J Opt Soc Am A Opt Image Sci Vis. Author manuscript; available in PMC 2014 March 27.

Published in final edited form as:
J Opt Soc Am A Opt Image Sci Vis. 2013 May 1; 30(5): 923–931.

Empirical variability in the calibration of slope-based eccentric photorefraction

Shrikant R. Bharadwaj[1,2,*], **N. Geetha Sravani**[2], **Julie-Anne Little**[3], **Asa Narasaiah**[2], **Vivian Wong**[4], **Rachel Woodburn**[4], and **T. Rowan Candy**[4]

[1]Prof. Brien Holden Centre for Eye Research, Hyderabad Eye Research Foundation, L V Prasad Eye Institute, Road no. 2, Banjara Hills, Hyderabad, Andhra Pradesh 500034, India

[2]Bausch & Lomb School of Optometry, L V Prasad Eye Institute, Road no. 2, Banjara Hills, Hyderabad, Andhra Pradesh 500034, India

[3]Vision Science Research Group, University of Ulster, Coleraine, Northern Ireland, UK

[4]Indiana University School of Optometry, Bloomington, Indiana 47405, USA

Abstract

Refraction estimates from eccentric infrared (IR) photorefraction depend critically on the calibration of luminance slopes in the pupil. While the intersubject variability of this calibration has been estimated, there is no systematic evaluation of its intrasubject variability. This study determined the within subject inter- and intra-session repeatability of this calibration factor and the optimum range of lenses needed to derive this value. Relative calibrations for the MCS PowerRefractor and a customized photorefractor were estimated twice within one session or across two sessions by placing trial lenses before one eye covered with an IR transmitting filter. The data were subsequently resampled with various lens combinations to determine the impact of lens power range on the calibration estimates. Mean (±1.96 SD) calibration slopes were 0.99 ± 0.39 for North Americans with the MCS PowerRefractor (relative to its built-in value) and 0.65 ± 0.25 Ls/D and 0.40 ± 0.09 Ls/D for Indians and North Americans with the custom photorefractor, respectively. The ±95% limits of agreement of intrasubject variability ranged from ±0.39 to ±0.56 for the MCS PowerRefractor and ±0.03 Ls/D to ±0.04 Ls/D for the custom photorefractor. The mean differences within and across sessions were not significantly different from zero ($p > 0.38$ for all). The combined intersubject and intrasubject variability of calibration is therefore about ±40% of the mean value, implying that significant errors in individual refraction/accommodation estimates may arise if a group-average calibration is used. Protocols containing both plus and minus lenses had calibration slopes closest to the gold-standard protocol, suggesting that they may provide the best estimate of the calibration factor compared to those containing either plus or minus lenses.

1. INTRODUCTION

The photorefraction technique provides a rapid and noninvasive estimate of the eye's refractive error from a remote distance [1]. These features are particularly advantageous when working with uncooperative subjects [2–4] and in situations where high sampling rates are required [3,5–7]. Accordingly, a number of commercial devices have been developed

---

[*]Corresponding author: bharadwaj@lvpei.org.

using this technique for vision screening of infants and young children, and also for monitoring accommodative performance (e.g., MCS PowerRefractor [8]).

The principle of photorefraction involves projecting light into the eye from a localized light source (e.g., an LED) and then examining the distribution of light reflected from the retina, through the pupil, and into the aperture of a camera [9–11]. In slope-based eccentric photorefraction, the version that is most commonly used currently, there are a number of LEDs positioned eccentrically from the camera aperture that, in combination, generate an approximately linear distribution of reflected light across the pupil. The slope of this light distribution can be calibrated to indicate the sign (myopic or hyperopic) and magnitude of the eye's defocus over a limited but useful operating range. This approach can be applied to both eyes concurrently at video frame rates, while pupil size and eye position are recorded simultaneously using first Purkinje image tracking [4,7,12].

There are a number of factors that can impact the luminance distribution of the reflected light across individuals. Some of them have been modeled and discussed: the distance between the subject and camera, the size of the camera's limiting aperture, pupil size, the eccentricity and design of the LED array, and the refractive error of the eye, for example [9,10,13,14]. Others are less well understood: reflectance properties of the retina [15,16], the distance between the retinal structures that reflect near infrared (IR) light and the photoreceptors that initiate the visual response [17], higher-order monochromatic aberrations [11,13], chromatic aberration and small variations in the vertex distance (VD) of the lens placed before the eye, for example. Given the inability to predict the light distribution across the pupil for a given refractive error, commercial photorefraction devices have developed an average instrument calibration derived from empirical measurements collected from numerous subjects [8,14]. Consequently, the average refractive measurement across a group of subjects is likely to be relatively accurate, but the measurement error for any individual subject is unknown without individual calibration.

Theoretically, the calibration error can manifest in two independent forms—an absolute error or a relative error. The absolute calibration is determined by comparing the refraction estimate from the photorefractor with an estimate from a gold standard technique (e.g., retinoscopy). The relative calibration is estimated by measuring the rate of change in photorefractor reading per diopter change in an eye's defocus [18,19]. Both need to be considered in generating an accurate estimate of refractive error. An absolute calibration is complex to conduct on a routine basis, as the photo-refractor reading must be compared with a simultaneous measurement taken with the gold standard technique, ideally with accommodation stabilized using a cycloplegic drug and pupil size matched to those under habitual viewing conditions [11,13]. Using retinoscopy, Blade and Candy (2006) found the absolute calibration for individuals to vary from −0.43D to +0.05D around the proprietary instrument (MCS PowerRefractor) value for four adult subjects (mean mean ± 1SD: − 0.28 ± 0.22D) [18].

The relative calibration is used to calibrate changes in the eye's refraction (in studies of accommodation, for example), rather than its true underlying defocus, and is obtained from the slope of the function when luminance slope in the pupil is plotted as a function of lens power placed before the eye. It does not require a gold standard estimate of true, absolute, refractive error and can also be performed easily under experimental conditions without cycloplegia, utilizing natural pupil sizes. Relative calibrations have been found to vary across individuals, with Blade and Candy (2006) observing a range of approximately±50% of the mean value across their 13 adult subjects (mean ± 1SD:0.90 ± 0.18; range: 0.55–1.14), relative to the default proprietary calibration of the instrument [18]. The goal of this study was to determine whether these relative calibration estimates are actually repeatable within the same subject, both within and across days, and to determine the potential impact of different ranges of lens powers on the calibration estimate. Several different ranges of lens powers have been used in the past (−5D to +5D lenses in 1D steps [14], −6D to +6D lenses in 2D steps [8], −1D to +5D lenses in 1D steps [7], +1D to +4D lenses in 1D steps [3,18], and only +2D and +4D lenses [20]).

The overall purpose of this study was to provide a clearer sense of how stable the relative calibration estimates are within an individual and which range of lens powers provides the best and quickest estimate of the relative calibration function. Data were collected with two eccentric photorefraction instruments. One, the commercially available MCS PowerRefractor, has been used in numerous studies and is therefore of wide relevance. Newer versions of this instrument are now available (e.g., PlusOptix PowerRef II and III). This instrument has a proprietary algorithm for converting photorefraction images into refraction estimates. The other, a photorefraction system based in MATLAB, was developed locally in our group. It has not been used extensively by other groups, but does provide access to the raw image data and therefore permits an assessment of the calibration characteristics that are general to the photorefraction technique.

2. METHODS

The study was conducted at the Indiana University School of Optometry (IUSO), Bloomington, USA and at the L V Prasad Eye Institute (LVPEI), Hyderabad, India. All data collected using the MCS PowerRefractor were collected at IUSO, while data collected using the custom-designed photorefractor were collected at IUSO and LVPEI. Forty-eight adults (age range: 19–55 years) from IUSO and 61 adults (age range: 18.5–31.3 years) from LVPEI were recruited. Approximately 65% of the subjects recruited for the study were female. Subjects were emmetropic (with uncorrected distance visual acuity of 20/20 or better and wearing no habitual refractive correction) or they were myopic and fully corrected using soft contact lenses. Astigmatism was <0.5D in all subjects recruited for the study. None of the subjects had clinical detectable ocular pathology, and it happened that none of the subjects were corrected hyperopes. The study adhered to the tenets of the Declaration of Helsinki and was conducted after approval by the local Institutional Review Boards of IUSO and LVPEI.

The overall design of the MCS PowerRefractor and the custom photorefractor was similar. Details of the MCS PowerRefractor can be found elsewhere [8]. The custom photorefractor consisted of a 4 × 6 array of IR LEDs (peak spectral emission at 850 nm) mounted directly below an IR sensitive digital camera (Table 1). The top row of LEDs was 4 mm below the camera aperture, with the other three rows evenly spaced at 7 mm from the center of one LED row to the next. The camera was connected to a computer and the photorefraction data were analyzed offline using MATLAB.

The relative calibration protocol for both instruments was as follows. Subjects were aligned with the camera while their left eye fixated on a visual target. The right eye was occluded using an IR transmitting filter (Table 1). Trial lenses of varying plus and minus powers were placed before the filter over the right eye at a VD of 10–14 mm (see Table 1 for lens range used in the two instruments). Four seconds of photorefraction data were collected from both eyes simultaneously for each lens power. Any frame containing a blink or reflections from the lens, collected from an eccentricity of >30° from the pupillary axis, or with pupil diameter <3 mm or >8 mm was excluded from the analysis. For the MCS PowerRefractor, the slope of the photorefraction luminance profile across the pupil was calculated in both eyes for each usable video frame and converted into refraction values using the proprietary software provided with the instrument. The difference in average refraction between the eyes for each lens power (i.e., measured anisometropia) was then plotted as a function of lens power and an ordinary least squares regression was performed on the linear portion of the data to determine the change in anisometropia per diopter change in lens power. This value provided the calibration factor for that dataset relative to the defocus estimates generated by the proprietary software. For the custom photorefractor, the luminance slope was calculated across the central 80% of the vertical pupil diameter for each eye and then scaled reciprocally by the mean luminance across the pupil in that frame. This scaling accounted for any multiplicative change in slope associated with pupil size. The difference in the mean raw slopes between the two eyes was plotted as a function of the lens power and the same ordinary least squares regression was performed on the linear portion of these data to obtain the defocus calibration estimate for the custom system.

The relative calibration factor for the MCS PowerRefractor was a unitless quantity as it was estimated relative to the inbuilt proprietary calibration of the instrument. The calibration estimate for the custom photorefractor was in units of luminance slope per diopter (Ls/D). The calibration protocol, looking at anisometropia as a function of lens power, minimized the impact of accommodative changes on the calibration measurements and permitted calibration under habitual viewing conditions, without cycloplegia and the associated mydriasis. Accommodative responses in the two eyes are yoked and any change in accommodative state of one eye would be accompanied by an equivalent change in accommodative state of the fellow eye, without impacting the induced difference in refraction [21]. In reality, the refraction of the uncovered eye varied very little during the measurements, indicating relatively stable accommodation.

A. Aim 1: Repeatability of Calibration Function

The intrasubject variability in the calibration factor of the MCS PowerRefractor was determined by repeating the protocol four times on 16 subjects at IUSO. Two calibration routines were recorded at the same visit and two more routines were recorded within 1 week of each other. The intrasubject variability of the custom photorefractor was determined by repeating the protocol three times on 23 Indians at LVPEI and twice on 20 North Americans at IUSO. At LVPEI, the two calibration routines were recorded at the same visit while the third was recorded between 2 and 4 weeks after the first session. At IUSO, the second routine was performed within 1 week of the first session. Data collected twice within the same session describe intrasession repeatability while those collected across different sessions describe intersession repeatability in the calibration estimate.

B. Aim 2: Impact of Lens Power Range on Calibration Function

The impact of the lens powers used on data collected using the MCS PowerRefractor was determined in a subset of 12 subjects at IUSO, by repeating the protocol with a larger range of lens powers (+8D to −6D in 1D steps; gold standard set). The data collected were then sampled in five different lens combinations and the results compared to those obtained with the full “gold standard” set. The protocols were (i) sampling every 2D across the full lens range instead of every 1D (the “alternating” protocol), (ii) only positive lenses in 1D steps (induced myopia, the “positive” protocol), (iii) only negative lenses in 1D steps (induced hyperopia, the “negative” protocol), (iv) 0D to +4D in 1D steps and −2D (the “five-lens” protocol), and (v) +2D and +4D, as used by Gabriel and Mutti [20] (the “two-lens” protocol). The impact of lens powers used on the calibration estimate from the custom photorefractor was explored in the same way using 20 subjects each from IUSO and LVPEI. The data obtained from the resampled protocols were compared with those obtained from the full “gold standard” +8D to −8D lens set.

Data analyses were performed using MATLAB, Microsoft Excel, and SPSS. Statistical analyses included ordinary least squares linear regression, student's *T*-test, and Bland–Altman style analysis for examining repeatability. The agreement between two measures of relative calibration factors were described using the mean difference ±1.96* standard deviation of the difference [±95% limits of agreement (LOA)].

3. RESULTS

Data collection was successful in all 16 subjects using the MCS PowerRefractor. Data collection was successful in 56 of the 61 Indians at LVPEI and all 20 North Americans at IUSO using the custom photorefractor. The data from 5 Indians were discarded due to poor image quality ($n = 2$) and/or small pupil size ($n = 3$). For both instruments, data collected from the eye with the lens changed linearly with lens power typically between ±5D, while the data from the eye without the lens remained constant (Fig. 1). The linear range found with the MCS PowerRefractor at IUSO and the custom photorefractors at IUSO and LVPEI is consistent with the MCS PowerRefractor recommended range of 4D of hyperopia to 6D of myopia [8,14].

A. Repeatability of the Calibration Function

Four sets of calibration data were collected using the MCS PowerRefractor. Lens powers ranging from 0D to +4D, in 1D steps, and a −2D lens were used to obtain these data. The mean (±1.96 * SD) calibration values across subjects (an indication of intersubject variability) were 0.99 (±0.39), 1.04 (±0.40), 1.09 (±0.58), and 1.12 (±0.44) for the first, second, third, and fourth routine, respectively. Comparison of data collected from routines 1 and 2 indicated the intrasession repeatability of calibration while comparison of data collected from routines 3 and 4 indicated the intersession repeatability of calibration. There was no significant difference in the means between routines 1 and 2 ($df = 15$; $t = 0.86$; $p = 0.40$) and between routines 3 and 4 ($df = 15$; $t = 0.42$; $p = 0.68$). Figure 2 shows Bland–Altman type plots of the intrasession and intersession variability of the calibration measurements. The mean (±95% LOA) intrasession signed difference between routines 1 and 2 was 0.04 (±0.39) and the difference between routines 3 and 4 was 0.03 (±0.56), indicating that the calibration estimates varied within approximately ±0.5 units across all routines. The paired intrasession and intersession differences were not significantly different from each other, either (df = 15; $t = 0.16$; $p = 0.87$).

For the custom photorefractor, three sets of data were obtained from Indian eyes at LVPEI and two sets of data were obtained from North American eyes at IUSO. Lens powers ranging from −8D to +8D, in 1D steps, were used to obtain these data. Frequency histograms of the defocus calibration factors obtained from session 1 at LVPEI and IUSO are shown in Fig. 3. Just like the MCS PowerRefractor [18], calibration values obtained using the custom photorefractor also showed intersubject variability, with a mean (±1.96 * SD) slope from set 1 of 0.65 (±0.25)Ls/D for the Indians and 0.40 (±0.09)Ls/D for North Americans. The mean slopes of the subjects in India and North America were statistically significantly different from each other ($df = 19$; $t = 3.45$; $p < 0.001$). To confirm that this difference was not related to instrument or experiment differences at LVPEI and IUSO, the calibration protocol was repeated in 6 European and 3 African subjects who were visiting LVPEI. The calibration factors obtained from the 6 Europeans (range: 0.37 Ls/D to 0.50 Ls/D) were similar to the data from the North American cohort [Fig. 3(a)] while the calibration factors obtained from the 3 Africans (0.68 Ls/D to 0.70 Ls/D) were similar to the data from the Indian cohort [Fig. 3(a)].

Comparison of data collected from sets 1 and 2 in the Indian eyes represented the intrasession repeatability of calibration. For the Indian data, the mean (±1.96 * SD) calibration factor from set 1 (0.74 ± 0.31 Ls/D) was not significantly different from that of set 2 (0.72 ± 0.31 Ls/D) (df = 23; paired $t = 0.92$; $p = 0.14$) [Fig. 4(a)]. The mean (±95% LOA) signed difference between the two measurements was 0.02 ± 0.11 Ls/D and the mean (±95% LOA) absolute difference was 0.05 ± 0.07 Ls/D. Comparison of data collected from sets 1 and 3 in Indians and from sets 1 and 2 in North Americans represented the intersession repeatability of calibration. The intersession variability in the Indian and North American data were qualitatively similar to each other. The mean (±1.96 * SD) calibration factors for sets 1 and 3 in Indians were not statistically significantly different from each other (set 1: 0.64±0.22 Ls/D; set 3: 0.66 ± 0.25 Ls/D) ($df = 23$; paired $t = 0.66$; $p = 0.38$), with a mean (±95% LOA) signed difference of 0.01 ± 0.14 Ls/D and absolute difference of 0.07 ± 0.09 Ls/D across subjects [Fig. 4(b)]. The means (±1.96 * SD) of set 1 (0.39 ± 0.19 Ls/D) and set 2 (0.42 ± 0.18 Ls/D) for the North Americans were marginally significantly different from each other with no correction for multiple comparisons (paired $t = 2.43$, $df = 19$, $p = 0.03$). The small individual differences tended to be consistently positive, with a mean (±95% LOA) signed difference of 0.03 ± 0.10 Ls/D and a mean absolute difference of 0.05 ± 0.06 Ls/D [Fig. 4(c)].

B. Impact of Lens Power Range on Calibration Function

Table 2 and Figs. 5(a)–5(e) show the results of resampling the range of lenses using the five different protocols relative to the "gold standard" routine (±8D to ±6D in 1D steps) for the MCS PowerRefractor. The ends of the linear region were determined by visual inspection of the data before performing the linear regression. The calibration slopes obtained in the alternating and five-lens re-sampling procedures were closest to and most strongly correlated with the gold standard values [Table 2, Figs. 5(a) and 5(d)]. While there was no significant difference between either the five-lens or the alternating lens protocols compared with the gold standard protocol, the 95% LOA range was wider in the five-lens protocol than in the alternating lens protocol. The positive and two-lens protocols underestimated the calibration factors significantly, relative to the gold standard and had 95% LOAs that were similar to each other [Table 2, Figs. 5(b) and 5(e)]. The negative protocol tended to overestimate the calibration factors (without reaching statistical significance), had the widest 95% LOA, and was least correlated with the gold standard [Table 2, Fig. 5(c)].

Table 2 and Figs. 5(f)–5(o) show the results of resampling the first set of the Indian and North American data obtained using the custom photorefractor. Again, the slopes obtained using the alternating protocol were closest to and not significantly different from the gold standard, with the smallest 95% LOA, in both populations ($p > 0.1$ for both) [Table 2 and Figs. 5(f) and 5(g)]. The correlation coefficient was also highest between the alternating protocol and the gold standard for both populations ($r > 0.94$ for both) and weakest between the two-lens protocol and the gold standard ($r = 0.8$ for Indians and $r = 0.44$ for North Americans). The trends in the other protocols were similar to those observed with the MCS PowerRefractor [Table 2 and Figs. 5(a)–5(e)].

4. DISCUSSION

The accuracy of defocus estimates obtained using eccentric photorefraction depends upon the calibration of the luminance slope in the pupil. This study determined (i) the intrasubject repeatability of the relative calibration of a commercial and a custom-designed photorefractor and (ii) the impact of lens range used on the relative calibration estimates obtained using these instruments. Data were obtained only on emmetropes or on myopes who were rendered emmetropic using soft contact lenses. Uncorrected refractive error will be, however, expected to only induce an absolute offset in the calibration function [i.e., refraction output or luminance slope plotted against induced lens power (Fig. 1)], without affecting its slope (i.e., the relative calibration estimate).

A. Repeatability of Relative Calibration Factor

Both photorefractors appeared to exhibit intersubject variability in their calibration factors [estimated from the 95% CI's (1.96 * SD)], with the ±95% CI's ranging from 22.5% to 38.5% of the mean value for the custom photorefractor and around 40% of the mean value for the MCS PowerRefractor. The variability found in this study with the MCS PowerRefractor is similar to that obtained by Blade and Candy using the same instrument [18]. This intersubject variability would have a significant impact on the final defocus estimates obtained using photorefraction. For example, the mean calibration slope of the custom photorefractor for Indian data was 0.65 Ls/D, with the steepest and flattest slopes being 1.14 Ls/D and 0.41 Ls/D, respectively [Fig. 3(a)]. A unit change in luminance slope would therefore correspond to a refraction change of 1.45D (1/0.69 Ls/D = 1.45 D/Ls), 0.88D or 2.4D, respectively. The true refraction change per unit change in slope would therefore be overestimated by 0.57D for the subject with the steepest slope and underestimated by 0.95D for the subject with the flattest slope if the group-average calibration factor were used.

How much of this variability is due to true variation between subjects, as both photorefractors also exhibited intrasubject variability in the calibration factor? The overall intrasubject variability of the custom photorefractor was somewhat smaller (±95% LOA between two sessions ranged from ±0.10 Ls/D to ±0.14 Ls/D) than that of the MCS PowerRefractor (±95% LOA between two sessions ranged from ±0.39 to ±0.56 units), indicating better repeatability of calibration in the former than in the latter. Data collected twice in the same visit or across different days resulted in mean signed differences that were not significantly different from zero [except for the North American data in the custom photorefractor, where the mean difference (±95% LOA) was 0.03 ± 0.10 Ls/D; $p = 0.03$], indicating that there was little or no intrasession or intersession bias in the calibration measurements of both instruments. The slightly larger intrasubject variability in the MCS PowerRefractor measurements, relative to the custom photorefractor, may arise due to small differences in the range of lenses used to determine the calibration factor. For the MCS PowerRefractor calibration, the five-lens protocol was used (0D to ±4D in 1D steps and ±2D lens) while the gold standard protocol (−8D to ±8D in 1D steps) was used in calibrating the custom photorefractor (Table 1). The resampling exercise undertaken in this study for the MCS PowerRefractor (Aim 2) indicated that the five-lens protocol resulted in slightly more variable results than the alternate lens protocol containing a larger range of plus and minus lens powers (Fig. 5; Table 2). This was true for North American eyes in the custom photorefractor also (Fig. 5; Table 2). Alternatively, the difference in variability between the two instruments may also reflect inherent differences in the LED array, cameras, or software algorithm used to analyze the pupil profiles. This question cannot be tested further due to the proprietary nature of the MCS PowerRefractor software.

The decision to perform an individual calibration for each subject will depend on the required accuracy and precision of the data and whether the study is designed to draw conclusions regarding the mean of a group or for an individual (in a screening protocol, for example). The benefit of performing a calibration routine on individual subjects will also depend on the ratio of intrasubject to intersubject variability. High intrasubject variability would suggest that performing a calibration routine is unlikely to provide a useful estimate of the subject's true calibration factor, while high intersubject variability implies that calibrating the instrument for an individual would be beneficial. The intersubject variability was, in fact, only marginally greater than the intrasubject variability for some calibration protocols. A simple single factor ANOVA performed on the repeatability data collected using the MCS PowerRefractor suggested that the intersubject variability was still significant though ($F(15, 48) = 3.54$, $p = 0.0004$). The ANOVA-based estimator of the intraclass correlation coefficient (ICC) (two-way random effects model [22]) was 0.39, indicating that the intersubject variability explained approximately 40% of the variance in this dataset. ICC for the intersession and intrasession repeatability for Indian eyes and intersession repeatability for North American eyes, all obtained using the custom photorefractor, were 0.74, 0.94, and 0.84, respectively. This indicated a smaller effect of intrasubject variability in these protocols.

B. Factors Causing a Change in Photorefraction Calibration

As listed in the Introduction, several factors could contribute to the observed variability in the defocus calibration estimate. These could have a combination of additive and multiplicative effects on the estimate. Measurements taken during the design and development of the custom photorefractor indicated that a mean luminance compensation for any multiplicative effect of pupil size changes worked robustly for pupil sizes >3 mm and that the pupil slope estimates for a given dioptric value did not change systematically during pupil size variation greater than this value. Similar compensation for the multiplicative effect of pupil size and mean luminance was also employed in the eccentric IR photorefractor developed by Schaeffel *et al.* (1993) [14]. Indeed, the relative calibration slopes of 20 Indian subjects obtained here using the custom photorefractor had no correlation with their pupil diameters (range: 3.5–5.7 mm; $r = -0.14$; $p = 0.87$) (data not shown).

Nonlinearities in the luminance distribution across the pupil (from higher-order monochromatic aberrations particularly at large pupil sizes [11] or from tear-film disturbances [23], for example) may also influence the defocus calibration. The pupil luminance profiles obtained in this study were typically linear, with only the far ends of the profile showing nonlinearity in many subjects (only the central 80% of the profile was included in the analyses for the custom-built system). While it may explain differences between subjects, this factor would need to be changing between two measurements performed on the same subject if it were to explain the intrasession and intersession variability seen in the current data (e.g., due to changes in monochromatic abberrations of the eye with accommodation [24]).

Measurement noise could also be introduced by small variations in the angle and VD at which the trial lens was held before the eye. In this study, trial lenses were held by hand at a VD of 10–14 mm from the eye. While manual positioning of the lens allowed the experimenter to make slight adjustments to avoid lens reflections, this might have induced small variations in VD across subjects and experimental sessions. Assuming that the calibration factor is unity for a VD of 12 mm, a change in VD of 10 mm toward (i.e., VD = 2 mm) or away (i.e., VD = 22 mm) from the subject, results in calibration factors of 1.06 and 0.94, respectively, based on changes in the effective power of the lenses. Given that the VD is unlikely to have changed by 10 mm in this study (at most, they changed by 2–3 mm), the calculations indicate that this might only account for a small amount of the observed variability.

Interestingly, in addition to the intersubject and intrasubject variability in calibration estimates, this study also observed systematic differences in the calibration slopes obtained using the custom photorefractor between Indian and North American eyes (Fig. 3). The mean (±1.96 * SD) calibration slope in Indian eyes [0.65 ± 0.25 Ls/D] was statistically significantly larger than those of North American eyes [0.40 ± 0.09 Ls/D]. The compensation for the multiplicative effect of mean pupil luminance incorporated into the analysis of the custom photorefractor suggests that this difference cannot be a simple result of brightness difference resulting from increased absorption by pigment or pupil size, however. The calibration slopes of European subjects (with similar pigmentation as the North Americans) visiting LVPEI were in the same range as the slopes of North American subjects obtained at IUSO [compare asterisk symbols in Figs. 3(a) and 3(b)]. This suggests that the difference in slopes between the two groups is unlikely to arise from differences in instrumentation of experiment setup. The Indian subjects were all of typical Indian racial and ethnic backgrounds, with a range of pigmentation. The North American subjects also had a range of racial and ethnic backgrounds, but on average were less pigmented than the Indian subjects. While fundal reflectance in the visible spectrum tends to be less in darker eyes than in lighter eyes, this difference with eye color reduces in IR light [15,16]. Further experiments are required to fully understand the origin of these differences in calibration factors obtained from subjects from different geographical regions.

C. Impact of Lens Power Range on the Relative Calibration Estimate

When collecting data from subjects with short attention spans (e.g., infants and children), any calibration protocol needs to be short in duration yet reliable. An analysis of different lens selections was performed to determine the impact of both number and range of lenses on the slope estimate (Fig. 5 and Table 2). The first resampling procedure used alternating lenses over the full range and resulted in minimal differences when compared with the full protocol [Figs. 5(a), 5(f), and 5(k)]. This protocol required eight or nine lenses instead of 15 or 17 lenses in the gold standard protocol, with the alternate protocol taking about 2–3 min to complete in a cooperative adult. The positive and negative lens protocols required nine or seven lenses each and resulted in differences in estimates that were more variable than for the alternating protocol [Figs. 5(b), 5(g), 5(l), 5(c), 5(h), and 5(m)]. If the purpose was to study accommodative performance using photorefraction, one might wish to perform the calibration using the positive lens protocol as it mimics myopic focus, but as noted above this protocol introduces additional variability relative to the alternating protocol using a similar number of lenses. The final two protocols reduced the number of lenses to two and five and therefore made the data collection shorter in duration [Figs. 5(d), 5(i), 5(n), 5(e), 5(j), and 5(o)]. They also concentrated the data collection in the myopic region. Taking the number of lenses down to two appeared to result in additional variability and bias, while the five-lens protocol including one lens inducing hyperopia (−2D) tended to result in estimates that were more similar to the alternating or positive lens protocols.

Overall, if the goal was to estimate the calibration function over the entire linear region, it would seem that the lenses should span both the myopic and hyperopic regions, with the alternating protocol with approximately eight lenses, or the five-lens protocol with four lenses in the myopic region and one lens in the hyperopic region resulting in the least error relative to the gold standard condition. Gabriel and Mutti [20] used the two-lens protocol to correct refraction data obtained using photorefraction for comparison with cycloplegic retinoscopy. They concluded that the use of this calibration protocol did not improve the agreement between the two datasets. This example illustrates the dilemma of trying to achieve a valid calibration in the shortest time possible. The intraobserver variability for cycloplegic retinoscopy is also on the order of 0.25 to 0.5D [25].

5. CONCLUSIONS

Due to its many advantages—speed, remote location, ease of operation, relatively large working range, and the potential to assess eye alignment—the eccentric IR photorefraction technique has been used to screen for pediatric refractive errors and to monitor the accommodative response. Refraction estimates obtained using this technique depend critically on the conversion of luminance profile into diopters of defocus. This study observed both intersubject and intrasubject variability in the relative calibration estimates. The combined intersubject and intrasubject variability was on the order of ±40% around the mean, implying that significant errors in refraction estimates may arise at an individual level if a group-average calibration is used. The lens range used also had an impact on the variability of calibration estimates, with a calibration protocol spanning both myopic and hyperopic defocus over the linear operating range of the instrument resulting in least variability and providing the best estimate of the full calibration factor.

Acknowledgments

The authors thank all the subjects who participated in the study. We also thank Kevin Haggerty, William Monette, and Tom Kemerly at IUSO for developing the customized photorefractor, and Hope Rittichier, Vaishnavi Raghuram, Gouthami Deshapadhini, and Praveen Kumar for assistance with the data analysis. We also thank Lynn Mitchell for discussion of statistical analysis. The authors acknowledge the following funding sources: Government of India, Department of Biotechnology Ramalingaswami fellowship to S. R. Bharadwaj, Champalimaud Foundation grant to the Prof. Brien Holden Centre for Eye Research, NIH R01 EY014460 to T. R. Candy, NIH P30 EY019008 Core Grant to the Indiana University School of Optometry, NIH T35 EY013937 training support for Rachel Woodburn, and College of Optometrists Research Fellowship, UK to Julie-Anne Little.

REFERENCES

1. Howland HC. Photorefraction of eyes: history and future prospects. Optom. Vis. Sci. 2009; 86:603–606. [PubMed: 19390473]
2. Anderson HA, Manny RE, Glasser A, Stuebing KK. Static and dynamic measurements of accommodation in individuals with Down syndrome. Invest. Ophthalmol. Visual Sci. 2011; 52:310–317. [PubMed: 20739471]
3. Bharadwaj SR, Candy TR. Cues for the control of ocular accommodation and vergence during postnatal human development. J. Vis. 2008; 8(16):1–16. 14. [PubMed: 19146280]
4. Turner JE, Horwood AM, Houston SM, Riddell PM. Development of the response AC/A ratio over the first year of life. Vis. Res. 2002; 42:2521–2532. [PubMed: 12445846]
5. Anderson HA, Glasser A, Manny RE, Stuebing KK. Age-related changes in accommodative dynamics from preschool to adulthood. Invest. Ophthalmol. Visual Sci. 2010; 51:614–622. [PubMed: 19684002]
6. Harb E, Thorn F, Troilo D. Characteristics of accommodative behavior during sustained reading in emmetropes and myopes. Vis. Res. 2006; 46:2581–2592. [PubMed: 16545421]
7. Suryakumar R, Meyers JP, Irving EL, Bobier WR. Application of video-based technology for the simultaneous measurement of accommodation and vergence. Vis. Res. 2007; 47:260–268. [PubMed: 17174376]
8. Choi M, Weiss S, Schaeffel F, Seidemann A, Howland HC, Wilhelm B, Wilhelm H. Laboratory, clinical, and kindergarten test of a new eccentric infrared photorefractor (PowerRefractor). Optom. Vis. Sci. 2000; 77:537–548. [PubMed: 11100892]
9. Bobier WR, Braddick OJ. Eccentric photorefraction: optical analysis and empirical measures. Am. J. Optom. Physiol. Opt. 1985; 62:614–620. [PubMed: 4050963]
10. Howland HC. Optics of photoretinoscopy: results from ray tracing. Am. J. Optom. Physiol. Opt. 1985; 62:621–625. [PubMed: 4050964]
11. Roorda A, Campbell MC, Bobier WR. Slope-based eccentric photorefraction: theoretical analysis of different light source configurations and effects of ocular aberrations. J. Opt. Soc. Am. A. 1997; 14:2547–2556.
12. Bharadwaj SR, Wang J, Candy TR. Pupil responses to near visual demand during human visual development. J. Vis. 2011; 11(4):1–14. 6.
13. Roorda A, Campbell MC, Bobier WR. Geometrical theory to predict eccentric photorefraction intensity profiles in the human eye. J. Opt. Soc. Am. A. 1995; 12:1647–1656.
14. Schaeffel F, Wilhelm H, Zrenner E. Inter-individual variability in the dynamics of natural accommodation in humans: relation to age and refractive errors. J. Physiol. 1993; 461:301–320. [PubMed: 8350267]

15. Delori FC, Pflibsen KP. Spectral reflectance of the human ocular fundus. Appl. Opt. 1989; 28:1061–1077. [PubMed: 20548621]
16. Elsner AE, Burns SA, Weiter JJ, Delori FC. Infrared imaging of sub-retinal structures in the human ocular fundus. Vis. Res. 1996; 36:191–205. [PubMed: 8746253]
17. Glickstein M, Millodot M. Retinoscopy and eye size. Science. 1970; 168:605–606. [PubMed: 5436596]
18. Blade PJ, Candy TR. Validation of the PowerRefractor for measuring human infant refraction. Optom. Vis. Sci. 2006; 83:346–353. [PubMed: 16772892]
19. Seidemann A, Schaeffel F. An evaluation of the lag of accommodation using photorefraction. Vis. Res. 2003; 43:419–430. [PubMed: 12535999]
20. Gabriel GM, Mutti DO. Evaluation of infant accommodation using retinoscopy and photoretinoscopy. Optom. Vis. Sci. 2009; 86:208–215. [PubMed: 19165126]
21. Ball EA. A study of consensual accommodation. Am. J. Optom. Arch. Am. Acad. Optom. 1952; 29:561–574. [PubMed: 12996635]
22. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 1979; 86:420–428. [PubMed: 18839484]
23. Himebaugh NL, Wright AR, Bradley A, Begley CG, Thibos LN. Use of retroillumination to visualize optical aberrations caused by tear film break-up. Optom. Vis. Sci. 2003; 80:69–78. [PubMed: 12553546]
24. Cheng H, Barnett JK, Vilupuru AS, Marsack JD, Kasthurirangan S, Applegate RA, Roorda A. A population study on changes in wave aberrations with accommodation. J. Vis. 2004; 4(4):272–280. 3. [PubMed: 15134474]
25. Hirsch MJ. The variability of retinoscopic measurements when applied to large groups of children under visual screening conditions. Am. J. Optom. Arch. Am. Acad. Optom. 1956; 33:410–416. [PubMed: 13354710]

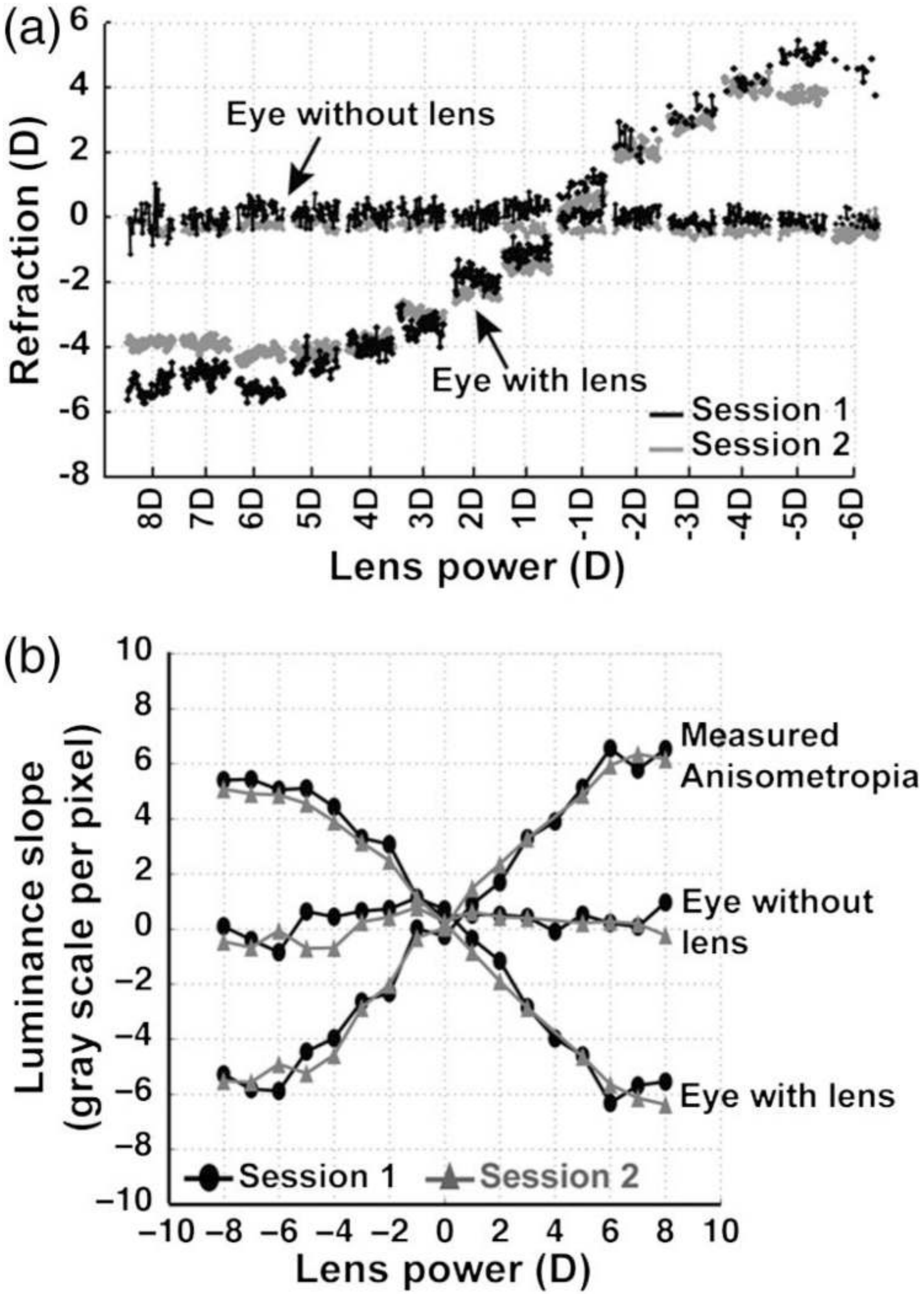

Fig. 1.

Raw datasets collected from one representative subject using the MCS PowerRefractor (panel *a*) and for another subject using the custom photorefractor (panel *b*). Data collected from the right and left eyes are plotted as a function of the lens placed over the right eye. Data collected from two sessions are shown to demonstrate repeatability of the measurements.

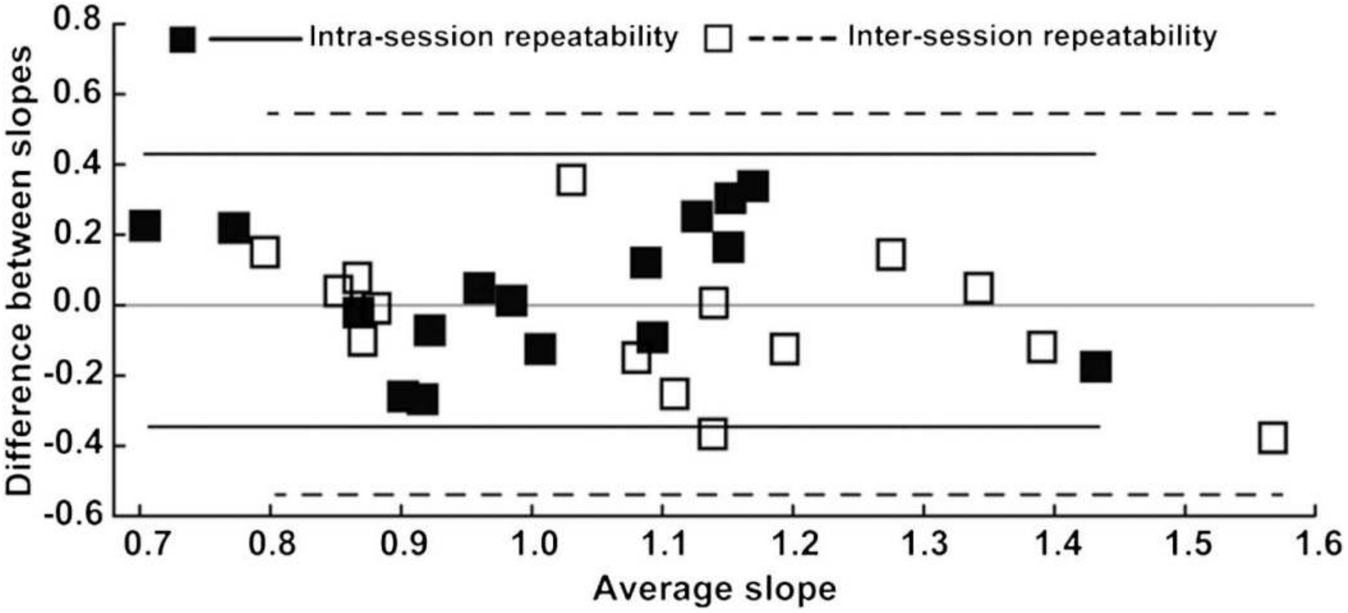

Fig. 2.

Difference in calibration factors obtained using the MCS PowerRefractor between the first and second sessions (intrasession repeatability) and the third and fourth sessions (intersession repeatability) plotted as a function of the mean calibration factor. Solid and dashed lines indicate 95% LOA.

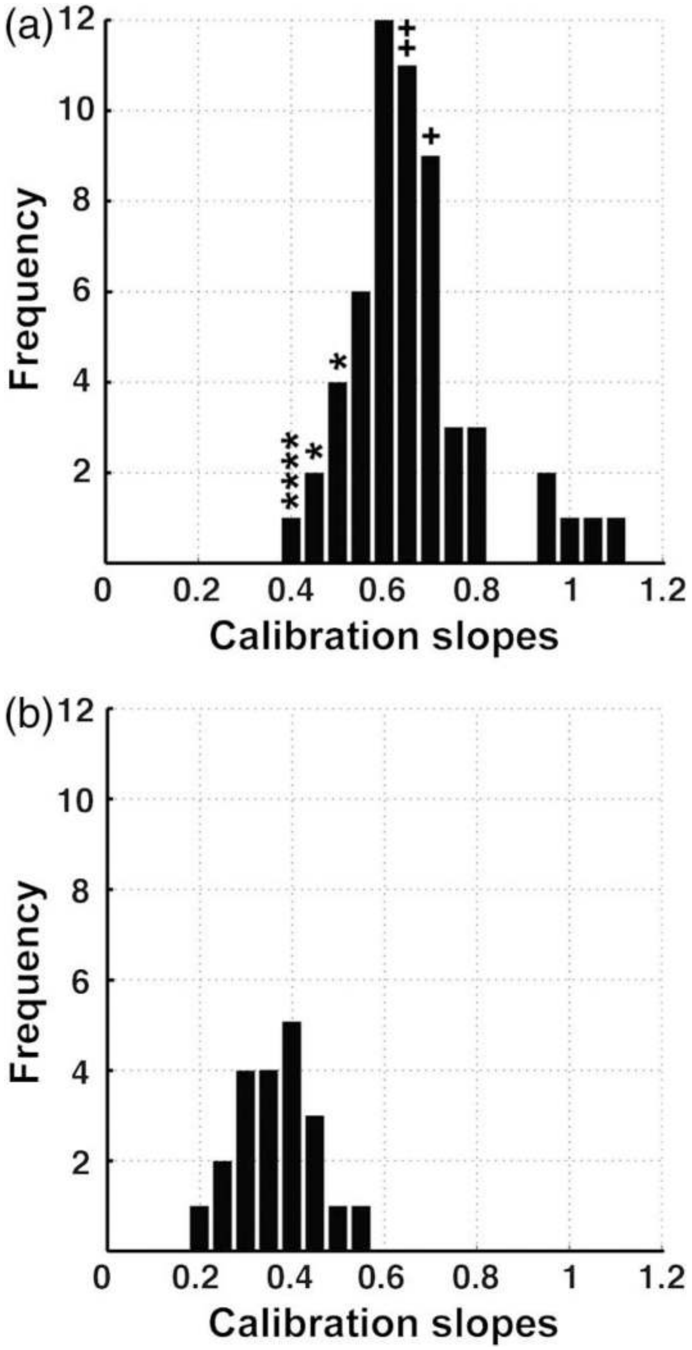

Fig. 3.

Frequency histograms of the calibration factor obtained using the custom photorefractor for all (a) Indian and (b) North American subjects.

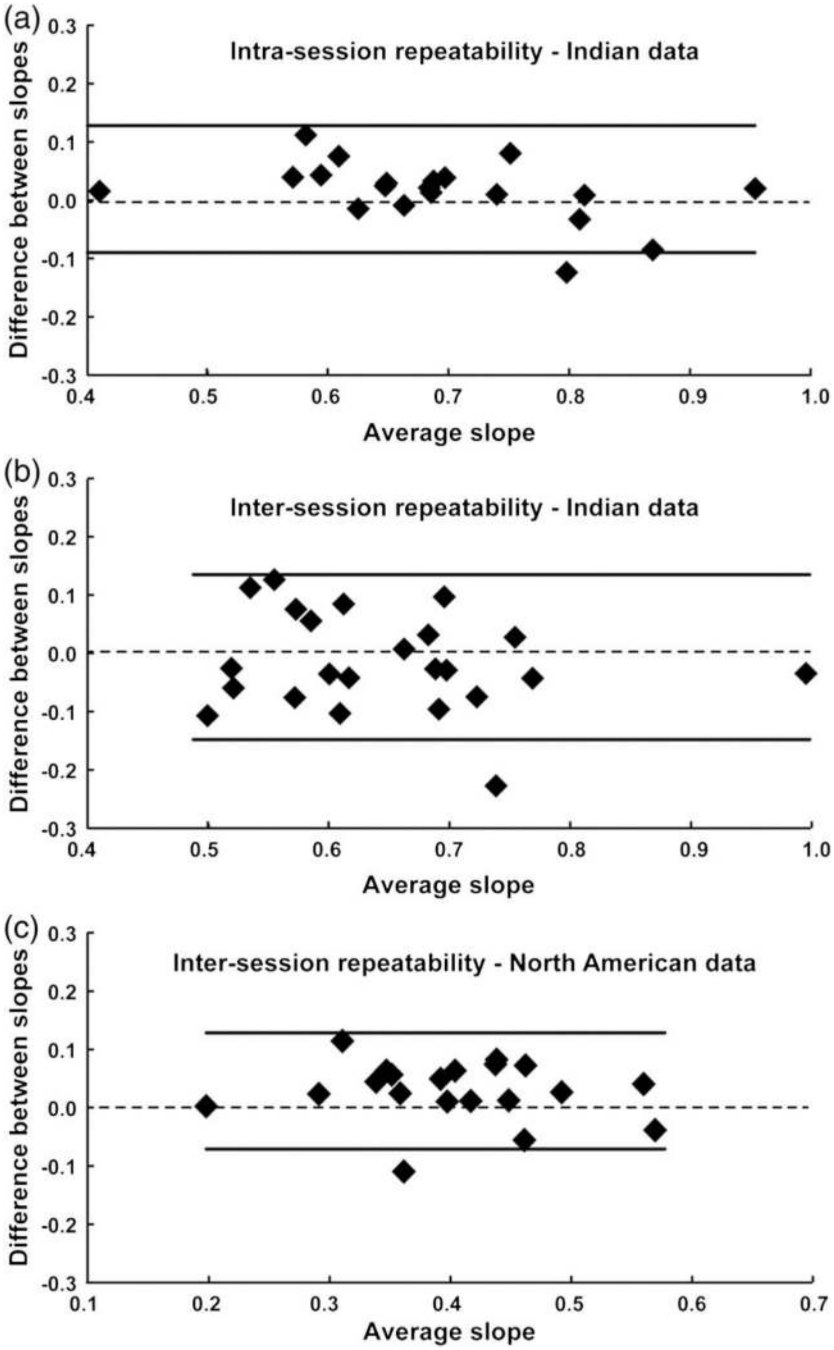

Fig. 4.

(a) Intrasession and (b) intersession repeatability of the calibration factor for Indians, and (c) the intersession repeatability of the calibration factor for North Americans obtained using the custom photorefractors. All other details are the same as Fig. 2.

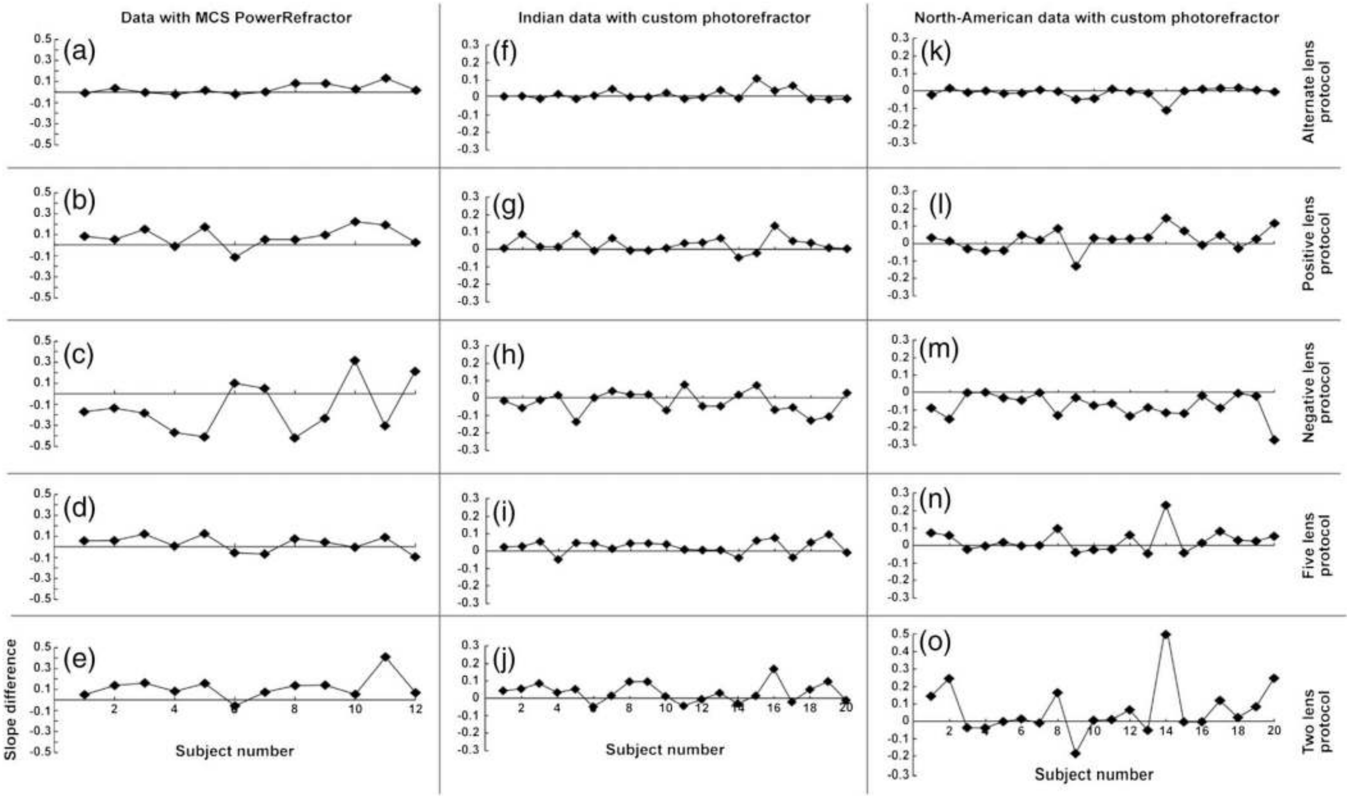

Fig. 5.

Difference in defocus calibration factor between the gold standard and a resampling protocol plotted for each subject. (a)–(f) Data obtained using the MCS PowerRefractor, (f)–(j) data obtained from Indians, and (k)–(o) data obtained from North Americans, both using the custom photorefractor. Subjects are arranged in ascending order of their calibration slope obtained in the gold standard protocol. Positive numbers on the ordinate indicate that the calibration slope obtained from the resampled protocol was larger than that in the gold standard protocol.

Table 1

Design Specifications, Experimental Details and the Data Collection Protocols Followed with the MCS PowerRefractor at IUSO and with the Custom Photorefractors at IUSO and LVPEI

		Custom Photorefractor	
	MCS PowerRefractor	**IUSO**	**LVPEI**
Design specifications			
i. Camera used	–	Point Grey Research FireFly	Point Grey Research Dragonfly Express
ii. Shape of LED array	Trapezoid	Rectangle	Rectangle
iii. Sampling frequency	25 fps	30 fps	30 fps
iv. Viewing distance	100 cm	75 cm	75 cm
Experiment details			
i. Lens range used	0D to +4D in 1D steps and −2D for repeatability; −6D to +8D in 1D steps for resampling	−8D to +8D in 1D steps for repeatability and resampling	−8D to +8D in 1D steps for repeatability and resampling
ii. Vertex distance	10–14 mm	10–14 mm	10–14 mm
iii. IR filter used	Kodak Wratten #87B	Optcast filter Edmund Optics NT43–954	Optcast filter Edmund Optics NT43–954
iv. Calibration units	Unitless	Luminance slope per diopter (Ls/D)	Luminance slope per diopter (Ls/D)
Data collection protocols followed			
i. Intrasession repeatability	Yes (*n* = 16)	No	Yes (*n* = 23)
ii. Intersession repeatability	Yes (*n* = 16)	Yes (*n* = 20)	Yes (*n* = 23)
iii. Resampling	Yes (*n* = 12)	Yes (*n* = 20)	Yes (*n* = 20)

Table 2

Mean ±95% LOA and Minimum to Maximum Range of the Calibration Slopes Obtained in the Gold Standard and in Each of the Five Resampling Protocols for the MCS PowerRefractor and for Indians and North Americans Using the Custom Photorefractor[a]

	Protocol	**Mean ±1.96 SD**	***p*-Value**	**Mean Difference +95% LOA**	***r*-Value**
MCS PowerRefractor data	Gold standard	0.97 ± 0.30			
	Alternate protocol	0.94 ± 0.27	0.35	0.03 ± 0.09	0.95
	Positive protocol	0.89 ± 0.34	0.05	0.08 ± 0.19	0.84
	Negative protocol	1.10 ± 0.44	0.50	−0.13 ± 0.48	0.20
	Five-lens protocol	0.94 ± 0.39	0.95	0.03 ± 0.14	0.95
	Two-lens protocol	0.85 ± 0.34	0.02	0.12 ± 0.22	0.78
Custom photorefractor—Indian data	Gold standard	0.64 ± 0.19 Ls/D			
	Alternate protocol	0.63 ± 0.24 Ls/D	0.80	0.01 ± 0.09 Ls/D	0.94
	Positive protocol	0.61 ± 0.19 Ls/D	0.15	0.02 ± 0.11 Ls/D	0.84
	Negative protocol	0.66 ± 0.26 Ls/D	0.25	−0.03 ± 0.14 Ls/D	0.86
	Five-lens protocol	0.62 ± 0.21 Ls/D	0.10	0.02 ± 0.08 Ls/D	0.92
	Two-lens protocol	0.61 ± 0.23 Ls/D	0.30	0.03 ± 0.14 Ls/D	0.81
Custom photorefractor—North American data	Gold standard	0.39 ± 0.19 Ls/D			
	Alternate protocol	0.40 ± 0.19 Ls/D	0.55	−0.01 ± 0.06 Ls/D	0.95
	Positive protocol	0.37 ± 0.19 Ls/D	0.55	0.02 ± 0.12 Ls/D	0.80
	Negative protocol	0.46 ± 0.26 Ls/D	0.0005	−0.07 ± 0.13 Ls/D	0.87
	Five-lens protocol	0.36 ± 0.22 Ls/D	0.35	0.03 ± 0.13 Ls/D	0.82
	Two-lens protocol	0.32 ± 0.33 Ls/D	0.30	0.07 ± 0.30 Ls/D	0.44

[a] *p*-values = statistical significance for a paired *t*-test; *r*-values = Pearson's correlation coefficient.