Eckstein MK, Guerra-Carrillo B, Miller Singley AT, Bunge SA. Beyond eye gaze: What else can
384 eyetracking reveal about cognition and cognitive development? Dev Cogn Neurosci 2017;25:69-91.

385 6. Poletti B, Carelli L, Solca F, Lafronza A, Pedroli E, Faini A, Zago S, Ticozzi N, Ciammola A, Morelli C,
386 Meriggi P, Cipresso P, Lule D, Ludolph AC, Riva G, Silani V. An eye-tracking controlled
387 neuropsychological battery for cognitive assessment in neurological diseases. Neurol Sci
388 2017;38(4):595-603.

389 7. Brodie S. Photographic calibration of the hirschberg test. Invest Ophthalmol Vis Sci
390 1987;28(4):736-42.

391 8. Barry J, Backes A. Limbus versus pupil center for ocular alignment measurement with corneal
392 reflexes. Invest Ophthalmol Vis Sci 1997;38(12):2597-607.

393   9. Riddell PM, Hainline L, Abramov I. Calibration of the hirschberg test in human infants. Invest

394   Ophthalmol Vis Sci 1994;35(2):538-43.

395   10. Carter AJ, Roth N. Axial length and the hirschberg test. Am J Optom Physiol Opt June

396   1978;55:361-64.

397   11. Model D, Eizenman M, Sturm V. Fixation-free assessment of the hirschberg ratio. Invest

398   Ophthalmol Vis Sci 2010;51(8):4035-039.

399   12. Jagini KK, Vaidyanath H, Bharadwaj SR. Utility of theoretical hirschberg ratio for gaze position

400   calibration. Optom Vis Sci 2014;91(7):778-85.

401   13. Brodie S. Corneal topography and the hirschberg test. Appl Opt 1992;31(19):3627-631.

402   14. Bharadwaj SR, Candy TR. Cues for the control of ocular accommodation and vergence during

403   postnatal human development. J Vis 2008;8(16):14-6.

404   15. Eskridge J, Wick B, Perrigin D. The hirschberg test - a double-masked clinical-evaluation. Am J

405   Optom Physiol Opt 1988;65(9):745-50.

406   16. Bharadwaj SR, Candy TR. Accommodative and vergence responses to conflicting blur and

407   disparity stimuli during development. J Vis 2009;9(11):1-18.

408   17. Eskridge J, Perrigin D, Leach N. The hirschberg test - correlation with corneal radius and axial

409   length. Optom Vis Sci 1990;67(4):243-47.

410   18. Roorda A, Campbell MCW, Bobier WR. Slope-based eccentric photorefraction: Theoretical

411   analysis of different light source configurations and effects of ocular aberrations. J Opt Soc Am A Opt

412   Image Sci Vis 1997;14(10):2547-556.

413   19. Horwood AM, Riddell PM. The use of cues to convergence and accommodation in naive,

414   uninstructed participants. Vision Res 2008;48(15):1613-24.

415   20. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for

416   reliability research (vol 15, pg 155, 2016). Journal of Chiropractic Medicine 2017;16(4):346-.

417   21. Hasebe S, Ohtsuki H, Tadokoro Y, Okano M, Furuse T. The reliability of a video-enhanced

418   hirschberg test under clinical conditions. Invest Ophthalmol Vis Sci 1995;36(13):2678-685.

419   22. Lackner B, Schmidinger G, Skorpik C. Validity and repeatability of anterior chamber depth

420   measurements with pentacam and orbscan. Optom Vis Sci 2005;82(9):858-61.

421   23. Crawford AZ, Patel DV, McGhee CNJ. Comparison and repeatability of keratometric and corneal

422   power measurements obtained by orbscan II, pentacam, and galilei corneal tomography systems.

423   Am J Ophthalmol 2013;156(1):53-60.

424   24. Quick M, Boothe R. A photographic technique for measuring horizontal and vertical eye

425   alignment throughout the field of gaze. Invest Ophthalmol Vis Sci 1992;33(1):234-46.

426   25. Thaler L, Schuetz AC, Goodale MA, Gegenfurtner KR. What is the best fixation target? the effect

427   of target shape on stability of fixational eye movements. Vision Res 2013;76:31-42.

428   26. Thompson JT, Guyton DL. Ophthalmic prisms. Ophthalmology 1983;90(3):204-10.

429   27. Toole AJ, Fogt N. The forced vergence cover test and phoria adaptation. Ophthalmic Physiol Opt

430   2007;27(5):461-72.

431   28. Ukwade M, Bedell H. Stability of oculomotor fixation as a function of target contrast and blur.

432   Optom Vis Sci 1993;70(2):123-26.

433   29. Steinman R, Pizlo Z, Forofonova T, Epelboim J. One fixates accurately in order to see clearly not

434   because one sees clearly. Spat Vis 2003;16(3-4):225-41.

435   30. Lam A, Chan R, Pang P. The repeatability and accuracy of axial length and anterior chamber

436   depth measurements from the IOLMaster((TM)). Ophthalmic Physiol Opt 2001;21(6):477-83.

437   31. Santodomingo-Rubido J, Mallen E, Gilmartin B, Wolffsohn J. A new non-contact optical device for

438   ocular biometry. Br J Ophthalmol 2002;86(4):458-62.

439   32. Wang Q, Savini G, Hoffer KJ, Xu Z, Feng Y, Wen D, Hua Y, Yang F, Pan C, Huang J. A

440   comprehensive assessment of the precision and agreement of anterior corneal power

441   measurements obtained using 8 different devices. Plos One 2012;7(9):e45607.

442

443

Figure legends

Figure 1: Raw data of eye position recorded by the PowerRef 3™ plotted as a function of time in the in eccentric viewing calibration (panel A) and prism-based calibration techniques (panel B) for one representative emmetropic subject. Positive and negative values in the y-axis indicate leftward and rightward gaze rotations, respectively, in the eccentric viewing calibration technique. Positive and negative values in the y-axis indicate the effect of base-out and base-in prisms, respectively, in the prism-based calibration technique. The right eye was occluded in the eccentric viewing calibration technique and therefore the instrument did not record any data in this eye (panel A). The right eye fixated on a distant target while prisms were placed before the left eye that was occluded using an infrared transmitting filter in the prism-based calibration technique (panel B). Mean (95% CI) of the eye position data recorded by the PowerRef 3™ plotted as a function of eccentric stimulus position (panel C) and as a function of prism power (panel D) for the same representative subject. The solid line through the data represents the best-fit linear regression equation while the dashed line indicates the unity line.

Figure 2: Box and Whisker plots of the baseline Hirschberg ratios obtained using the eccentric viewing, prism-based, and theoretical techniques (panel A) and the anterior chamber biometric properties of the eye (panel B) for calculating the Hirschberg ratio using the theoretical technique in the UU cohort. The solid horizontal line within the box indicates median value, lower and upper edges of the box indicate the 25th and 75th interquartile range and lower and upper whiskers show the 1st and 99th quartiles. The squares represent individual data points.

Figure 3: Bland-Altman type plots show the agreement between the Hirschberg ratios obtained using the three calibration techniques in the UU cohort. Panel A shows the agreement between the prism-based and eccentric viewing techniques, panel B shows the agreement between the prism-based and theoretical techniques and panel C shows the agreement between the eccentric viewing and theoretical techniques. The solid black lines in all panels indicate the mean absolute difference between the two measurements while the dashed black lines indicate the 95% limit of agreement. The mean difference (MD) and the limits of agreement (LOA) obtained for each comparison is noted in the figure panel.

Figure 4: Bland-Altman type plots of repeatability of three calibration techniques in the UU cohort. Panel A shows repeatability of the eccentric viewing technique, panel B shows repeatability of the prism-based technique and panel C shows repeatability of the theoretical technique. The solid black

478   lines in all panels indicate the mean difference between the two measurements while the dashed

479   black lines indicate the 95% limit of agreement. The mean difference and the limits of agreement

480   obtained for each comparison is noted in the figure panel.

481

482   **Figure 5:** Box and Whisker plots of the baseline Hirschberg ratios obtained using the eccentric viewing,

483   prism-based, and theoretical techniques (panel A) and the anterior chamber biometric properties of

484   the eye (panel B) for calculating the Hirschberg ratio using the theoretical technique in the LVPEI

485   cohort. The solid horizontal line within the box indicates median value, lower and upper edges of the

486   box indicate the 25th and 75th interquartile range and lower and upper whiskers show the 1st and

487   99th quartiles. The squares represent individual data points.

488

489   **Figure 6**: Bland-Altman type plots show the agreement between the Hirschberg ratios obtained using

490   the three calibration techniques (panels A – C) and the repeatability of the prism-based technique

491   (panel D) in the LVPEI cohort. The solid black lines in all panels indicate the mean absolute difference

492   between the two measurements while the dashed black lines indicate the 95% limit of agreement.

493   The mean difference and the limits of agreement obtained for each comparison is noted in the figure

494   panel.

495

496 **Table 1:** Characteristics, advantages, and disadvantages of the three gaze position calibration techniques.

497 PI – 1st Purkinje image, PC – Centre of entrance pupil, AC- Anterior Chamber and HR – Hirschberg Ratio.

Technique	Characteristics	Advantages	Disadvantages
Eccentric-Viewing	• Subject fixates on targets placed at known angular eccentricities. • Separation between PI and PC for each target eccentricity is measured. • Reciprocal of slope of linear regression fit of measured separation between PI and PC against target eccentricity gives HR. • Standard calibration routine in most eye trackers.	• Easy to perform in adults and healthy subjects. • Requires minimal technology.	• Assumes that subject is fixating accurately at the expected target location. • Unsteady head position can affect measurement. • Resistance to monocular occlusion in some subjects can make data collection difficult. • Data acquisition can be difficult in uncooperative subjects like infants and children.
Prism-based	• Involves the use of prisms of known base powers to create a separation between PI and PC while one eye is occluded with IR filter • A reciprocal of the slope of the linear regression fit of the separation between the PI and PC against prism power gives the HR.	• Requires minimal technology (loose prisms in a trial case can be used). • Requires minimal participation from subject. • Can be used in infants and children.	• Can be time consuming (e.g. if reflections are present during measurements). • Resistance to monocular occlusion in some subjects can make data collection difficult. • Chance of binocular fusion if monocular occlusion technique is inappropriate.
Theoretical	• HR is derived from anterior chamber biometry of the eye (i.e. corneal curvature and AC depth). • Corneal curvature and AC depth converted into HR using a formula described by Brodie.[7]	• HR can be obtained more quickly than other two techniques. • Less reliant on participant's cooperation. • Less reliant on gaze changes	• Dependency on the availability of technology for biometric measures. • Accuracy of HR estimates depends on accuracy and repeatability of the biometric device.

498

499 **Table 2:** Subject characteristics, experimental set-up and data collection protocols used at the
500 University of Ulster (UU) and at the L V Prasad Eye Institute (LVPEI).

Subject details	UU	LVPEI
Sample size	28	30
Age	18 - 40	18 – 40
Refractive error details		
Overall	†0.44 D (25th -75th IQR: -0.13-0.88D)	0.00 D (25th-75th IQR: -2.25 – 0.00D)
i. Myopia	CL corrected, n=9	Uncorrected, n=10
ii. Emmetropia	n=19	n=20
Experimental set up		
Eccentric viewing		
i. Target used	6Maltese crosses	6 LEDs
ii. Viewing distance	2m	3m
iii. Visual angle subtended	± 12° (in 4° steps)	± 15° (in 5° steps)
iv. Fixating eye	Left eye	Left eye
v. Illumination	Dim	Dim
Prism-based		
i. Prism range used	4 – 16ΔD	4 – 25ΔD
ii. Vertex distance	10 – 14mm	10 – 14 mm
iii. IR filter used	Optcast Filter NT3-953	Optcast Filter NT3-953
iv. Fixating eye	Dominant eye	Left eye
Theoretical		
CC and AC depth Instrument used	Zeiss IOL Master™	Wavelight® Oculyzer™ II diagnostic device
Data collection protocol used		
Eccentric-viewing		
i. Intra-subject repeatability	Yes (n=28)	No
Prism-based		
i. Intra-subject repeatability	Yes (n=28)	Yes (n=30)
Theoretical Intra-subject repeatability	Yes (n=28)	No

501 † Median refractive error for all participants

Table 3: Repeat measures of median regression slopes, and Hirschberg ratios (HRs) (full range) for three calibration techniques and Intraclass Correlation Coefficient (ICC) test of agreement between baseline and repeat measures. Baseline values represent first visit measurements. Intra-subject variability in each technique was calculated from the MEAN difference (95% Limit of agreement) between the baseline and repeat measurements.

	Regression slope (unitless)	Median (full range) HR (°/mm)	Intra-subject variability Mean difference (95% limits of agreement)	Intraclass correlation Coefficient (ICC) rho (95% CI)
UU				
Eccentric viewing				
Baseline	0.99	11.90 (10.61 – 14.63)	0.05 (-0.30 to 0.40)	0.99 (0.98-0.997)
Repeat measurement	0.99	11.94 (10.28 – 14.29)	(t=1.63, P=.12)	P<.001
Prism-based				
Baseline	0.89	13.30 (11.47 – 16.93)	0.09 (-1.91 to 2.08)	0.88 (0.74-0.944)
Repeat measurement	0.87	13.59 (11.34 – 17.83)	(t=0.44, P=.66)	P<.001
Theoretical				
Baseline	N/A	11.43(9.84 – 13.44)	0.04 (-0.20 to 0.28)	0.99 (0.99-0.998)
Repeat measurement	N/A	11.45(9.82 – 13.18)	(t= 1.93, P= .07)	P<.001
LVPEI				
Eccentric viewing				
Baseline	1.04	11.41 (9.52 –14.94)	-	
Repeat measurement	-	-		
Prism-based				
Baseline	0.97	12.16 (10.54 – 15.45)	0.08 (-1.70 to 1.90)	0.83 (0.64 - 0.92)
Repeat measurement	0.99	11.92 (10.54 – 14.57)	(t=0.49, P=.63)	P<.001
Theoretical				
Baseline	N/A	11.72 (10.14 – 13.25)	-	
Repeat measurement	-	-		

t represents paired t test of the mean difference between baseline measures and repeat measurement, and P represents the statistical significance.

511

Table 4. Results of mean signed and mean absolute difference between the three techniques. The mean signed difference was computed using one-way ANOVA, with post-hoc test employing Scheffe method. P-values represent probability of mean difference being statistically significantly different from zero.

	Mean difference (signed) Mean $^{\circ}$/mm (95% LOA)	Mean difference(absolute) Mean $^{\circ}$/mm (95% LOA)
UU		
Prism-based vs Eccentric viewing technique	-1.63 (-1.17 to 4.43) P<.001	1.87 (-0.27 to 4.01) P<.0001
Prism-based vs Theoretical technique	-2.54 (-0.32 to 5.40) P<.001	2.54 (-0.32 to 5.40) P<.0001
Eccentric viewing vs Theoretical technique	-0.91 (-1.40 to 3.22) P<.024	0.98 (-1.22 to 3.18) P<.001
LVPEI		
Prism-based vs Eccentric viewing technique	-0.77 (-1.58 to 3.1) P<.03	1.11 (-0.63 to 2.85) P<.0001
Prism-based vs Theoretical technique	-0.67 (-0.98 to 2.30) P<.06	0.91 (-0.19 to 2.01) P<.0001
Eccentric viewing vs Theoretical technique	0.10 (-1.73 to 1.53) P=.93	0.57 (-0.61 to 1.75) P<.0001

Negative values observed in the 95% LOA for the mean absolute difference column is due to the mean difference data being skewed, although individual data points were all positive (see Figures 3A-C and 6A-C).
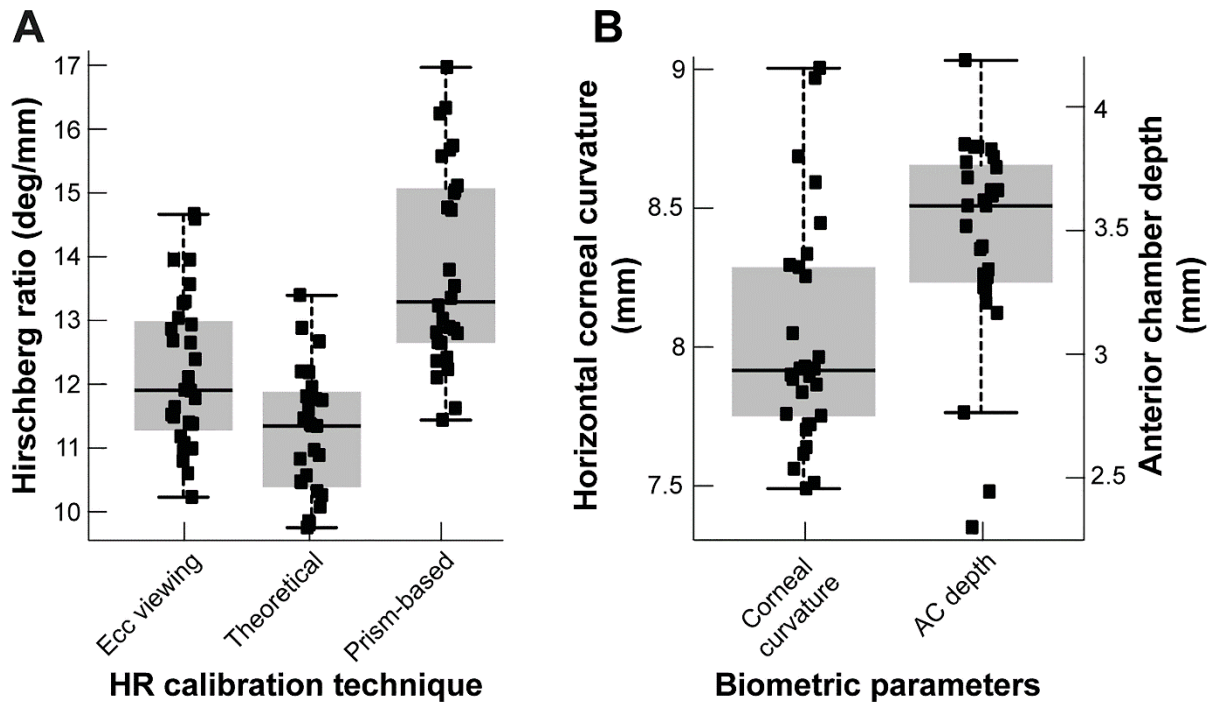
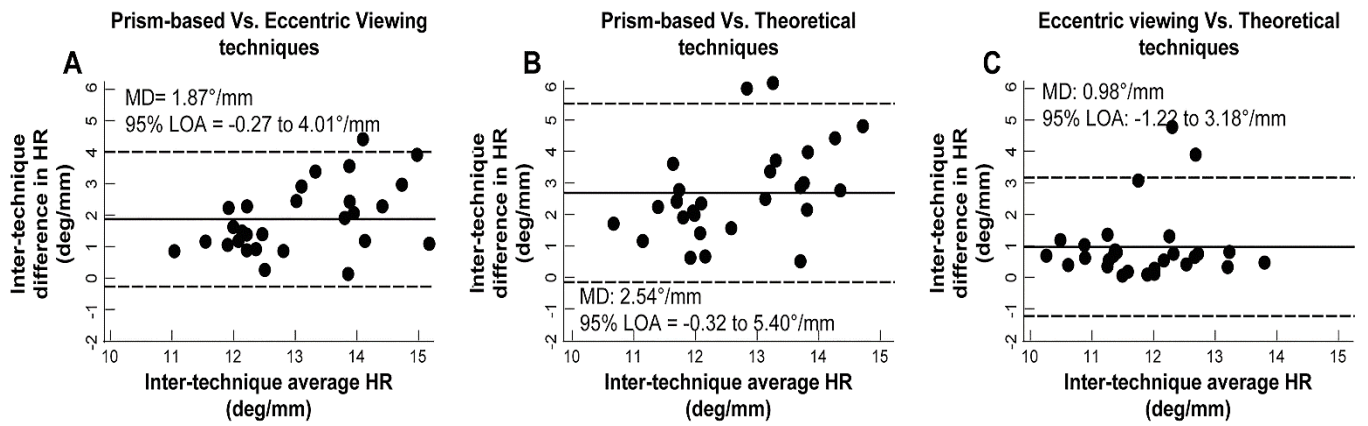519
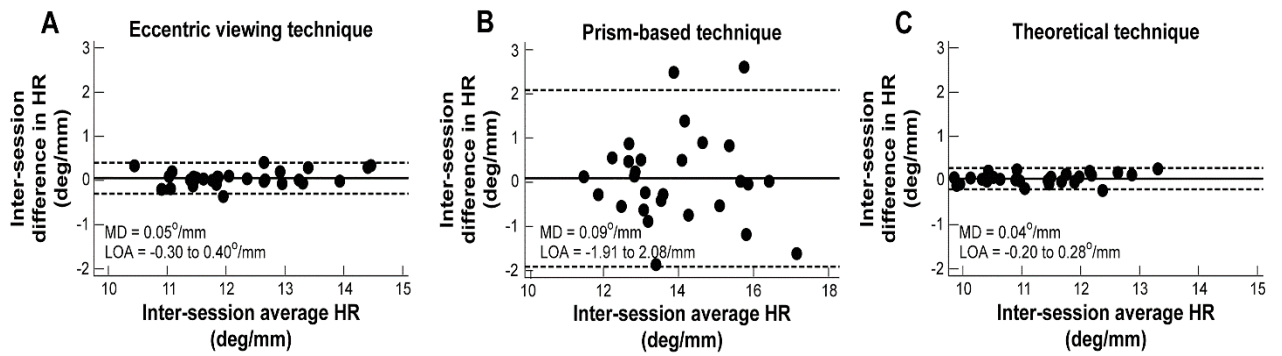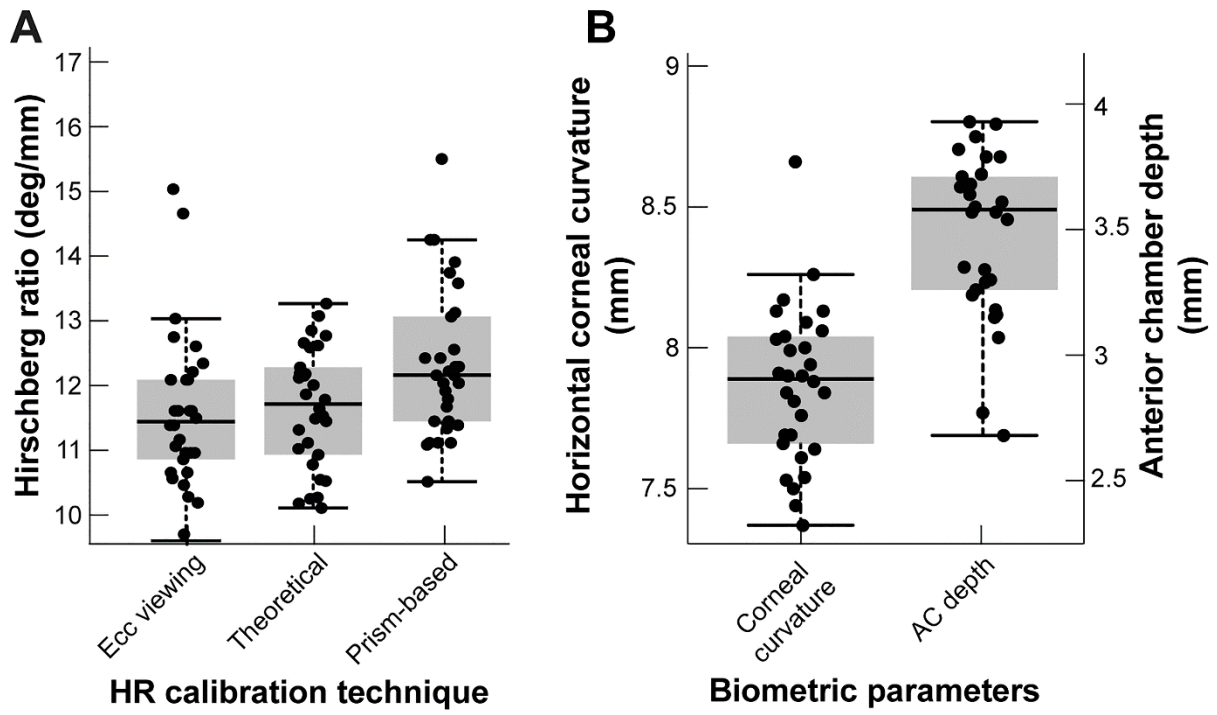
A

B

C

D
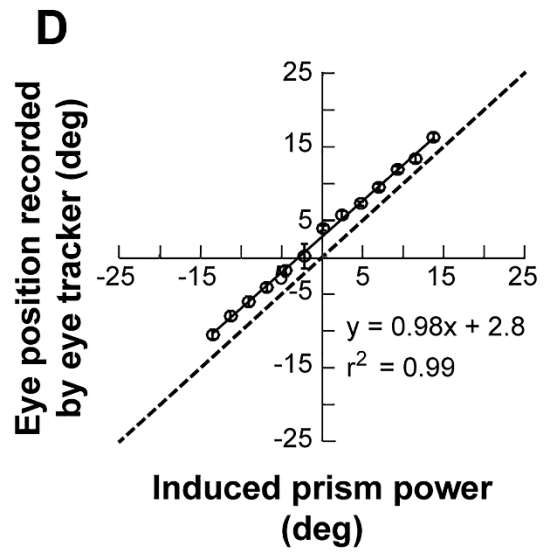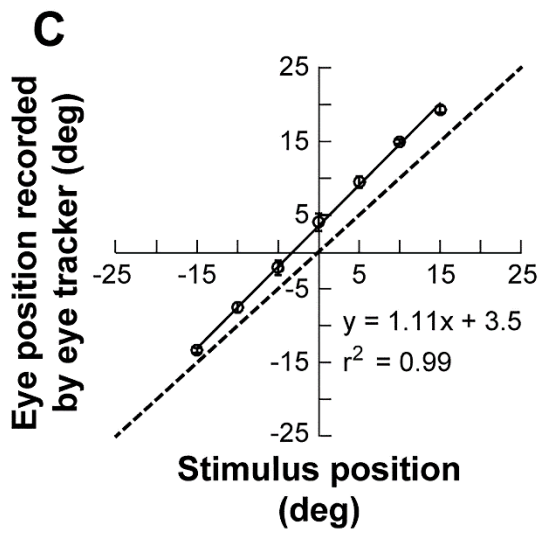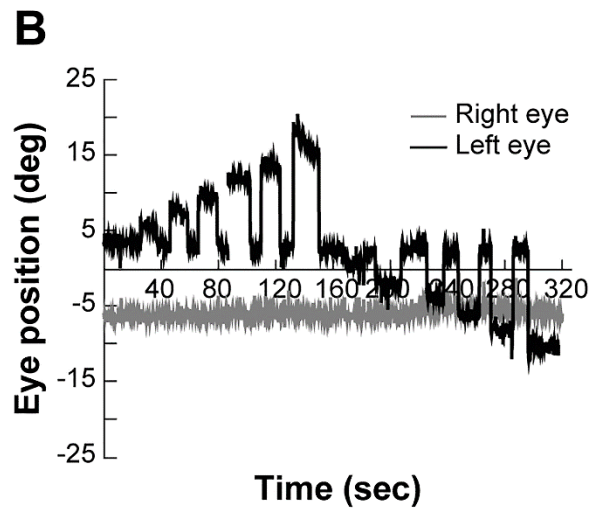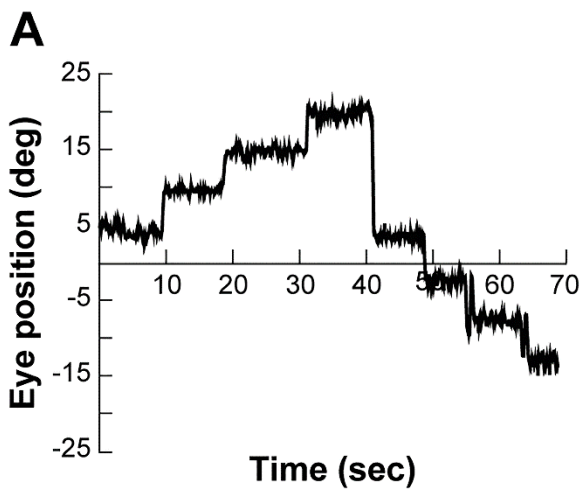
523   Figure 2A-B

524

525

526   Figure 3A-C

527

528

529   Figure 4A-C

530

531

532

533   Figure 5A-B

534

535