



Published in final edited form as:

*Optom Vis Sci.* 2009 December ; 86(12): 1337–1345. doi:10.1097/OPX.0b013e3181be9d9c.

## Modified Bell Retinoscopy: Measuring Accommodative Lag in Children

**Kristina Tarczy-Hornoch, MD, DPhil**

Childrens Hospital Los Angeles, Los Angeles, California and the Department of Ophthalmology, University of Southern California, Los Angeles, California

### Abstract

**Purpose**—To describe a modified bell retinoscopy (MBR) method for quantifying accommodative lag in children, and to assess its repeatability and comparability to other techniques.

**Methods**—In MBR, the target is advanced toward the patient until the retinoscopic reflex is neutralized. A “standardized 40 cm target estimate” of lag was derived for each child using data from three retinoscope distances. Within-visit repeatability was assessed in normal children 5–23 months of age, a heterogeneous group of clinic patients, and a group of children with Down syndrome. Clinic patients were tested on separate days for between-visit repeatability, and also with Nott retinoscopy (NR) and the monocular estimate method (MEM) on Day 2.

**Results**—MBR correlated with NR ( $R=0.84$ ) and MEM ( $R=0.82$ ). MBR and NR estimates were lower than MEM for high lags. Within-visit repeatability of the standardized 40 cm target estimate of MBR in normal children and clinic patients varied with the amount of lag ( $p<0.0001$ ). The repeatability index (RI) for 0.50 D lag was 0.49 D, and for 1.00 D lag it was 0.80 D. Repeatability was similar in children with Down syndrome. In clinic patients, the between-visit RI for 0.50 D lag was 0.60 D for the second estimate of each day, with lower repeatability for the first measure of each day. Repeatability did not vary with age or refractive error. The decrease in repeatability with high lag may be attributable to spatial measurement error.

**Conclusions**—MBR estimates of accommodative lag correlate with traditional dynamic retinoscopy measures over a wide range of lags, and show comparable repeatability. MBR may be a useful addition to the repertoire of clinical tools available for assessing accommodation in young children.

### Keywords

Accommodation; lag of accommodation; retinoscopy; infants; children

---

“Dynamic” retinoscopy techniques quantify accommodative lag by assessing the refractive state of the freely-accommodating eye at a single point in time. Traditional techniques include Nott retinoscopy (NR),<sup>1</sup> the monocular estimate method (MEM)<sup>2, 3</sup> and bell retinoscopy.<sup>4</sup> Clinical applications of measuring accommodation include optimizing refractive correction in children at risk of accommodative deficits, from Down syndrome,<sup>5, 6, 7</sup> cerebral palsy,<sup>8, 9</sup> or retinal disorders,<sup>10</sup> and evaluating asthenopia and convergence insufficiency.<sup>11</sup> Accommodative deficits may be important to identify in children with amblyopia,<sup>12, 13, 14, 15</sup> since hyperopic or bifocal correction could theoretically enhance treatment in these children.<sup>16, 17</sup> Accommodative lag may also help distinguish between well-compensated hyperopes and

those requiring correction to prevent esotropia or amblyopia;<sup>18, 19, 20</sup> a straight-eyed hyperope who is accommodating well may not need spectacles,<sup>17, 21</sup> whereas habitual under-accommodation in a hyperope may foreshadow esotropia<sup>22</sup> and/or bilateral amblyopia.<sup>23, 24</sup>

Dynamic retinoscopy is challenging, however, in young children. With a moving or inattentive child, it is difficult to introduce MEM lenses briefly enough not to influence accommodation. NR, using a fixed target, is inconvenient to perform in a typical clinic setting without an assistant. The examiner's withdrawal can distract the patient, and the retinoscopic reflex seen from a distance becomes hard to judge when lag is high. Hunter<sup>17</sup> proposed evaluating accommodation in children qualitatively by starting from a fixed retinoscope position, and advancing the target toward the child until the retinoscopic reflex is neutralized. The present article describes a similar, but quantitative procedure for assessing accommodation. Because the underlying principle is the same as that in bell retinoscopy,<sup>4</sup> the procedure is called modified bell retinoscopy (MBR). This study includes an assessment of the repeatability of measurements of accommodative lag in young children using MBR, and an evaluation of the comparability of MBR to other dynamic retinoscopy methods.

## METHODS

### Modified Bell Retinoscopy, Nott Retinoscopy and MEM

In MBR, the retinoscopist maintains a fixed distance from the child; the target is advanced until the retinoscopic reflex is neutralized (Figure 1). The target is attached to the end of retractable tape measure, which is attached to the retinoscope head, to measure the distance advanced. The accommodative demand, accommodative response, and accommodative lag are calculated from the known, fixed distance between retinoscope and child ( $D_2$  in Figure 1), and the final recorded distance between retinoscope and target.  $D_1$  in Figure 1 is the difference between  $D_2$  and the recorded retinoscope-target distance; the inverse of  $D_1$  is the dioptric viewing distance between child and target, and is henceforth termed the accommodative demand. When the retinoscopic reflex is neutralized, the dioptric distance between child and retinoscope (inverse of  $D_2$ ) is the accommodative response corresponding to that demand. The difference between accommodative demand and response is the accommodative lag.

The variable in MBR is the accommodative demand yielding a given accommodative response, with the response defined by  $D_2$ , set by the examiner. A measurement at one retinoscope position yields one point on the child's accommodative demand-response function, as in NR. NR differs from MBR only in that the demand is fixed by the examiner, and the corresponding response is the measured variable. The nomogram for MBR shown in Figure 2 illustrates how a given combination of retinoscope distance ( $D_2$ ) and target advancement corresponds to a given accommodative demand and associated lag of accommodation.

The MBR target, an internally-illuminated cube<sup>9</sup> with high-contrast black-and-white cartoon images containing a range of spatial frequencies, or letter/number optotypes for older children, was viewed in dim illumination, drawing attention to the illuminated images. The tape-measure on the retinoscope was used to establish the distance between child and examiner. Holding the target at the retinoscope, the examiner evaluated the retinoscopic reflex, eliciting the child's attention with verbal cues and by wiggling the target. "Against-motion" (accommodative lead) was recorded, but not quantified (the target could not be moved further away from the child than the retinoscope). If "with-motion" was observed, the target was first advanced toward the child until the motion was neutralized or reversed, then withdrawn again until "with-motion" was just seen. The final retinoscope-target distance was recorded. Secondary withdrawal of the target was designed to avoid overestimating lag, in children who were initially inattentive but then showed a step increase in accommodation in response to the moving target. Withdrawal identified the low neutral point once the child's attention was engaged.

The variable assessed in MBR is the accommodative demand, so different measurements from the same child recorded at a given retinoscope distance reflect not only different amounts of lag, but different accommodative demands (see Figure 2). Because lag may vary with accommodative demand, 9, 25 this study used normalized measurements. Lag was measured at three retinoscope distances (33 cm; 50 cm; 67 cm), giving three different points on the accommodative demand-response curve (targets rotated between measures to maintain interest). Regression was used to derive an accommodative demand-response function, and a lag estimate for a hypothetical 2.50 D demand. This is termed a “standardized 40 cm target estimate,” while the measurement obtained at a single retinoscope distance, e.g. 67 cm, is a “67 cm retinoscope measure.”

The nominal distance between child and retinoscopist may in reality vary, causing measurement variability in both accommodative response (calculated from the retinoscopy distance) and accommodative demand (calculated from the retinoscopy distance and the tape measure reading). Orthogonal rather than linear regression was therefore used to derive accommodative demand-response functions. 26

For NR, the same target was used, viewed at 40 cm in dim illumination. The examiner drew attention to it verbally and by wiggling the target’s fixed support, and withdrew until the retinoscopic reflex was neutralized, measuring the distance with a string. The accommodative response was the inverse of the final distance between child and examiner.

The MEM target consisted of stickers (range of spatial frequencies similar to the NR/MBR target images), and letter/number optotypes for older children, on a black background surrounding a central aperture, attached to the front of the retinoscope, with the examiner viewing through the aperture. This was presented at 40 cm with normal illumination. The examiner was masked to the lenses, drawing lenses from a shuffled set of 4 boxes each containing a 1.00 D range in 0.25 D increments. “Against-motion” with the first lens, or “with-motion” with the last, prompted selection of another box. The lowest-neutral lens value represented the accommodative lag.

Retinoscopy was performed in the horizontal meridian, usually the more hyperopic in the population of children studied.

### Subject Populations and Procedures

The “screening” population consisted of 5- to 23-month-old children in Early Head Start, Women Infants and Children centers, and community medical clinics in Los Angeles County. MBR was performed at retinoscope distances of 33 cm, 50 cm, and 67 cm; the sequence was repeated, yielding two standardized 40 cm target estimates. Cycloplegic (cyclopentolate/mydriacyl) retinoscopy was then performed.

The “clinic” population comprised 30 patients  $\leq 16$  years undergoing MBR on a routine visit, returning 3–30 days later for MBR, NR and MEM, performed in randomized sequence. On Day 2, the examiner was masked to Day 1 findings. Two measures were taken with each method (2 standardized 40 cm target estimates, for MBR). Children were usually tested uncorrected, maximizing lag from hyperopia and astigmatism, to assess a wide range of accommodative lags. The right eye was measured, with 3 exceptions for media opacity or strabismus. The same eye and spectacle-wear conditions were used for all methods on both days.

MBR was also performed in a cohort of 11 children with Down syndrome, without refractive correction, to evaluate repeatability in children with a high prevalence of accommodative deficiency.

This research followed the tenets of the Declaration of Helsinki, and was approved by the institutional review board of Childrens Hospital Los Angeles. After the nature of the study was explained, written informed consent for participation was obtained from parents, and assent was obtained from children who were able to provide assent.

### Statistical Analysis

The repeatability index (RI) of a measurement is the half-width of the 2-sided 95% agreement limits: 1.96 times the standard deviation (SD) of the inter-measurement difference. Stratified RI were calculated for lag  $<0.50$  D and lag  $\geq 0.50$  D. Corresponding 95% agreement limits are shown as horizontal dashed lines in Bland-Altman plots.

A regression method described by Bland and Altman<sup>27</sup> was used to describe inter-measurement agreement continuously as a function of measurement amplitude and other parameters of interest (see Appendix, Supplemental Digital Content 1, which describes the method in detail). This approach determines whether systematic inter-measurement difference (bias) and/or residual inter-measurement variability (repeatability) change as a function of a parameter of interest. It was used to (1) describe bias as a function of measurement amplitude when comparing dynamic retinoscopy methods, (2) assess inter-measurement agreement for MBR as a function of age and refractive error, and (3) assess inter-measurement agreement for MBR as a function of measurement amplitude. In the latter context, the analysis models the 95% limits of agreement as functions of measurement amplitude (sloping dotted lines in Bland-Altman plots), defining an amplitude-specific RI for any given lag (“modeled RI”).

The R programming environment was used for orthogonal regression analysis.<sup>28</sup> Excel was used for linear regression, and online software for t-tests and RI 95% confidence intervals (CI) (<http://graphpad.com/quickcalcs>).

## RESULTS

MBR was attempted in 203 children aged 5–23 months, a majority of whom (59%) were of Hispanic ethnicity. Six were uncooperative. Two sets of 3 retinoscope distances were completed in 176 children (87%), and two sets of  $\geq 2$  distances in 186 children (92%). Twelve children with unmeasurable lags (retinoscopic reflex not neutralizable with target advancement) and two showing leads were excluded, leaving a screening cohort of 172 children (median age 12 months). Ophthalmic findings were mostly normal (2 cases of strabismus, 5 of anisometropia). Mean right eye spherical equivalent (SE) refractive error was 0.68 D (SD 1.03 D).

The clinic cohort comprised 30 children (median age 60 months, range 15 months to 16 years) tested with MBR on two separate days, and NR and MEM on Day 2 (NR unavailable for one child). Amblyopia and/or motility disorders were present in 69% of children. Mean right eye SE was 1.99 D (SD 1.74 D).

The Down syndrome cohort comprised 11 children (median age 45 months); one was excluded for accommodative lead, and one did not cooperate with repeat measurement.

MBR accommodative demand-response functions in the screening cohort appeared linear, and minimally impacted by testing sequence (see Appendix, Supplemental Digital Content 1, for analysis of linearity and of the effect of testing sequence). The mean slope of the accommodative demand-response function was 0.98 (SD 0.22), not significantly different from 1.00 ( $t(171)=1.19$ ;  $p=0.23$ ). Seventy-eight percent of children had slopes between 0.80 and 1.20.

In pair-wise comparisons between clinic cohort measurements using MEM, NR and MBR (standardized 40 cm target estimates), all methods were correlated (Pearson's  $R=0.84$ ,  $0.82$ , and  $0.80$ , respectively, for MBR and NR, MBR and MEM, and NR and MEM) (Figure 3A–C). The difference between MBR and NR was independent of lag (Figure 3D) (see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement). MBR averaged  $0.32$  D (SD  $0.37$  D) less lag than NR ( $t(28)=4.66$ ;  $p<0.0001$ ). The difference between MBR and MEM or NR and MEM increased with increasing lag (Figure 3E–F) ( $p<0.0001$ ; see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement). For lag around  $1.00$  D, MEM estimates averaged  $0.78$  D higher than MBR and  $0.31$  D higher than NR; for  $2.00$  D lag, they averaged  $1.59$  D higher than MBR and  $1.14$  D higher than NR.

For MBR (standardized 40 cm target estimates), the impact of age and astigmatic and SE refractive error on within-visit inter-measurement agreement was studied in the screening cohort (Figure 4) (see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement). Neither bias nor repeatability varied with age or astigmatism.

Repeatability did not vary significantly with SE refractive error, but bias did ( $p=0.0005$ ); for higher SE, the second estimate showed slightly less lag on average than the first ( $0.13$  D less lag for a SE of  $2.00$  D).

To analyze within-visit repeatability of MBR as a function of lag, screening and clinic (Day 1) cohorts were pooled, excluding 2 clinic outliers (see Appendix, Supplemental Digital Content 1, which describes analysis justifying pooling of cohorts). Figure 5 shows the Bland-Altman plot for sequential standardized 40 cm target estimates. The stratified RI was  $0.33$  D (95% CI  $0.30$ – $0.38$  D) for lag  $<0.50$  D, and  $0.69$  D (95% CI  $0.58$ – $0.90$  D) for lag  $\geq 0.50$  D (dashed lines, Figure 5). Measurement repeatability decreased with increasing lag ( $p<0.0001$ ; see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement). Figure 5 shows 95% limits of agreement as a function of lag (sloping dotted lines). The modeled RI was  $0.49$  D for  $0.50$  D lag, and  $0.80$  D for  $1.00$  D lag. Bias did not vary with lag (middle dotted line, Figure 5) (see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement). The first estimate of lag slightly exceeded the second by an average of  $0.04$  D (SD  $0.24$  D) ( $t(199)=2.58$ ;  $p=0.01$ ).

For children with Down syndrome, within-visit MBR repeatability was comparable to other cohorts, after accounting for lag level. The overall RI was  $0.69$  D (95% CI  $0.47$ – $1.32$  D), for lag averaging  $1.48$  D (range  $0.92$ – $3.41$  D).

Within-visit repeatability of MBR was also assessed for individual 33 cm, 50 cm and 67 cm retinoscope measures, for which the overall (unstratified) within-visit RI (screening and clinic cohorts pooled) were  $1.24$  D (95% CI  $1.10$ – $1.35$  D),  $0.98$  D (95% CI  $0.88$ – $1.06$  D), and  $0.68$  D (95% CI  $0.61$ – $0.74$  D), respectively. The distribution of lag measured at each distance, meanwhile, was the same: mean (SD) lags were  $0.45$  D ( $0.75$  D),  $0.43$  D ( $0.65$  D), and  $0.45$  D ( $0.70$  D), respectively. Thus, for a given distribution of accommodative lags, greater proximity of the retinoscope resulted in decreased repeatability.

For the 67 cm retinoscope measure, stratified RI for lag  $<0.50$  D and  $\geq 0.50$  D were  $0.29$  D (95% CI  $0.25$ – $0.33$  D) and  $1.18$  D (95% CI  $0.98$ – $1.45$  D), respectively (Figure 6, dashed lines). The latter was higher than for standardized 40 cm target estimates, because of a wider distribution of lags. Repeatability decreased with increasing lag ( $p<0.0001$ ; see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement). The 95% agreement limits as a function of lag are shown in Figure 6 (sloping dotted lines). The modeled RI was  $0.53$  D for  $0.50$  D lag, and  $0.80$  D for  $1.00$  D lag (similar

to standardized 40 cm target estimates). Bias also varied with lag ( $p < 0.0001$ ; see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement); the second measurement yielded slightly higher lag estimates than the first, for high lags (bias of 0.34 D for 2.00 D lag).

The simple average of the 33 cm, 50 cm and 67 cm lag measures (2 outliers excluded) showed lower within-visit repeatability than the standardized 40 cm target estimate derived through regression using the same three measures. The modeled RI was 0.59 D for 0.50 D lag, and 1.03 D for 1.00 D lag.

Between-visit repeatability (standardized 40 cm target estimates) was analyzed in the clinic cohort. Figure 7 shows Bland-Altman plots for the first estimate of Day 1 compared to the first estimate of Day 2 (Panel A; two outliers excluded), and the second estimates of Day 1 and Day 2 (Panel B). For lag  $< 0.50$  D, the stratified between-visit RI were 0.45 D (95% CI 0.32–0.70) and 0.34 D (95% CI 0.24–0.54 D), for first and second estimates of each day, respectively. For lag  $\geq 0.50$  D, the stratified RI were 1.74 D (95% CI 1.22–2.81) and 0.87 D (95% CI 0.63–1.32 D) for first and second estimates of each day, respectively. Repeatability for the first estimate of the day decreased with increasing lag (Panel A, sloping 95% limits of agreement) ( $p < 0.0001$ ; see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement). The modeled RI was 0.84 D for 0.50 D lag, and 1.37 D for 1.00 D lag. For the second estimate of each day, the relationship between repeatability and lag was not quite significant (see Table, Supplemental Digital Content 2, detailing analysis of variability of inter-measurement agreement), and repeatability was better: the modeled RI was 0.60 D for 0.50 D lag, and 0.77 D for 1.00 D lag, similar to within-visit repeatability.

To explore why between-visit repeatability differed for the first and second measurements of each day, all six pair-wise comparisons of the four standardized 40 cm target estimates obtained over 2 days were examined. The three comparisons involving the first Day 2 estimate all showed poor repeatability (overall RI 1.01–1.23 D); the other three showed better repeatability (overall RI 0.57–0.70 D), indicating that the first Day 2 estimate was anomalous.

Between-visit repeatability was worse for 67 cm retinoscope measures than for standardized 40 cm target estimates: for the first or second measure of each day, or their average, the modeled RI was 0.83–0.84 D for 0.50 D lag, and 1.23–1.24 D for 1.00 D lag.

Simulation was performed to assess the effect of spatial measurement error on repeatability. Assuming no source of measurement variability other than error in the presumed distance between child and retinoscope (D2 in Figure 1), 95% inter-measurement agreement limits were modeled as a function of lag. For a nominal 67 cm retinoscope distance, the observed slopes of the agreement limits in the within-visit repeatability data for 67 cm retinoscope measures (slopes of dotted lines in Figure 6, adjusted for bias function) were matched by a model simulating 10% error (i.e., 95% of actual values for D2 falling within  $\pm 10\%$  of 67 cm).

## DISCUSSION

The modified bell retinoscopy (MBR) technique is easy to use in young children. Lag estimates correlate strongly with results using other dynamic retinoscopy techniques, over a wide range of accommodative lags. Measurements are reproducible, and although repeatability decreases at higher levels of accommodative lag, it is independent of age and refractive error.

In a series of clinic patients enriched for ophthalmic pathology and tested without correction, ensuring a broad and clinically relevant range of accommodative lags, lag estimates by MBR were highly correlated with findings from Nott retinoscopy (NR) and the monocular estimate method (MEM). MBR results were most similar to NR. Both MBR and NR yielded lower

estimates of lag than MEM. The differences were greater with higher amounts of lag. This is similar to previous observations comparing MEM and NR.<sup>29</sup> A study finding that NR and MEM were equivalent<sup>30</sup> did not include any children with  $\geq 1.00$  D of lag. Normally, cognitively demanding targets are recommended for MEM; with the less demanding targets used in this study, geared toward younger children, MEM may lead to greater relaxation of accommodation than other methods, and be more influenced by latent hyperopia. Because MEM, unlike MBR and NR, was performed in normal illumination, smaller pupil size may also have influenced accommodative accuracy. Compared to NR, MBR underestimated lag by a third of a diopter, independent of the level of lag. Target movement in MBR may elicit more interest than the static target of NR, or directly stimulate accommodation through looming cues.<sup>31</sup> The secondary withdrawal of the target until “with-motion” is seen may contribute to systematically lower lag estimates compared to NR, and also may bias measurements toward the child’s best performance in children with variable accommodation and/or attention.

The repeatability of MBR was assessed over a wide range of accommodative lags, in children from a screening cohort of mostly normal children, in clinic patients with a broad range of disorders, and in a cohort of children with Down syndrome. Repeatability decreased with higher lag, which, for a given retinoscope position, meant greater target proximity to the patient. Repeatability also worsened with retinoscope proximity to the patient; individual measures at 67 cm were more repeatable than individual measures at 33 cm. Spatial measurement error may explain both findings, because movement of the child introduces error in the nominal retinoscope distance ( $D_2$  in Figure 2) that is used to calculate lag, and the diopter-equivalent error is greater nearer the child. Indeed, modeling revealed that just  $\pm 10\%$  variation in the child’s nominal distance from the examiner accounts for the observed relationship between lag and repeatability of 67 cm retinoscope lag measures. The repeatability of MBR would likely be greater, especially for high lags, if the actual distance between child and examiner could be measured simultaneously with the target-retinoscope distance when retinoscopic reflex neutrality is observed.

Other factors may contribute to lower repeatability of high lag measures. Poor accommodation may be associated with intrinsically greater variability of accommodation. Microfluctuations in accommodative response increase with the stimulus to accommodation,<sup>32, 33</sup> and in MBR, higher lag measurements correspond to higher stimulus demands. There is also measurement ambiguity in the width of the zone of retinoscopic reflex neutrality;<sup>34</sup> however, while this may contribute to inter-measurement variability, the effect is the same at all levels of lag, because retinoscopy conditions are the same for all measurements.

McLelland and Saunders reported between-visit RI for a single NR measurement to be 0.56 D for a 25 cm target (mean lag 0.72 D), and 1.09 D for a 17 cm target (mean lag 1.43 D), in adults and children aged 6 years and older.<sup>35</sup> The within-visit RI of a standardized 40 cm target MBR estimate was similar for comparable levels of lag (0.49 D for 0.50 D lag, and 0.80 D for 1.00 D lag). The between-visit RI for MBR was also in a similar range, when considering the second estimate of each day (0.60 D for 0.50 D lag, and 0.77 D for 1.00 D lag). Between-visit repeatability for MBR was worse for the first measure of each day, however, because of anomalous MBR performance on the first measurement on Day 2. That measurement may have been impacted by the change in testing circumstances on Day 2, when multiple testing methods were intermingled. After adjusting for the level of lag, repeatability was similar for normal children, clinic patients and children with Down syndrome, supporting the clinical relevance of these results. Although behavioral changes that might affect attention occur between 5 and 23 months of age, MBR was equally repeatable in infants and toddlers. Astigmatism might be predicted to impact repeatability if astigmatic children vary their focus over the astigmatic interval,<sup>36</sup> but repeatability did not vary measurably with astigmatism. SE refractive error might impact repeatability if hyperopes sometimes show lag and sometimes not, but no such

effect was seen in the screening cohort comprising mostly normal children. Hyperopes showed slightly less lag on average on the second estimate, however, possibly indicating a practice effect.

The standardized 40 cm target estimate used in this study was derived from three measures at different retinoscope distances, and required data processing. Therefore, repeatability was also assessed for the lag estimate from a single retinoscope distance. The within-visit RI for 67 cm retinoscope measures, adjusted for lag, was similar to standardized 40 cm target estimates. Between-visit repeatability, however, was lower. Repeatability was also assessed for the simple average of lag estimates from three retinoscope distances; it was lower than for the standardized 40 cm target estimate derived by regression using data from the same three retinoscope distances. The use of orthogonal rather than linear regression for accommodative demand-response functions, however, was not critical (see Appendix, Supplemental Digital Content 1, comparing findings using both approaches).

MBR, unlike MEM, requires no lenses, with less attendant risk of modifying the accommodative response while measuring it. Unlike NR, MBR does not require an assistant or free-standing target. The target is the only moving object, which helps maintain the child's attention, and it is easier to judge the retinoscopic reflex in children with large lags, because the examiner remains near the child. MBR was not designed to quantify accommodative leads, but could be adapted to do so; however, as with NR, positioning the retinoscopist between the child and the target to measure leads might distract the child.

In MBR, the accommodative demand is not fixed across measurements, whereas in NR and MEM, the accommodative demand is fixed. Hence, a standardized 40 cm target estimate derived from the accommodative demand-response function was employed here. However, the accommodative demand-response functions of normal children in this study had slopes near 1.00, consistent with classic studies in infants and young children.<sup>37,38</sup> Slopes near 1.00 imply that changes in accommodative demand drive equal changes in output, i.e., lag varies little over a range of demands, and lag measured at one demand will be representative of lag at other levels of accommodative demand.

For routine clinical purposes, then, the simplest way to use MBR is to consult a nomogram (Figure 2) to convert a raw measurement (distance of target advancement) directly into a measure of accommodative lag. For a 50 cm working distance, if the target must be advanced 10 cm for retinoscopic reflex neutrality for one child, this corresponds to 0.50 D of lag, at a demand of 2.50 D. If, for another child, the target must be advanced around 21 cm, this child clearly has more than 0.50 D lag at a demand of 2.50 D (more lag than the first child), because the reflex was not neutralized with only 10 cm target advancement. The second child shows 1.50 D of lag, measured at a demand of 3.50 D. Regardless of the difference in accommodative demands, this observation again indicates that he accommodates less well than the first child, because, as discussed above, children typically show similar lags of accommodation over a range of demands. Using single measures at one retinoscope distance together with the nomogram of Figure 2, it is easy to assess lag relative to a threshold level of lag deemed to be of concern; if the target must be advanced more than the distance corresponding (for the examiner's chosen working distance) to that threshold, then the child's accommodative lag is greater than the threshold of interest.

Accommodative demand-response function slopes may not approach 1.00 in children with very poor accommodation. In addition, as discussed above, the between-visit repeatability of single retinoscope measures taken at one distance is lower than for 40 cm target estimates derived from accommodative demand-response functions. Thus, for precise quantification, especially



of high lags, calculated 40 cm target estimates of lag are superior to single measures at one retinoscope distance.

A few children in this study had unquantifiable lag, because the retinoscopic reflex could not be neutralized with any degree of target advancement. Some of these children might have been measurable using the less hyperopic meridian. In an attentive child, inability to neutralize the retinoscopic reflex despite using the less hyperopic meridian is evidence of high lag; for example, if at a 67 cm working distance, the retinoscopic reflex is not neutralized even when the target is 33 cm from the child (3.00 D demand), then the lag in the measured meridian must be at least 1.50 D or greater.

In conclusion, this study describes a new technique for quantifying accommodative lag, validated through comparison with existing methods, and showing comparable repeatability to other methods. For clinical purposes, a nomogram is useful for distinguishing between normal and deficient accommodation on the basis of how far a target must be advanced toward the subject to achieve retinoscopic reflex neutrality. Modified bell retinoscopy may be a useful addition to existing tools for assessing accommodation, especially in infants and young children.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This research was funded by NEI 5 K23 EY016699. Velma Dobson, Rowan Candy and Joseph Miller provided valuable input during the preparation of the manuscript. The artwork in Figure 1 is by Joel Schechter, PhD.

## References

1. Nott I. Dynamic skiametry, accommodation and convergence. *Am J Physiol Opt* 1925;6:490–503.
2. Rouse MW, London R, Allen DC. An evaluation of the monocular estimate method of dynamic retinoscopy. *Am J Optom Physiol Opt* 1982;59:234–9. [PubMed: 7072827]
3. Pascal JJ. Neutralization in dynamic retinoscopy. *Br J Ophthalmol* 1931;15:589–90. [PubMed: 18168988]
4. Apell RJ. Clinical application of bell retinoscopy. *J Am Optom Assoc* 1975;46:1023–7. [PubMed: 1078340]
5. Woodhouse JM, Meades JS, Leat SJ, Saunders KJ. Reduced accommodation in children with Down syndrome. *Invest Ophthalmol Vis Sci* 1993;34:2382–7. [PubMed: 8505220]
6. Haugen OH, Hovding G, Lundstrom I. Refractive development in children with Down's syndrome: a population based, longitudinal study. *Br J Ophthalmol* 2001;85:714–9. [PubMed: 11371494]
7. Stewart RE, Woodhouse JM, Trojanowska LD. In focus: the use of bifocal spectacles with children with Down's syndrome. *Ophthalmic Physiol Opt* 2005;25:514–22. [PubMed: 16343127]
8. McClelland JF, Parkes J, Hill N, Jackson AJ, Saunders KJ. Accommodative dysfunction in children with cerebral palsy: a population-based study. *Invest Ophthalmol Vis Sci* 2006;47:1824–30. [PubMed: 16638987]
9. Leat SJ, Gargon JL. Accommodative response in children and young adults using dynamic retinoscopy. *Ophthalmic Physiol Opt* 1996;16:375–84. [PubMed: 8944182]
10. Leat SJ, Mohr A. Accommodative response in pre-presbyopes with visual impairment and its clinical implications. *Invest Ophthalmol Vis Sci* 2007;48:3888–96. [PubMed: 17652765]
11. Marran LF, De Land PN, Nguyen AL. Accommodative insufficiency is the primary source of symptoms in children diagnosed with convergence insufficiency. *Optom Vis Sci* 2006;83:281–9. [PubMed: 16699440]

12. Abraham SV. Accommodation in the amblyopic eye. *Am J Ophthalmol* 1961;52:197–200. [PubMed: 13681091]
13. Wood IC, Tomlinson A. The accommodative response in amblyopia. *Am J Optom Physiol Opt* 1975;52:243–7. [PubMed: 1130487]
14. Ciuffreda KJ, Rumpf D. Contrast and accommodation in amblyopia. *Vision Res* 1985;25:1445–57. [PubMed: 4090279]
15. Ukai K, Ishii M, Ishikawa S. A quasi-static study of accommodation in amblyopia. *Ophthalmic Physiol Opt* 1986;6:287–95. [PubMed: 3822469]
16. Guyton DL, O'Connor GM. Dynamic retinoscopy. *Curr Opin Ophthalmol* 1991;2:78–80. [PubMed: 10149292]
17. Hunter DG. Dynamic retinoscopy: the missing data. *Surv Ophthalmol* 2001;46:269–74. [PubMed: 11738434]
18. Ingram RM, Arnold PE, Dally S, Lucas J. Results of a randomised trial of treating abnormal hypermetropia from the age of 6 months. *Br J Ophthalmol* 1990;74:158–9. [PubMed: 2182103]
19. Atkinson J, Braddick O, Robier B, Anker S, Ehrlich D, King J, Watson P, Moore A. Two infant vision screening programmes: prediction and prevention of strabismus and amblyopia from photo- and videorefractive screening. *Eye* 1996;10(Pt 2):189–98. [PubMed: 8776448]
20. Tarczy-Hornoch K. The epidemiology of early childhood hyperopia. *Optom Vis Sci* 2007;84:115–23. [PubMed: 17299341]
21. Kaakinen K, Kaseva H, Kause ER. Mass screening of children for strabismus or ametropia with two-flash photoskiagraphy. *Acta Ophthalmol (Copenh)* 1986;64:105–10. [PubMed: 3962614]
22. Ingram RM, Gill LE, Goldacre MJ. Emmetropisation and accommodation in hypermetropic children before they show signs of squint—a preliminary analysis. *Bull Soc Belge Ophthalmol* 1994;253:41–56. [PubMed: 7633630]
23. Friedman Z, Neumann E, Abel-Peleg B. Outcome of treatment of marked ametropia without strabismus following screening and diagnosis before the age of three. *J Pediatr Ophthalmol Strabismus* 1985;22:54–7. [PubMed: 3989639]
24. Friedman Z, Neumann E, Hyams SW, Peleg B. Ophthalmic screening of 38,000 children, age 1 to 2 1/2 years, in child welfare clinics. *J Pediatr Ophthalmol Strabismus* 1980;17:261–7. [PubMed: 7411341]
25. McClelland JF, Saunders KJ. Accommodative lag using dynamic retinoscopy: age norms for school-age children. *Optom Vis Sci* 2004;81:929–33. [PubMed: 15592117]
26. Parish RC. Comparison of linear regression methods when both variables contain error: relation to clinical studies. *Dicp* 1989;23:891–98. [PubMed: 2596133]
27. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135–60. [PubMed: 10501650]
28. R Development Core Team. R: A Language and Environment for Statistical Computing [computer program]. Vienna, Austria: R Foundation for Statistical Computing; 2008.
29. del Pilar Cacho M, Garcia-Munoz A, Garcia-Bernabeu JR, Lopez A. Comparison between MEM and Nott dynamic retinoscopy. *Optom Vis Sci* 1999;76:650–5. [PubMed: 10498007]
30. Locke LC, Somers W. A comparison study of dynamic retinoscopy techniques. *Optom Vis Sci* 1989;66:540–4. [PubMed: 2771345]
31. McLin LN Jr, Schor CM, Kruger PB. Changing size (looming) as a stimulus to accommodation and vergence. *Vision Res* 1988;28:883–98. [PubMed: 3250084]
32. Kotulak JC, Schor CM. Temporal variations in accommodation during steady-state conditions. *J Opt Soc Am (A)* 1986;3:223–7. [PubMed: 3950795]
33. Stark LR, Atchison DA. Pupil size, mean accommodation response and the fluctuations of accommodation. *Ophthalmic Physiol Opt* 1997;17:316–23. [PubMed: 9390376]
34. Boeder P, Kolder HE. Neutralization at infinity in streak retinoscopy. *Arch Ophthalmol* 1984;102:1396–9. [PubMed: 6477256]
35. McClelland JF, Saunders KJ. The repeatability and validity of dynamic retinoscopy in assessing the accommodative response. *Ophthalmic Physiol Opt* 2003;23:243–50. [PubMed: 12753480]

36. Stark LR, Strang NC, Atchison DA. Dynamic accommodation response in the presence of astigmatism. *J Opt Soc Am (A)* 2003;20:2228–36.
37. Haynes H, White BL, Held R. Visual accommodation in human infants. *Science* 1965;148:528–30. [PubMed: 14263778]
38. Banks MS. Infant refraction and accommodation. *Int Ophthalmol Clin* 1980;20:205–32. [PubMed: 7399818]

## APPENDIX 1

### Analysis of Variation in Inter-Measurement Agreement

Variation in inter-measurement agreement as a function of measurement amplitude was analyzed according to the regression method of Bland and Altman (manuscript reference 27). First, linear regression was performed for the difference between measures against their average, describing systematic bias between first and second measurements as a function of measurement amplitude. The result, henceforth termed the “bias function,” is the regression fit to a standard Bland-Altman plot (e.g. middle dotted line in Figure 5 of manuscript). If the slope of this function is significantly different from zero, bias varies with measurement amplitude. Otherwise, the mean bias can be described as a constant, the average of all inter-measurement differences, and a t-test can be used to assess its significance.

The inter-measurement differences predicted by the bias function are then subtracted from those actually observed. The resulting residuals take into account any systematic differences between the measurements being compared, and thus reflect inter-measurement variability alone. Linear regression is performed for the absolute values of the residuals against measurement amplitude. The resulting function, herein termed the “variability function,” describes inter-measurement variability as a function of measurement amplitude. If the y-dimension scatter on a Bland-Altman plot increases along the x-axis, this is reflected as a positive slope of the variability function. If the slope is significantly different from zero, repeatability varies with measurement amplitude.

For any given measurement amplitude, the mean inter-measurement difference is predicted by the bias function. The SD around this mean is predicted by the variability function, multiplied by a factor of  $\sqrt{(\pi/2)}$ . For any given measurement amplitude, an amplitude-specific RI is defined according to the definition of RI: 1.96 times the SD of the inter-measurement difference.

Graphically, the 95% limits of agreement are represented by linear functions describing (mean difference +1.96 SD) and (mean difference – 1.96 SD) as a function of measurement amplitude (e.g., sloping dotted lines in Figure 5 of manuscript).

The same methodology is applicable to the study of inter-measurement variability as a function not only of measurement amplitude, but of any parameter of interest, and thus was also used to assess whether repeatability varied significantly with age or refractive error. The bias functions described above were also assessed for inter-measurement differences as a function of measurement amplitude when comparing lag measurements using different dynamic retinoscopy methods.

The Table in Supplemental Digital Content 2 summarizes the findings of these analyses, and the tests of whether the slopes of bias and variability functions are significantly different from zero.

## Impact of Testing Sequence; Verification of Linearity of Accommodative Demand-Response Functions

The testing sequence for different retinoscope distances (33 cm, 50 cm, 67 cm, 33 cm, 50 cm, 67 cm) might impact accommodative demand-response slopes through fatigue or practice effects. The mean difference between successive lag estimates at a given retinoscope distance was assessed in the screening cohort. It was 0.15 D (SD 0.59 D), 0.04 D (SD 0.41 D), and -0.02 D (SD 0.32 D), for 33 cm, 50 cm and 67 cm retinoscope distances, respectively. The 33 cm position showed a small but significant decrease in lag with repeated testing ( $t(163)=3.26$ ;  $p=0.001$ ), but other positions did not ( $p>0.20$ ).

To evaluate the linearity of accommodative demand-response functions, an average “low-response” slope was calculated for each child in the screening cohort, using the data from 67 cm and 50 cm retinoscope distances, and a “high-response” slope using data from 50 cm and 33 cm retinoscope distances. The mean difference between individual children’s low-response and high-response slopes was 0.04 (s.d. 1.11), not significantly different from zero ( $t(167)=-0.47$ ;  $p=0.64$ ), indicating no systematic non-linearity in accommodative demand-response functions.

## Repeatability Analysis: Pooling of Screening and Clinic Cohorts

First and second MBR standardized 40 cm target estimates the same day were correlated, for both screening and clinic (Day 1) cohorts: Pearson’s  $R=0.79$  ( $t(170)=16.7$ ,  $p<0.0001$ ) and  $R=0.86$  ( $t(26)=8.7$ ,  $p<0.0001$ ), respectively. Two outliers whose cooperation level changed between measurements were excluded from the clinic cohort. Measured with 4 D and 2.4 D uncorrected hyperopia, respectively, their initial Day 1 lag estimates were 3.8 D and 4.4 D, while subsequent Day 1 estimates and all Day 2 estimates were between 1.1 and 1.7 D.

Stratified according to the amount of lag, the screening and clinic cohorts’ repeatability indices (RI) were comparable: 0.33 D (95% CI 0.30–0.38 D) and 0.29 D (95% CI 0.20–0.45 D), respectively, for lag  $<0.5$  D, and 0.61 D (95% CI 0.51–0.77 D) and 0.90 D (95% CI 0.64–1.42 D), respectively, for lag  $\geq 0.5$  D. The cohorts were therefore pooled for further analysis.

## Comparison of Orthogonal and Linear Regression for Standardized 40cm Target Estimates

Orthogonal regression was used for the primary analysis of accommodative demand-response functions, to derive standardized 40 cm target estimates, but ordinary linear regression of accommodative response (y) on accommodative demand (x) was also performed for comparison. In the screening cohort (first estimate), the mean difference between lag estimates using linear regression and orthogonal regression was 0.004 D (SD 0.016 D). The methods agreed within  $\pm 0.03$  D. For clinic cohort measurements (Day 1 or Day 2, first or second estimate), the mean difference between estimates using linear regression and orthogonal regression (excluding one outlier) ranged from 0.01 D to 0.03 D (SD range 0.06 D to 0.12); the methods agreed within  $\pm 0.12$  D to  $\pm 0.23$  D.

The within-visit inter-measurement RI for screening and clinic cohorts pooled, using linear instead of orthogonal regression, was 0.35 D (95% CI 0.31–0.39 D) for lag  $<0.5$  D, and 0.65 (95% CI 0.55–0.80) for lag  $\geq 0.5$  D, similar to results using orthogonal regression.

## APPENDIX 2

Table A1

Analysis of inter-measurement variability.

Lag measurements compared	Parameter studied	Slope of bias function	Slope of bias function, test of significant difference from zero			Slope of variability function	Slope of variability function, test of significant difference from zero		
			t	Dof	p		t	Dof	p
MBR and NR	Mean lag	-0.003	0.02	27	0.98	NA			
MBR and MEM	Mean lag	-0.82	8.2	28	<0.0001	NA			
NR and MEM	Mean lag	-0.82	7.7	27	<0.0001	NA			
MBR (standardized 40 cm target estimates): first and second measurements <i>within</i> visit	Age	-0.0002	0.06	170	0.95	-0.004	1.6	170	0.11
	Astigmatism	0.02	0.8	170	0.45	0.01	0.6	170	0.53
	SE	-0.05	3.5	170	<b>0.0005</b>	0.02	1.8	170	0.07
MBR (standardized 40 cm target estimates): first and second measurements <i>within</i> visit	Mean lag	-0.01	0.2	198	0.81	0.25	9.8	198	<0.0001
MBR (67 cm retinoscope measures): first and second measurements <i>within</i> visit	Mean lag	0.22	6.9	190	<0.0001	0.22	12.5	190	<0.0001
MBR (standardized 40 cm target estimates): Day 1 and Day 2 ( <i>first</i> measurement of each visit)	Mean lag	0.33	1.9	26	0.06	0.43	6.3	26	<0.0001
MBR (standardized 40 cm target estimates): Day 1 and Day 2 ( <i>second</i> measurement of each visit)	Mean lag	-0.08	0.8	28	0.43	0.11	1.8	28	0.09

MBR: modified bell retinoscopy

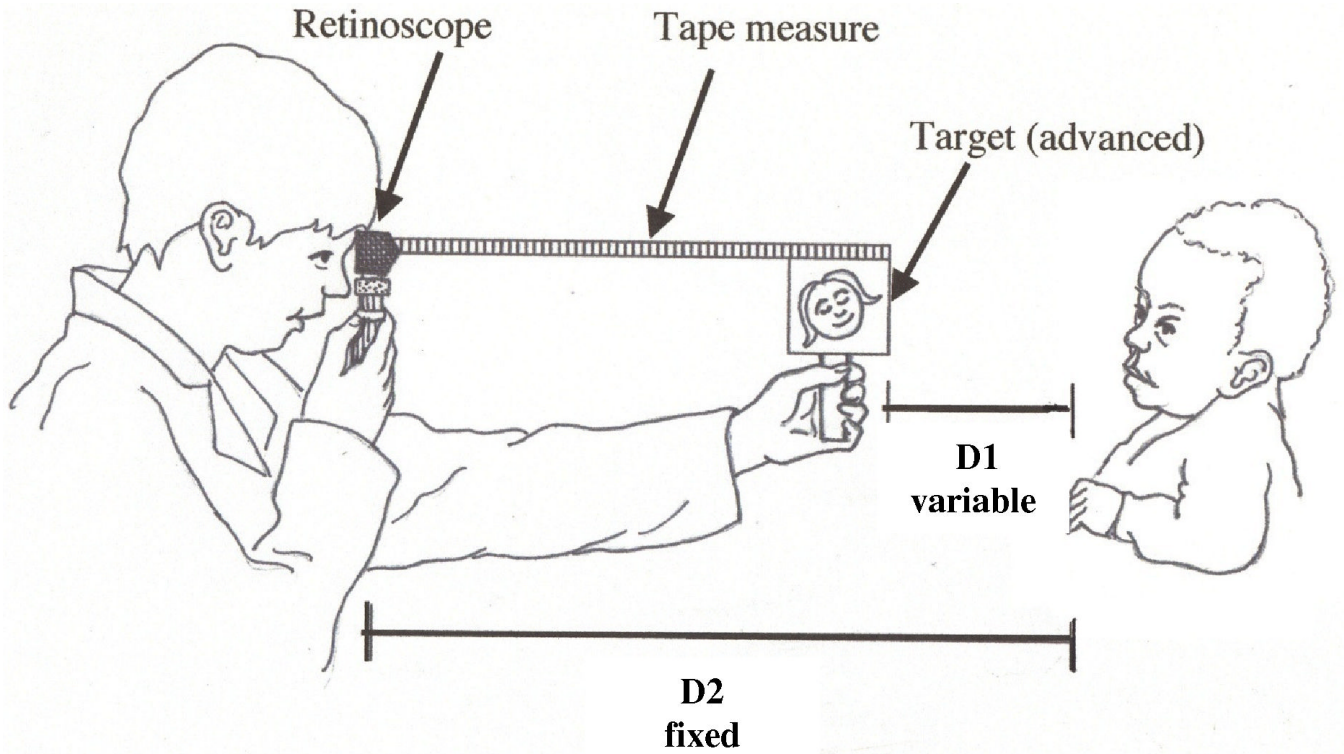
NR: Nott retinoscopy

MEM: monocular estimate method

NA: not applicable

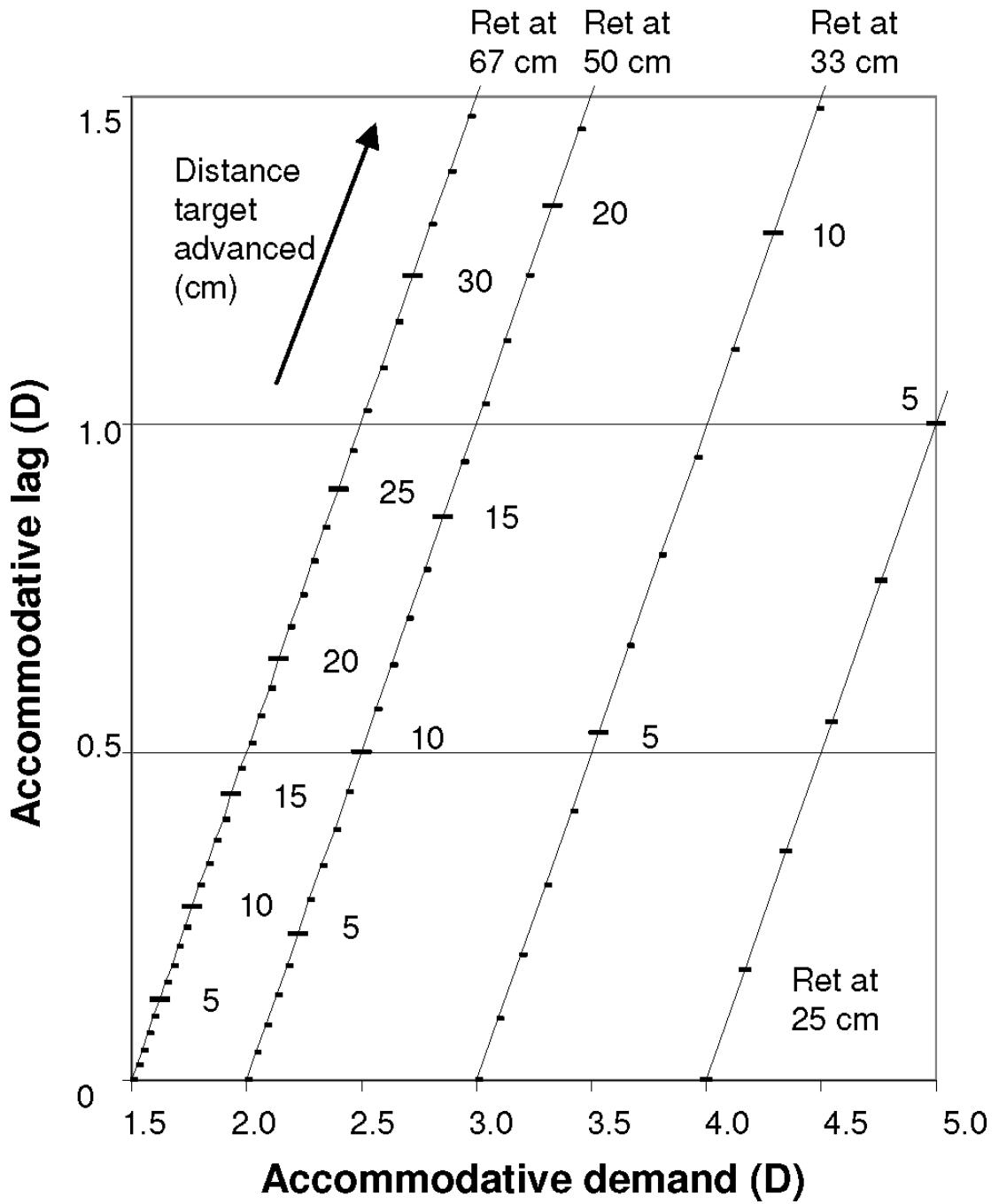
SE: spherical equivalent refractive error

Dof: degrees of freedom.

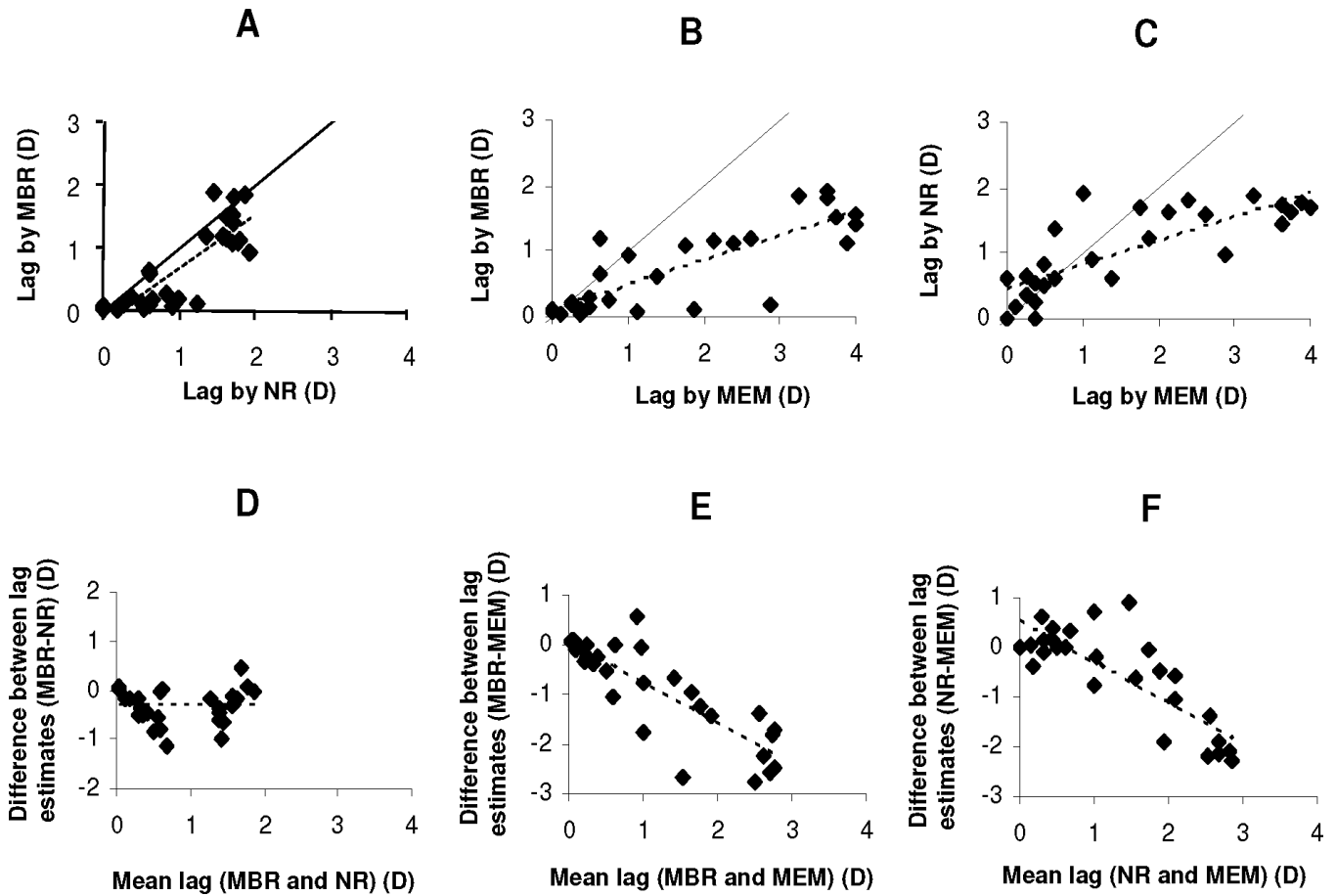


**Figure 1.**

Modified bell retinoscopy technique. D1: distance between child and target, corresponding to accommodative demand; this is variable, because the target is advanced until retinoscopic reflex neutrality is observed. D2: distance between child and retinoscope, corresponding, at reflex neutrality, to the accommodative response; this distance is set by the examiner at the start of measurement.



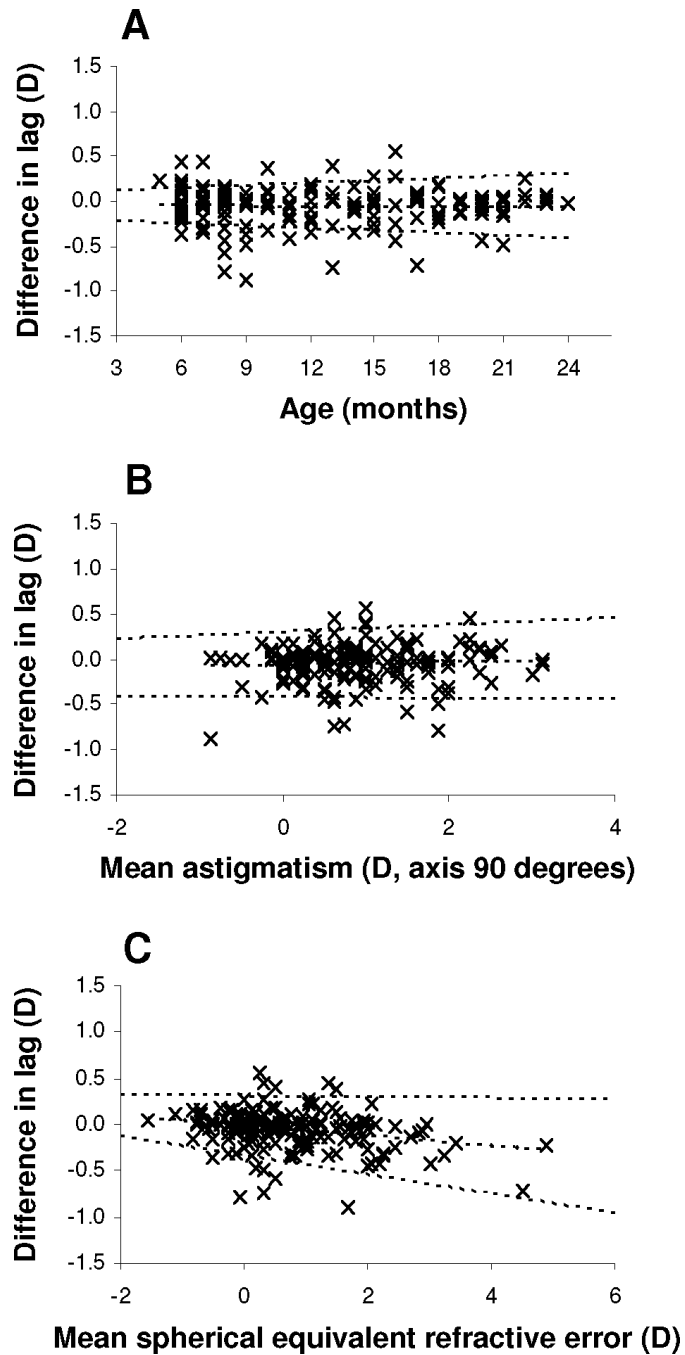
**Figure 2.** Modified bell retinoscopy nomogram. Relationship between the distance that the target must be advanced from the retinoscope to achieve retinoscopic reflex neutrality (recorded from tape measure; see Figure 1), the accommodative demand (corresponding to D1 in Figure 1), and the accommodative lag, for four different retinoscope (“Ret”) distances (corresponding to D2 in Figure 1).



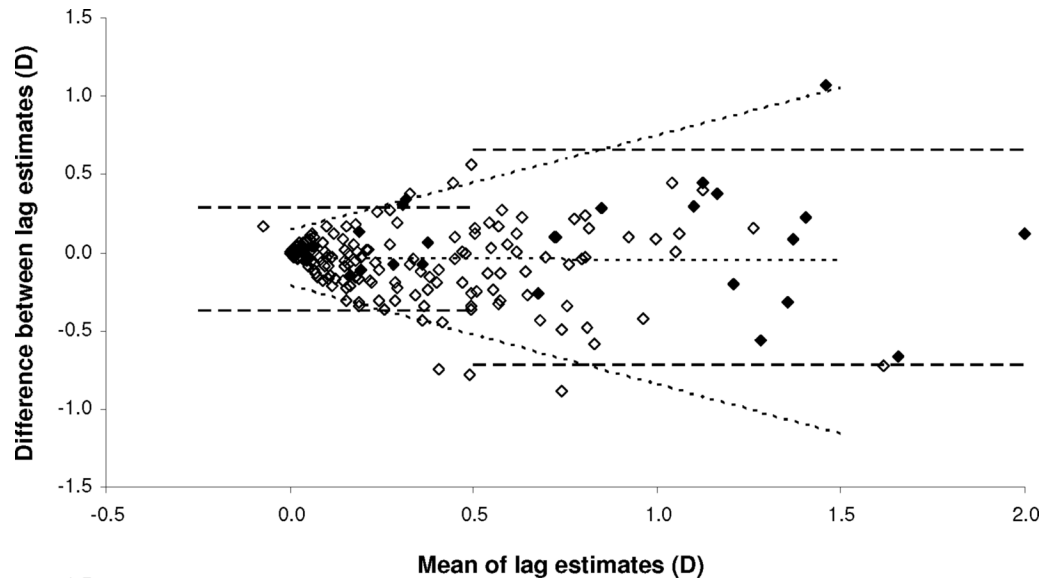
**Figure 3.**

Comparison of modified bell retinoscopy (MBR), Nott retinoscopy (NR), and monocular estimate method (MEM). (A–C) correlation between lag measured by MBR and NR, MBR and MEM, and NR and MEM; solid lines indicate line of perfect agreement; dotted lines show linear regression fits. Panels D–F: Bland-Altman plots of inter-method difference in lag plotted against mean lag measurement for MBR compared to NR, MBR compared to MEM, and NR compared to MEM; dotted lines indicated linear regression fits (bias functions).

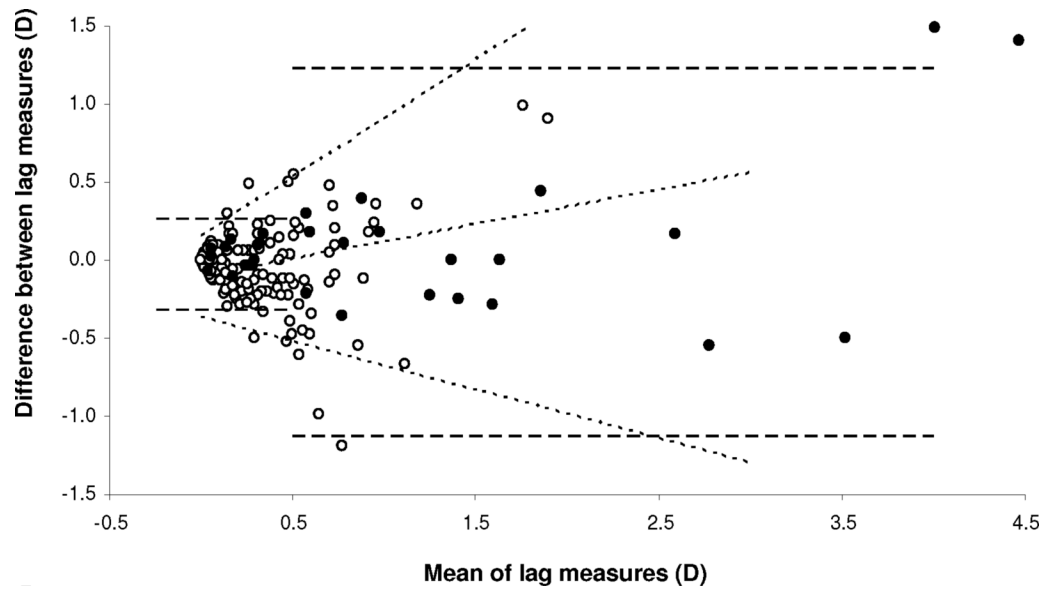




**Figure 4.** Repeatability of modified bell retinoscopy as a function of age and refractive error. Bland-Altman plots of differences in lag estimates (standardized 40 cm target estimates; second estimate minus first estimate) plotted against age (Panel A), cylindrical refractive error (axis 90 degrees) (Panel B), and spherical equivalent refractive error (Panel C), for the screening cohort. Middle dotted lines: bias functions; outer dotted lines: linear functions describing 95% limits of agreement as a function of age or refractive error.

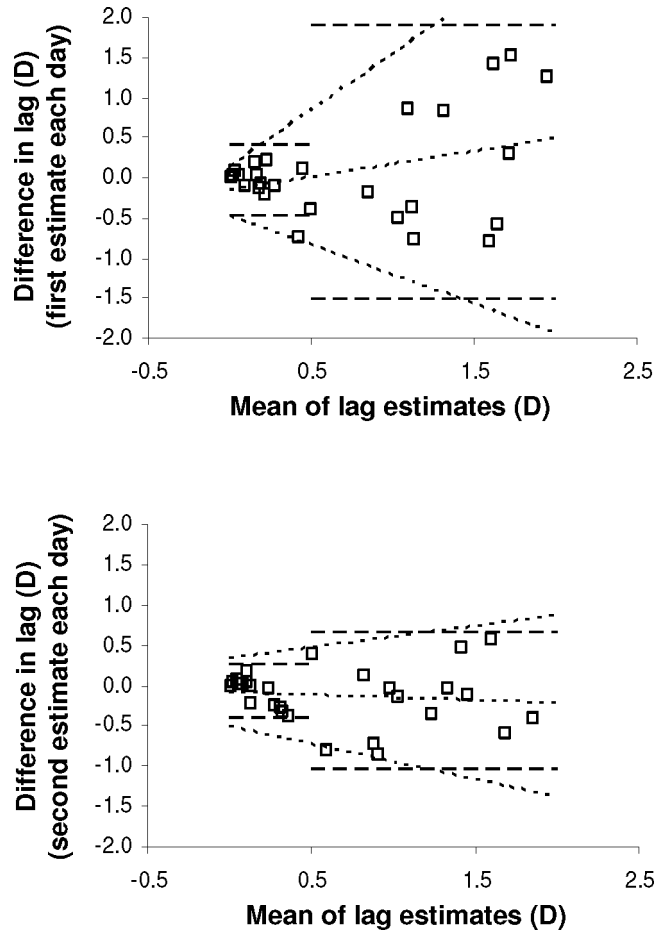


**Figure 5.** Within-visit repeatability of modified bell retinoscopy (standardized 40 cm target estimates). Bland-Altman plot of differences in lag estimates (standardized 40 cm target estimates; second estimate minus first estimate) plotted against the mean of two estimates. Dashed lines: 95% limits of agreement from stratified analysis, for mean lag above and below 0.50 D; middle dotted line: bias function; outer dotted lines: linear functions describing 95% limits of agreement as a function of mean lag; open symbols: screening cohort; filled symbols: clinic cohort.



**Figure 6.**

Within-visit repeatability of modified bell retinoscopy (67 cm retinoscope measures). Bland-Altman plot of inter-measurement differences in lag estimates (67 cm retinoscope measures; second measure minus first measure) plotted against the mean of two measurements. Dashed lines: 95% limits of agreement from stratified analysis, for mean lag above and below 0.50 D; inner dotted line: bias function; outer dotted lines: linear functions describing 95% limits of agreement as a function of mean lag; open symbols: screening cohort; filled symbols: clinic cohort. Note that scale of x-axis differs from Figure 5.



**Figure 7.** Between-visit repeatability of modified bell retinoscopy (standardized 40 cm target estimates). **(A)** Between-visit repeatability for the first estimate of each day in the clinic cohort. **(B)** Between-visit repeatability for the second estimate of each day in the clinic cohort. Difference in lag: Day 2 estimate minus Day 1 estimate. Dashed lines: 95% limits of agreement from stratified analysis, for mean lag above and below 0.50 D; middle dotted line: bias function; outer dotted lines: linear functions describing 95% limits of agreement as a function of mean lag.