Compensation Comparison Method for Assessment of Retinal Straylight

Luuk Franssen, Joris E. Coppens, and Thomas J. T. P. van den Berg

Purpose. Presently, no instrument or method exists that is generally accepted for routine clinical assessment of (functional) retinal straylight. Yet retinal straylight is the cause of major patient complaints, such as hindrance from glare and loss of contrast. It results from disturbances in the optical media that increase light–scattering over angles of 1° to 90°. Its assessment would help to decide whether to perform surgery for (early) cataract and would help in the evaluation of corneal or vitreal turbidity.

Methods. The psychophysical technique of the "direct compensation" method was adapted to make it suitable for routine clinical assessment. In the new approach, called "compensation comparison, " the central test field is subdivided into two half fields: one with and one without counterphase compensation light. The subject's task is a forced-choice comparison between the two half fields, to decide which half flickers more strongly. A theoretical form for the respective psychometric function was defined and experimentally verified in a laboratory experiment involving seven subjects, with and without artificially increased light scattering. The method was applied in a separate multicenter study. Its reliability was additionally tested with a commercial implement (C-Quant; Oculus Optikgeräte, Wetzlar-Dutenhofen, Germany).

Results. A repeated-measures SD of 0.07 log units was achieved, to be compared with differences in the young normal population of 0.4 log units and an increase with healthy aging by 0.5 log units at 80 years and by 1.0 or more log units with (early) cataract or corneal disturbances. Reliability was further found to be high when using the commercial version of the method.

Conclusions. The compensation comparison method for measuring retinal straylight is suited for clinical use to diagnose patients with complaints caused by large angle light scattering in the eye such as early cataract. (*Invest Ophthalmol Vis Sci.* 2006;47:768–776) DOI:10.1167/iovs.05-0690

Since the beginning of the 20th century, the importance of retinal straylight for visual function has been well recognized by many investigators. After Cobb[1] introduced the concept of equivalent veiling luminance, Holladay[2] and Stiles[3] applied the concept in a disability glare formula, which has been well accepted and widely used. Retinal straylight was first studied in the normal aging population (reviewed by Vos[4]), and it was found to increase with age. Subsequently, it was also studied in eyes with various ocular diseases, such as corneal diseases,[5] cataract,[6,7] and corneal edema,[8] where the straylight was found to increase with an increase in opacity and irregularity of ocular media. The extent to which retinal straylight from headlights of oncoming cars impairs visual function in night traffic has been investigated by many researchers (see Ref. 9 for a review).

Retinal straylight can be seen as the outer skirt of the point–spread function,[10] outside, say, 1°. It causes a veiling luminance over the whole retina that adds to the retinal projection of the visual scene, thereby reducing the contrast of the retinal image. Disability glare, as defined by the Commission International d'Eclairage,[4] corresponds to retinal straylight, which is quantified by means of the concept of equivalent luminance (i.e., the [external] luminance that has the same visual effect as the glare source at some angular distance).[4]

The first attempts to measure intraocular straylight by means of equivalent luminance involved the comparison of two threshold measurements: one threshold in the presence of a distant glare source and one threshold in the presence of a homogeneous background (equivalent) luminance.[4] Van den Berg and IJspeert[11] compared the results from various groups, all using this method, and concluded that these results varied considerably. Moreover, the method was not widely used, because it was not easily accessible for clinical application. As easy-to-use alternatives, so-called glare testers were introduced that usually consisted of visual acuity (e.g., ETDRS,[12] Ferris-Bailey,[13] or Regan[14] charts) or contrast sensitivity (e.g., sinusoidal gratings,[7,12,14,15] Landolt rings,[12,16] or Pelli-Robson charts[13,14,17]) test, with and without a glare source presented at some angular distance in the visual field. Although glare testers were occasionally appraised favorably,[14] more often provided unreliable results, demonstrated by their outcomes correlating badly with various validity measures such as outdoor visual acuity in bright sunlight,[12,15] a questionnaire assessing perceived visual disability,[13,16] or directly measured forward light scatter.[14,16] Also, the repeatability and discriminative ability of studied glare tests were found to be inadequate.[14,16] A particular example is the omission of the glare measurement results, performed with the Miller-Nadler glare tester, in the final results of the large multicenter PERK study,[18] because the glare tester was not sensitive enough to detect small but significant amounts of light scattering,[19] which was also mentioned in later studies.[14,20] As a result of these issues with glare testers, a standard way of glare measurement was never adopted, and some overview papers discussing glare test problems appeared.[21–25]

To improve on this situation, Van den Berg[5] proposed a new psychophysical method, called the direct compensation method. In short, this method works as follows (Fig. 1): A bright, ring-shaped, flickering light source is presented at a certain angular distance (θ) from a (dark) test field. Because of intraocular scatter, part of the light from the bright straylight source is projected on the retina at the location of the test field, inducing a (weak) flicker in the test field. To determine the exact amount of straylight, variable *counterphase* compensation light is presented in the test field. By adjustment of the amount of compensation light, the flicker perception in the test field can be extinguished. In this way, there is "direct

---

From The Netherlands Ophthalmic Research Institute, Royal Netherlands Academy of Arts and Sciences, Amsterdam, The Netherlands.

Submitted for publication June 2, 2005; revised August 4, October 6, and November 1, 2005; accepted December 16, 2005.

Disclosure: **L. Franssen**, None; **J.E. Coppens**, None; **T.J.T.P. van den Berg**, (P)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Luuk Franssen, The Netherlands Ophthalmic Research Institute, Royal Netherlands Academy of Arts and Sciences, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands; l.franssen@ioi.knaw.nl.

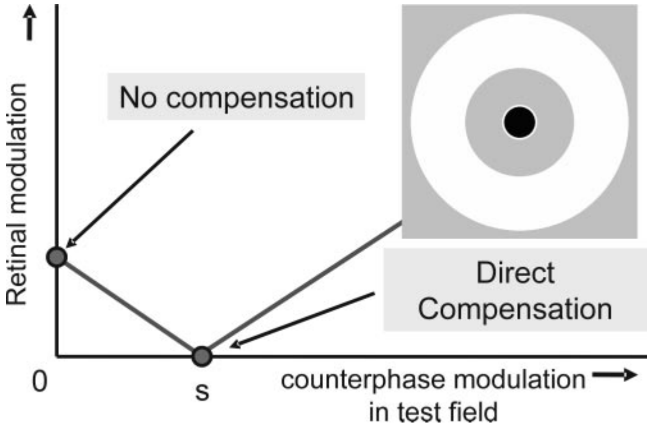

FIGURE 1. The direct compensation method. Retinal modulation in a foveal test field (*inset*: *black* field), resulting from scattered light from a constantly flickering annulus (*white*) is plotted against the amount of counterphase modulation in that test field. At point *s*, the flicker is extinguished, and the precise value of straylight found.

compensation" for the straylight modulation caused by light scattered from the glare source.

Using this technique, straylight was found to increase with age.[26] A new finding was that straylight depends on pigmentation of the eye. In further studies on pigmentation dependence, iris and ocular wall transmission was found to decrease with pigmentation in normal subjects[27] and was found to be significantly increased in patients with Fuchs' heterochromatic cyclitis.[28] Furthermore, intraocular light-scattering was found to be increased in retinitis pigmentosa[29] and cataract.[7]

In 1990, the direct compensation technique was implemented in a small portable device, called a straylight meter, to accommodate other researchers.[30–32] This method led to publications, notably by Elliott et al.,[6] on a variety of subjects, such as the already mentioned cataract and disability glare test evaluation[14] studies. Furthermore, they found increased straylight values after induced corneal edema,[8] in contact lens wearers[33] as well as in 25% of the subjects 1 year after excimer laser photorefractive keratectomy.[34] Advantages of the direct compensation method over alternative methods of assessing wide angle-scatter were mentioned in a paper discussing these methods for use in evaluating visual function in cataract.[24] Other researchers used the straylight meter after refractive surgery and found increases in small scatter angles and dilated pupils after radial keratotomy,[35] but no increases more than 2 weeks after photorefractive keratectomy,[20,36] except in some individuals.[37,38] Ocular lubricants were reported to have no adverse effects on the optical quality of the eye.[39] The straylight meter again showed increases in patients with retinitis pigmentosa[40,41] and also in those with choroideremias.[42] Straylight meter readings were found to correlate significantly with clinical grading of lens opacities and lens back scatter,[17] with corneal swelling,[43] with lens opacity measurements in patients with glaucoma,[44] and with posterior capsule opacification.[45] Most recently, the direct compensation method was used in a field study investigating the suitability of several glare tests for driver's license applications and was found to be the most promising candidate.[16,46]

In general, the direct compensation method has given a great boost to the study of retinal straylight. Moreover, it was emphasized in the literature that this technique has much greater sensitivity than do glare tests, for example in patients with corneal edema[8] and posterior capsular opacification.[45] It was also the gold standard for assessing the validity of glare tests.[14] However, outside the laboratory, it was a difficult technique to use.[38,45] In a field study[16,46] involving 112 subjects drawn from the patients and visitors of the outpatient departments of three clinics, the standard deviations of differences between repeated measurements found in such a field study were 0.15 and 0.18 log units, for two different implementations of the direct compensation method. It appears that the method has some major drawbacks for routine clinical or large-scale use: (1) Judgment of the weak flicker in the test field often appeared to be difficult for untrained subjects. This seemed to be caused by the presence of the strong flicker of the straylight source. (2) Usually, visual tests are based on what subjects actually see. On the contrary, in the direct compensation method, the subjects have to indicate whether the flicker perception has disappeared. The continuous flickering of the straylight source in the periphery made this contraintuitive task even more difficult. (3) The accuracy of the measurement seemed to depend on the adjustment strategy, which could differ considerably between subjects, and on proper explanation of the test. (4) There was no control over an individual's measurement reliability. (5) Subjects had the ability to influence the test outcome. This aspect is particularly important in the field of driver testing.

As a result of these drawbacks, the straylight meter largely remained limited to laboratory use. The instrument could not be used on a large scale, such as clinical diagnosis or occupational health testing. For these applications, the test must be easy to understand, easy and quick to perform, easy to explain, and fraud resistant. Also it should be criterion independent, so that the values have universal validity and results from different locations can be compared.

To overcome these limitations, we proposed a new method to measure retinal straylight, the "compensation comparison" method. In essence, this method presents exactly the same stimuli to the subject as the direct compensation method. Note that in the direct compensation method, the amount of compensation light is varied until the straylight flicker has disappeared. In other words, in the direct compensation method, the subject compares different stimuli *sequentially*. In contrast, in the compensation comparison method, two stimuli of the direct compensation method are presented to and compared by the subject *simultaneously*. In this way, the direct compensation method is implemented as a two-alternative forced-choice (2AFC) approach. The characteristics of the psychometric function for this 2AFC method will be reported in this article. This function determines what comparisons would be the best to use. The compensation comparison method has been summarized in abstract form (Van den Berg TJTP, et al. *IOVS* 2005;46:ARVO E-Abstract 4315) and in a patent.[47]

The compensation comparison method has been used in a field study involving 2422 subjects (GLARE, see www.glare.be) and in other projects such as a study investigating the wavelength dependence of retinal straylight.[48] Some results from the GLARE study, pertinent to the present question, will be used in this report. In this study, several visual tests, including straylight measurements, were performed among a population of drivers in Europe, spread over five age categories. Data were collected in clinics in The Netherlands, Austria, Germany, Spain, and Belgium. Since the study was designed to assess the prevalence of vision impairments in the driving population, the only inclusion criterion was being an active driver. As a result, the measured population consisted of a wide range of subjects, including ages from 20 to 85, visual acuities below 0.5 (logMAR [logarithm of the minimum angle of resolution] 0.3) to more than 1.0 (logMAR 0.0), visual field defects, and other ocular diseases such as glaucoma and cataract. This huge variation in ocular conditions provided an ideal opportunity to evaluate the compensation comparison method in clinical practice.

In the present paper, the principles, design considerations, and advantages of the compensation comparison method with respect to the direct compensation method are discussed, and a model for flicker comparison using this method is proposed and tested in a laboratory experiment. This model comprises a psychometric function designed to describe the (stochastic) characteristics of the responses in a compensation comparison experiment. For simplicity, real error responses (false-positive and false-negative mistakes of the subjects) were not included in the formulas that follow. These values are very low (on the order of 1% or less) in laboratory experiments. Their inclusion is straightforward though, and they were included in the final formulas used for the field study. The reliability of the compensation comparison method was tested with a commercially available embodiment of the method (C-Quant, manufactured by the Germany based firm Oculus Optikgeräte GmbH (Wetzlar-Dutenhofen, Germany).[47]

METHODS

The compensation comparison method was tested on seven subjects (age range, 21–57 years; mean 30). They were laboratory students and coworkers, including the authors. All subjects were without ocular defects. Testing was performed monocularly on the subject's preferred eye. Refraction ranged from −7 to emmetropic. Habitually worn glasses were allowed, but contact lenses were replaced by trial glasses. It must be noted that the test does not require refractive correction to be precise. Corrections were chosen for comfortable viewing, resulting in a +2 near addition for the older subjects, since the tests were performed at a distance of 32 cm from the stimulus screen. The study adhered to the guidelines of the Declaration of Helsinki for research in human subjects.

To test the compensation comparison method also for conditions of increased scattering, the same subjects were measured with a light–diffusing filter (Black Pro Mist 2 BPM2; Tiffen Manufacturing, Hauppauge, NY) in front of the tested eye. This filter, among a collection of 23 commercially available light–diffusing filters, was found to have the best light-scattering characteristics for mimicking (early) cataract or aging effects in the human crystalline lens.[49]

As mentioned before, the compensation comparison method was evaluated in the European GLARE study. In the course of this study, some improvements were made on the implementation of the test, as will be described at the end of this section.

For stimulus generation, a computer system with either a CRT monitor or combination of digital light processing (DLP) projector and back-projection screen was used. The straylight source was a white light annulus extending from 7° to 14°. Because of the approximate $1/\theta^2$ dependence of retinal straylight, this corresponds to a 10° scattering angle.[10]

To test the reliability of the clinical version of the compensation comparison method (C-Quant; Oculus Optikgeräte), 17 subjects with no experience in the direct compensation and compensation comparison measuring techniques were recruited from a neighboring institute. The average age was 44 years (range, 28–81). Except for the oldest subject, all were without ocular defects. Refraction ranged from −7 to +3 D. All measurements were performed monocularly on the subjects' preferred eyes, without glasses or contacts. Thoroughly cleaned trial glasses were used when appropriate. All subjects performed six measurements: three without and three with the BPM2 filter in front of the studied eye.

Basics of the Compensation Comparison Method

The test screen layout of the compensation comparison–based straylight meter is similar to that of the direct compensation method, only the test field is now divided in two halves (Fig. 2). Compensation light is presented in one of the two test field halves (randomly chosen, referred to as field b in the remainder of the article), whereas no compensation light is present in the other test field half (referred to as field a). As a result, two flickers are perceived, that differ in modulation depth: one results from straylight only (field a), the other is a combination of straylight and compensation light (field b), flickering in counterphase with this straylight. Simplified, the procedure runs as follows (Fig. 3): during the test, a series of limited–duration stimuli are presented that differ in the amount of compensation light in test field b. After a 2AFC paradigm, the task for the subject is to decide for each stimulus which test field half flickers stronger. The subject's responses are recorded by means of two push buttons, representing the left and right test fields. Using the psychophysical model for this flicker comparison task, which will be described in detail later in the article, a psychometric curve is fitted to the subject's responses, from which both the straylight parameter and a measure for the quality of the measurement can be deduced.

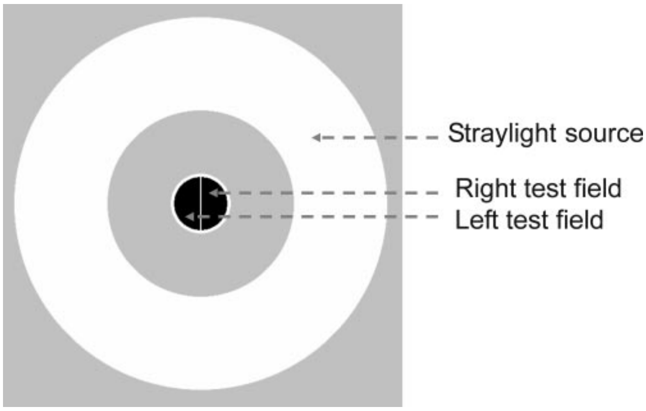

FIGURE 2. Stimulus layout for the compensation comparison method for retinal straylight measurement.

The fitting process makes use of a maximum-likelihood procedure that can briefly be described as follows. Assume that an experiment consists of *n* stimuli, to which *n* binary answers (e.g., yes or no, left or right) are obtained. Given a specific psychometric function, each answer of a subject has a certain likelihood, ranging between 0 and 1, since a psychometric function gives a certain probability (between 0 and 1) for a certain answer (yes or no) to the stimulus. For a complete experiment, the *n* answers correspond to *n* likelihoods. The total likelihood of the experiment is defined as the product of these *n* likelihoods, and the value of this total likelihood depends on the assumed psychometric function. The best-fitting psychometric function is the one that gives the highest total–likelihood value (Fig. 4 shows an example of an actual compensation comparison experiment). For a more complete description of the maximum-likelihood concept,[50] a separate paper discussing reliability assessment in the compensation comparison method, using the likelihood function, is in preparation (Coppens JE, et al., manuscript submitted).

It must be noted that the compensation light added in field b results in a change in average luminance. This change may confuse the subject, cause bias, or form a clue to manipulate the test outcome. To ensure that the two test fields are only different in retinal modulation, an offset of half the compensation value is added to field a (Fig. 5). This equates the average luminance in both test fields, while maintaining the (absolute) modulation.

Trial Strategy

The measurement procedure (Fig. 4 shows an example of an actual measurement from the GLARE study) consists of two consecutive stages with different types of stimuli: the "dark" or "initial" phase and the "light" or "final" phase. The initial phase (Fig. 4, dots) serves to obtain a first coarse estimate of the straylight value and to make the task easy at first. The final phase (Fig. 4, ×'s) serves to refine the first coarse estimate. In the initial phase, the amount of straylight is varied by varying the intensity of the flickering ring, while the compensation light is kept constant. In this phase, the task is very easy at first, becoming gradually more difficult, until the straylight value of the respective individual is approached. In other words, apart from giving a first coarse estimate, the initial phase serves as a training phase for the flicker comparison task, in which the potentially disturbing peripheral

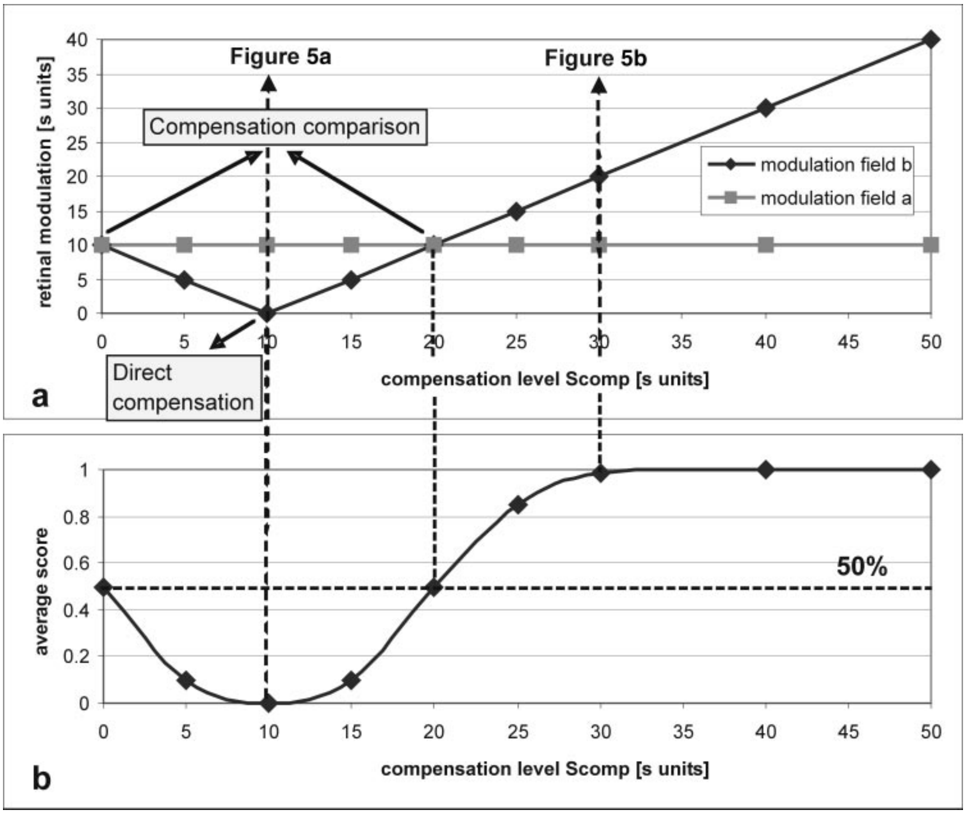

FIGURE 3. (**a**) A simplified straylight test with variable compensation in one test field and no compensation in the other field. Note that the V-shaped function (modulation field b) also corresponds exactly to the function of the direct compensation method shown in schematic form in Figure 1. (**b**) Probability of getting a 1 score (compensation test field flickers the most) as a function of the strength of the counterphase compensation light (psychometric function). The straylight value for this subject is 10.

flicker is very weak at first. In the final phase, maximum light intensity of the straylight source is used, to have maximum light intensity in the comparison task. At higher light intensities, the comparison task is performed more accurately (see also the Results section).

The stimuli in the initial phase are equidistant, with a step size of 0.1 log units (except for the first step which is 0.3 log units) and presented in order from high to low straylight (Fig. 4, increasing numbers). The absolute stimulus values of the initial phase can be placed differently and can be chosen by the operator, but in the GLARE study it was set to adjust for the known population averages as a function of age.[10] The example given in Figure 4 is for a 30-year-old subject. For a 70-year-old subject, all initial phase stimuli were shifted upward by 0.3 log units.

In the first stimulus, a very weakly flickering ring is presented (stimulus 1 in Fig. 4). Then it is very easy to recognize the test field half with compensation. Subsequently, the intensity of the ring is increased, thereby increasing the difficulty of the flicker comparison task. This relates to the real-life experience of being disturbed more and more by glare sources with higher intensities. In the final phase, the ring flickers at constant intensity, whereas the compensation luminance in field b is varied. The stimuli in the final phase (Fig. 4, ×'s) are logarithmically equidistant at 0.05 log units in a fixed interval around the first coarse estimate of the 50% point of the psychometric curve, as based on the data of the initial phase (Fig. 4, dots). In the final phase, the stimuli are presented in random order, according to the method of constant stimuli.[51]

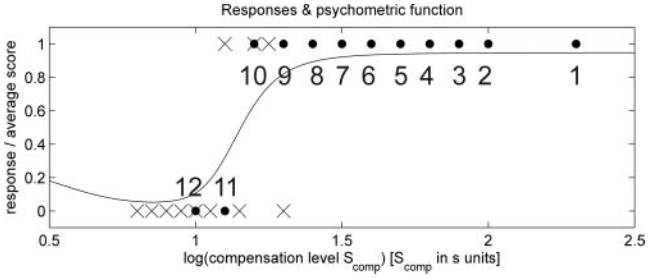

FIGURE 4. Example of an actual measurement in a 30-year-old subject from the GLARE study (emmetropic, no ocular disease, clear eye media, best corrected visual acuity 1.25 (logMAR −0.1)). The *dots* represent the initial phase of the measurement (stimuli presented to the subject in fixed order, 1–12). (×) The final-phase responses (stimuli presented in random order, centered around the initial phase estimate of the 50% value). A psychometric curve was fitted to the data by means of a maximum-likelihood procedure, with shape parameter MDC_c fixed at a certain pre-chosen level. The fit results in a straylight value $s = 7.14$, or $\log(s) = 0.85$.

The Psychometric Function

As a basis to describe the psychometric function, we started out from the well-known logistic function.[52] Comparing two flickering test fields a and b with different modulation depths, the chance probability (P) of choosing one of the test fields as having the stronger flicker was written as (Fig. 6)

$$P = \frac{1}{1 + e^{-[MDC/MDC_c]}}, \qquad (1)$$

where MDC_c is the parameter in the equation, giving a critical value for modulation depth contrast. MDC is the independent variable in the equation, giving the contrast between the two flickers, defined as

$$MDC = \frac{MD_b - MD_a}{MD_b + MD_a}, \qquad (2)$$

where MD_a and MD_b represent the retinal modulation depths in both test fields.

It must be noted that the light the fovea (the two half fields) receives, consists of two parts: light originating from the flickering annulus by the process of scattering and light originating from the half fields the subject is looking at. Both lights correspond to certain luminances in the outside world (in the two half fields). The light originating from scatter (i.e., the straylight) corresponds to an outside luminance (called equivalent luminance[4]) according to the equation[10]

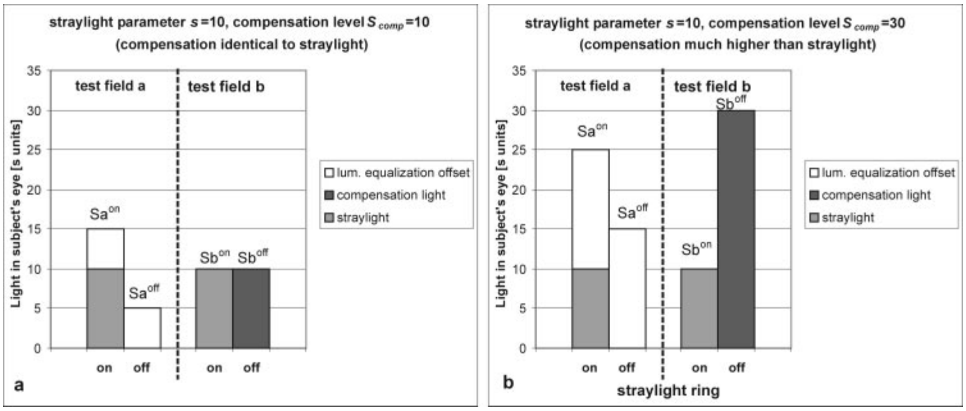

FIGURE 5. The average luminance is equalized between the two half fields. The result is shown for two different examples of a stimulus in a compensation comparison experiment (both indicated in Fig. 3): (**a**) Precise compensation for the straylight flicker; (**b**) Overcompensation for the straylight flicker by a factor of 3. *Light gray*: pure straylight flicker, which is the same in both half fields, but also in both stimuli of one experiment, since straylight, expressed in *s* units, did not change in one experiment (it depended only on the light-scattering characteristics of the eye under examination). In both cases, average luminance is equalized by adding half the compensation value as an offset to the other half field (*white* bars = 0.5 · *dark gray* bars).

$$L_{\mathrm{eq}} = 0.0013 \cdot s \cdot L_{\mathrm{src}}, \tag{3}$$

where L_{src} is the luminance of the straylight ring, and *s* is the "straylight parameter," a value that characterizes the amount of light-scattering in the eye under investigation. A more extensive explanation has been published.[10] Conversely, because L_{src} is known, we can use equation 3 to express the external luminance in the test fields (as seen by the fovea) in "equivalent" straylight parameter units. In other words, each given external luminance *L* corresponds to an equivalent *s* value. The modulation depths can then be written as

$$MD_{\mathrm{a}} = \left|\frac{La^{\mathrm{off}} - La^{\mathrm{on}}}{La^{\mathrm{off}} + La^{\mathrm{on}}}\right| \text{ and } MD_{\mathrm{b}} = \left|\frac{Lb^{\mathrm{off}} - Lb^{\mathrm{on}}}{Lb^{\mathrm{off}} + Lb^{\mathrm{on}}}\right|, \tag{4}$$

where *L* is the true or equivalent luminance, or eventually a combination of both. La^{off} and Lb^{off} represent the light in the off-phase of the straylight ring, whereas La^{on} and Lb^{on} represent the light in the on-phase of the straylight ring. By combining equations 3 and 4, we can express the retinal light levels in (equivalent) straylight parameter units, referred to as "*s* units" in this article

$$MD_{\mathrm{a}} = \left|\frac{Sa^{\mathrm{off}} - Sa^{\mathrm{on}}}{Sa^{\mathrm{off}} + Sa^{\mathrm{on}}}\right| \text{ and } MD_{\mathrm{b}} = \left|\frac{Sb^{\mathrm{off}} - Sb^{\mathrm{on}}}{Sb^{\mathrm{off}} + Sb^{\mathrm{on}}}\right|, \tag{5}$$

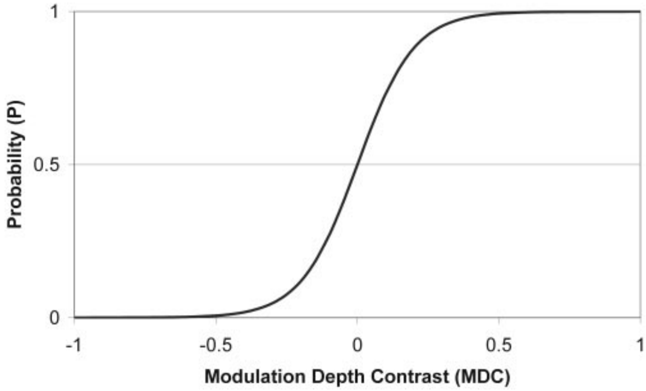

FIGURE 6. Function used as a psychometric function for flicker comparison (equation 1). MDC_{c} was set at 0.13 for this example.

since for any given situation the factor 0.0013-L_{src} drops out of equation 4. The on-phase light is the straylight *s* originating from the flickering ring, summed in field a with the luminance equalizing light which equals half of the compensation light in field b (Fig. 5). The off-phase light is the compensation light S_{comp} in field b. Half of this amount is again added as an offset to field a, serving as luminance-equalizing light. In formulas

$$Sb^{\mathrm{on}} = s \qquad Sb^{\mathrm{off}} = S_{\mathrm{comp}} \tag{6}$$

$$Sa^{\mathrm{on}} = s + 0.5 \cdot S_{\mathrm{comp}} \qquad Sa^{\mathrm{off}} = 0.5 \cdot S_{\mathrm{comp}}\,. \tag{7}$$

Plotting the probability against S_{comp} or log (S_{comp}) results in psychometric curves as in Figure 3b and Figure 4, respectively. The model parameters (*s* and MDC_{c}) were fitted by means of a maximum-likelihood procedure (described in short earlier) to the seven subjects' laboratory data.

Once the shape of the psychometric function has been established, estimation of the straylight parameter value *s* in individual subjects involves shifting of the psychometric function to fit the dataset of that individual. Fitting is achieved by means of the maximum-likelihood procedure, as just outlined. An example of such a fit is given in Figure 4. In this case, log(*s*) was found to be 0.85. The straylight value is determined by the horizontal position of the minimum of the curve, where $MD_{\mathrm{b}} = 0$ and $S_{\mathrm{comp}} = s$. This approach was applied in the European GLARE study involving 2422 subjects in total. In the course of the study, some improvements were made on the implementation of the straylight test: (1) A three-trial instruction phase was added before the real measurement, to familiarize the subject with the flicker–comparison task. (2) The subject's responses were displayed to the operator during the measurement, making it possible to interfere in case the response pattern was erratic and start a new measurement after additional explanation. (3) The luminance in the test fields was increased by a factor of 2 in the initial phase, making the measurement easier for older subjects. In total, 1073 subjects were measured with this final version (including these improvements). More detailed reports of this study are in preparation, but some preliminary data will be given herein to test the psychometric function (equation 1) and to illustrate the performance of the test.

RESULTS

Figure 7 shows the results of experiments performed to evaluate the model described in the previous section. All measure-

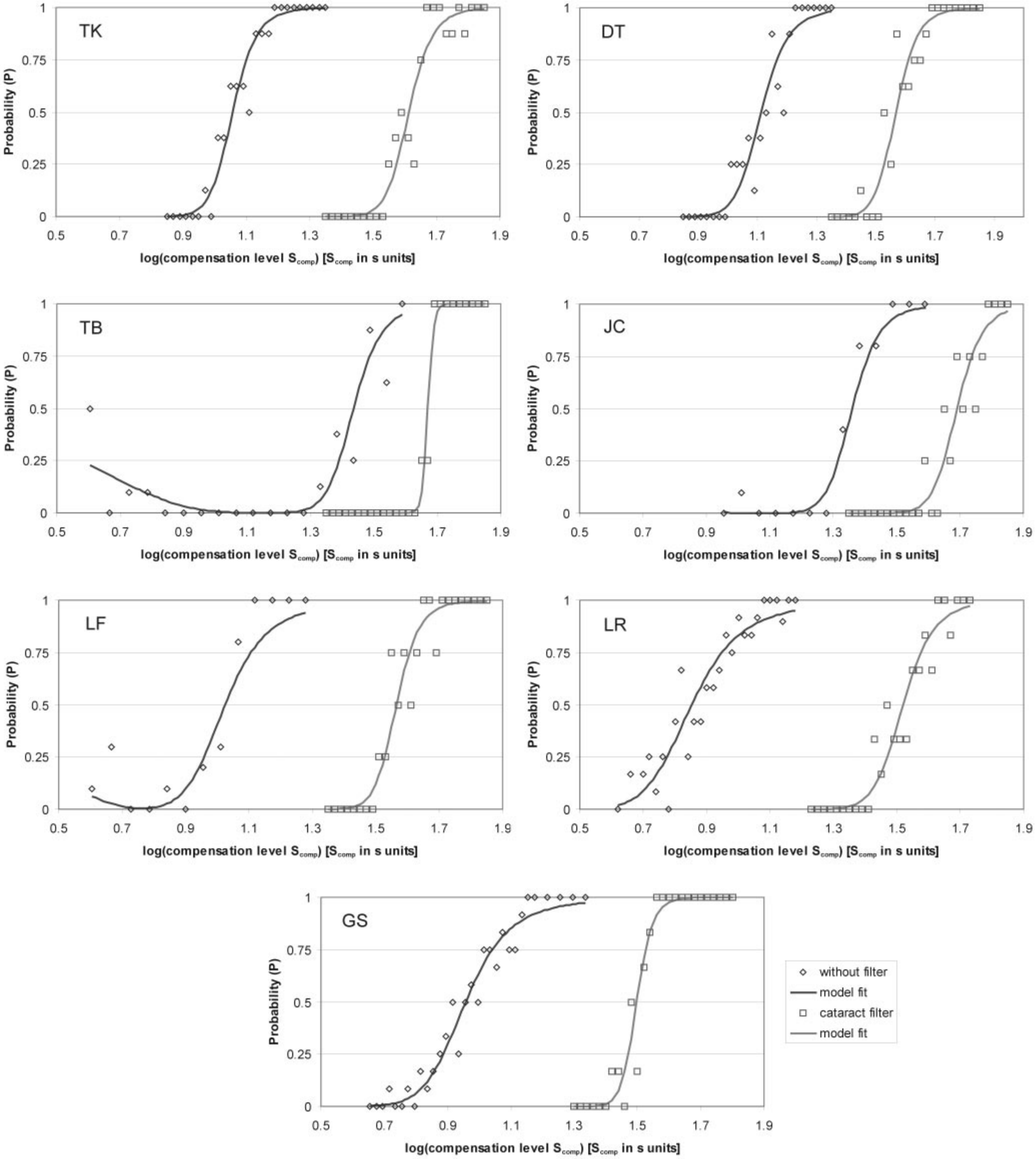

FIGURE 7. Measured psychometric curves and corresponding model curves (equation 1) in seven subjects. All measurements were performed monocularly. Each subject was measured without and with a BPM2 filter in front of the measured eye. Data points are averages over 8 (TK, DT), 10 (TB, JC, LF), or 12 (LR, GS) responses for the no-filter measurements, and averages over 4 (TB, JC, LF), 6 (LR, GS), or 8 (TK, DT) responses for the BPM2 measurements. In each subject, equation 1 was fitted to all data points with straylight parameter *s* and psychometric function shape parameter MDC_c as parameters (Table 1).

ments were repeated with the cataract simulating BPM2 filter in front of the eye. Values for the independently fitted parameters *s* and MDC_c are given in Table 1.

Figure 7 shows that the mathematical expression for the psychometric function, proposed in the Methods section (equation 1), performs very well in describing all measurements. Apart from the straylight value log(*s*), which determines the horizontal position of the curve, the differences in shape between the curves in Figure 7 all derive from differences in one parameter only, MDC_c (see also Table 1). Although the differences are not large, there seems to be a systematic effect of a steeper slope (MDC_c somewhat lower) with more straylight (cataract model curves). This may be understood by noting that with more straylight (curve shifted to the right), the

TABLE 1. Maximum-Likelihood Fits of Equation 1 to the Compensation Comparison Measurements

Subject Details			Fit Results			
			Without Filter		With BPM2 Filter	
Initials	Age (y)	Average Age Normal Log (Straylight Parameter) Log(s)	Log (Straylight Parameter) Log(s)	Shape Parameter Psychometric Function Log(MDC_c)	Log (Straylight Parameter) Log(s)	Shape Parameter Psychometric Function Log(MDC_c)
TK	24	0.85	0.76	−1.07	1.31	−1.07
DT	23	0.85	0.81	−0.98	1.27	−1.08
TB	57	1.00	1.15	−0.98	1.37	−1.68
JC	35	0.87	1.08	−0.99	1.39	−1.04
LF	29	0.86	0.76	−0.77	1.26	−1.09
LR	21	0.85	0.55	−0.75	1.22	−0.96
GS	22	0.85	0.66	−0.79	1.20	−1.25
Average	30.14			−0.90		−1.17
SD				0.13		0.24

Data were obtained in seven subjects, without and with a BPM2 filter (artificial straylight increase) in front of the measured eye. Fitted parameters are the straylight parameter s and the critical modulation depth contrast MDC_c, the latter being the shape parameter for the psychometric function.

flicker intensity in the central test fields is higher, which might make the flicker comparison task easier. This may turn out to be an advantage in practice. Speculatively, in eyes that are in worse ophthalmic condition, possible detrimental effects on psychometric behavior may be counteracted by this phenomenon.

The model was further validated by applying it to field measurements of 1073 subjects, performed in the European GLARE study, as described in the previous section. The wide variation in ocular conditions found in this population can be expected to reflect itself in different psychophysical behavior, and therefore in psychometric functions that differ between these 1073 individuals. To analyze this, all measurements were performed twice and divided in nine groups of equal size, sorted on the differences between the two repeated measurements. In each group, equation 1 was then fitted to all data, after normalizing each individual curve for the individual straylight value. Results are given in Figure 8. The best 67% (top six panels) of the 1073 subjects have a repeated-measurement SD of 0.036, and 89% (all but the last panel) of 0.059 log units, whereas the SD for all measurements is 0.099 log units.

The C-Quant measurements are summarized in Figure 9. The three separate measurements without a filter are plotted against the average of these three measurements. This figure shows that the repeatability of the compensation comparison method is very high. In addition to these repeatability data, the average of the three measurements with the BPM2 filter is also plotted against the average of the three measurements without filter. The dashed line represents the expected straylight value for the eye-plus-filter combination, obtained by adding the straylight values of the filter and the eye on a linear scale.

DISCUSSION

In this article, we have presented an approach toward retinal straylight measurement, intended to be feasible for routine

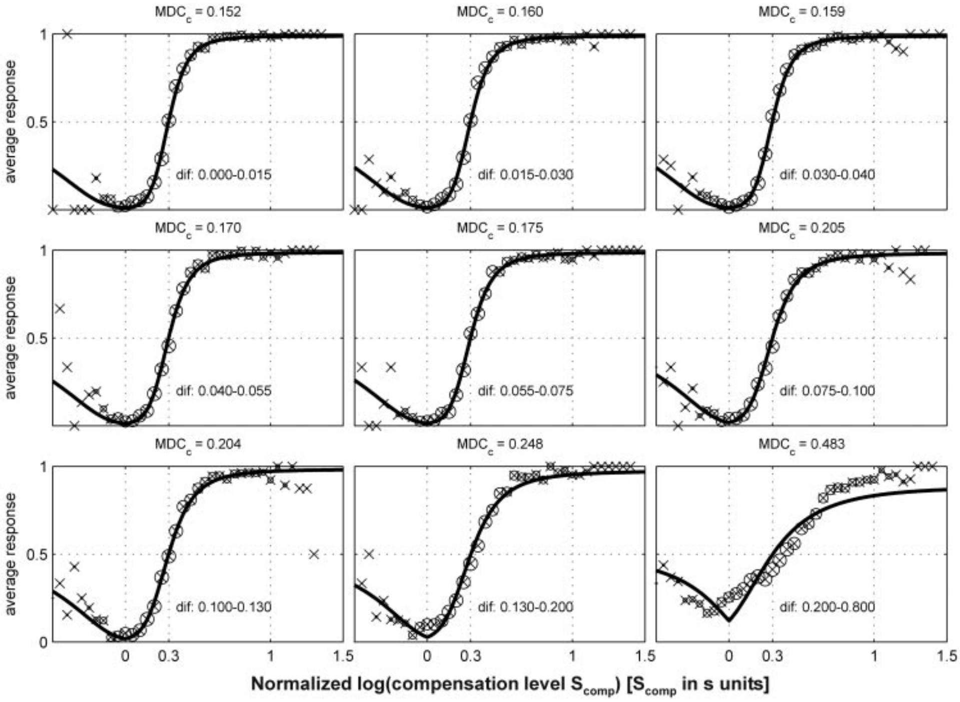

FIGURE 8. Equation 1 fitted to field measurements of 1073 subjects. Measurements were divided into nine groups of equal size, sorted on log(s) differences of repeated measurements (denoted as dif in the graphs). After normalization for log(s), the psychometric function model (equation 1) was fitted to the data, resulting in an MDC_c value in each group. The corresponding psychometric curves are drawn in each graph.

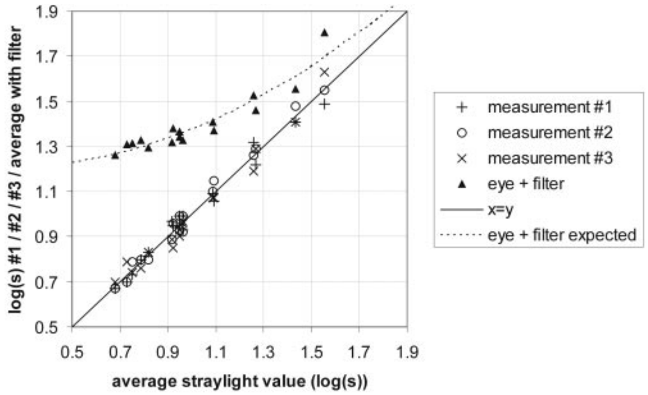

FIGURE 9. Reliability of the compensation comparison method, using the C-Quant (Oculus Optikgeräte GmbH, Wetzlar-Dutenhofen). Seventeen untrained subjects were measured three times without and three times with a BPM2 filter in front of the studied eye. The three separate measurements without filter are plotted against the average of these three measurements. In addition, the average of the three measurements with filter is plotted against the average without filter. *Dashed line*: expected straylight value for the eye-plus-filter combination.

clinical use. Reliability analysis of the population data (Fig. 8) can be refined by using the available response pattern (Fig. 4) of each individual to test for reliability. In fact, a reliability parameter has already been developed (described later). With this parameter, an overall repeated measurement SD between 0.06 and 0.1 log units can be obtained, depending on the filter criterion (percentage that is filtered out). Even without filtering, this SD is better than the repeatability of the direct compensation method, which was found to be 0.15 and 0.18 log units in field studies, for two different implementations of this method.[16] In practice we use a value of 0.07. The reliability parameter will be discussed in more detail in a separate paper (Coppens JE, et al., manuscript submitted). It should be noted that this parameter can also be used to evaluate the quality of a single measurement, which is important for clinical use. This ability to check and reinstruct or to exclude measurements based on an individual-specific measurement quality criterion is a main advantage of the compensation comparison technique with respect to the direct compensation technique.

This study was provoked, among others, by existing evidence in the literature that there is a clinical need for testing a patient's glare sensitivity. As outlined in the introduction, many different glare testers have been proposed, most of which have disappeared from the market. Some studies tried to validate glare testing against straylight as the gold standard, but with questionable results.[7,14,16,17] Repeatability was compared between different glare tests and the direct compensation method,[14] leading to the conclusion that the direct compensation method performs better. With the improved performance of the compensation comparison method, this will, a fortiori, be the case again. The present compensation comparison technique offers new opportunities to test and validate the performance of glare testers.

To obtain these results, some understanding of the underlying psychometric function was needed. The proposed model describes measured laboratory data well (Fig. 7) for a wide range of straylight values (Table 1). The log(s) values without BPM2 filter all fall within the normal population range, which has been shown to increase with age.[10] From this study, it follows that the relation between straylight parameter s and age can be approximated, in a white population with a 10° scattering angle, by the equation $s = 7(1+(\text{age}/70)^4)$, with an uncertainty of 0.1 log units. In Table 1, this average age-normal population value is given for each subject.

The log(s) values with BPM2 filter show less variation. This is because the total straylight is a combination of the filter (which itself has log(s) = 1.12) and the eye (log(s) values from 0.55 to 1.15). The experimental values for the eye–filter combinations (log(s) values from 1.20 to 1.39) correspond well to values that can be predicted by calculation (log(s) values from 1.22 to 1.44).

Figure 8 shows that the model is capable of accurately describing the psychophysical behavior of a population that varies widely with respect to physical condition of the eye. Subdividing the population according to differences between two repeated measurements reveals different slopes of the psychometric curves of the various subgroups, accounted for in the model by different MDC_c values. The model fits fairly well to all subgroups of Figure 8, except for the subgroup with the largest repeated measurement differences (lower right panel). For some cases in this subgroup, response behavior was so erratic that reliably fitting a psychometric curve and therefore reliably estimating the log(s) value, is not possible. To detect such erratic behavior automatically during measurements, we developed a reliability parameter, as mentioned earlier. This parameter must assume a certain shape of the psychometric function and was based on the analysis in the present paper. After the lowest-quality measurements were filtered out with this parameter, the overall SD of repeated measurements was between 0.06 and 0.1 log units, which is very good, considering the variation in straylight parameter in the (clinical) population.

Figure 9 shows that the compensation comparison method gives highly repeatable results in untrained subjects, over a wide range of straylight values. The measurements with BPM2 filter follow the additive model for eye plus filter very accurately, indicating that the instrument measures absolute straylight values very well. The fitted log(s) value for the filter (1.14) is very well in accordance with the objectively measured log(s) value for this filter of 1.12.[49]

The compensation comparison method for measuring retinal straylight was designed as an improvement on the direct compensation technique. According to feedback we got from the operators in the clinics who participated in the GLARE study and who also had earlier experience with the direct compensation method, the task is easier and more intuitive (mostly suprathreshold, short stimulus presentations), easier to explain and less dependent on explanation from the operator. The measurement time is fixed and limited, making the test more pleasant for both patient and operator. However, we did not collect systematic statistical data on these subjective assessments. Moreover, the reliability of the compensation comparison method was shown to be very good, and a reliability index was developed, based on the dataset of a tested individual.

Given these advantages, retinal straylight measurement is now possible on a large scale and in the clinical routine. As a result, the compensation comparison method described in this article has been implemented in a commercially available measurement device (the C-Quant; Oculus Optikgeräte). For future development, the model for flicker comparison gives a basis for improving on the measurement performance by studying different measurement strategies, such as adaptive methods.

References

1. Cobb PW. The influence of illumination of the eye on visual acuity. *Am J Physiol.* 1911;29:76–99.
2. Holladay LL. The fundamentals of glare and visibility. *J Opt Soc Am.* 1926;12:271–319.
3. Stiles WS. The effect of glare on the brightness difference threshold. *Proc Roy Soc.* 1929;104:322–355.
4. Vos JJ. Disability glare: a state of the art report. *Comm Int de l'Eclairage J.* 1984;3/2:39–53.
5. van den Berg TJTP. Importance of pathological intraocular light scatter for visual disability. *Doc Ophthalmol.* 1986;61:327–333.

6. Elliott DB, Hurst MA, Weatherill J. Comparing clinical tests of visual loss in cataract patients using a quantification of forward light scatter. *Eye*. 1991;5:601–606.
7. de Waard PW, IJspeert JK, van den Berg TJTP, de Jong PT. Intraocular light scattering in age-related cataracts. *Invest Ophthalmol Vis Sci*. 1992;33:618–625.
8. Elliott DB, Fonn D, Flanagan J, Doughty M. Relative sensitivity of clinical tests to hydrophilic lens-induced corneal thickness changes. *Optom Vis Sci*. 1993;70:1044–1048.
9. Mainster MA, Timberlake GT. Why HID headlights bother older drivers. *Br J Ophthalmol*. 2003;87:113–117.
10. van den Berg TJTP. Analysis of intraocular straylight, especially in relation to age. *Optom Vis Sci*. 1995;72:52–59.
11. van den Berg TJTP, IJspeert JK. Intraocular straylight, studied using the direct compensation technique. 22nd session (division 1). Vienna: CIE Proceedings. 1991;83–84.
12. Prager TC, Urso RG, Holladay JT, Stewart RH. Glare testing in cataract patients: instrument evaluation and identification of sources of methodological error. *J Cataract Refract Surg*. 1989; 15:149–157.
13. Elliott DB, Hurst MA, Weatherill J. Comparing clinical tests of visual function in cataract with the patient's perceived visual disability. *Eye*. 1990;4:712–717.
14. Elliott DB, Bullimore MA. Assessing the reliability, discriminative ability, and validity of disability glare tests. *Invest Ophthalmol Vis Sci*. 1993;34:108–119.
15. Neumann AC, McCarty GR, Locke J, Cobb B. Glare disability devices for cataractous eyes: a consumer's guide. *J Cataract Refract Surg*. 1988;14:212–216.
16. van Rijn LJ, Nischler C, Gamer D, et al. Measurement of stray light and glare: comparison of Nyktotest, Mesotest, stray light meter, and computer implemented stray light meter. *Br J Ophthalmol*. 2005;89:345–351.
17. Yager D, Chih Ling Liu, Kapoor N, Yuan R. Relations between three measures of ocular forward light scatter and two measures of backward light scatter. *OSA Technical Digest Series*. 1993;3:174–177.
18. Waring GO III, Lynn MJ, Fielding B, et al. Results of the prospective Evaluation of Radial Keratotomy (PERK) Study 4 years after surgery for myopia. Perk Study Group. *JAMA*. 1990;263:1083–1091.
19. Waring GO III, Lynn MJ, Gelender H, et al. Results of the prospective evaluation of radial keratotomy (PERK) study one year after surgery. *Ophthalmology*. 1985;92:177–198, 307.
20. Veraart HG, van den Berg TJ, Hennekes R, Adank AM. Stray light in photorefractive keratectomy for myopia. *Bull Soc Belge Ophtalmol*. 1993;249:57–62.
21. Prager TC, Urso RG, Lewis JW, Ruiz RS. Methodological considerations in glare testing in patients with cataracts. *Arch Ophthalmol*. 1988;106:1501–1502.
22. American Academy of Ophthalmology. Contrast sensitivity and glare testing in the evaluation of anterior segment disease. *Ophthalmology*. 1990;97:1233–1237.
23. van den Berg TJTP. On the relation between glare and straylight. *Doc Ophthalmol*. 1991;78:177–181.
24. Elliott DB. Evaluating visual function in cataract. *Optom Vis Sci*. 1993;70:896–902.
25. van den Berg TJTP. On the relation between intraocular straylight and visual function parameters. *Invest Ophthalmol Vis Sci*. 1994; 35:2659–2661.
26. IJspeert JK, de Waard PW, van den Berg TJTP, de Jong PT. The intraocular straylight function in 129 healthy volunteers; dependence on angle, age and pigmentation. *Vision Res*. 1990;30:699–707.
27. van den Berg TJTP, IJspeert JK, de Waard PW. Dependence of intraocular straylight on pigmentation and light transmission through the ocular wall. *Vision Res*. 1991;31:1361–1367.
28. La Hey E, IJspeert JK, van den Berg TJ, Kijlstra A. Quantitative analysis of iris translucency in Fuchs' heterochromic cyclitis. *Invest Ophthalmol Vis Sci*. 1993;34:2931–2942.
29. van den Berg TJTP. Red glasses and visual function in retinitis pigmentosa. *Doc Ophthalmol*. 1989;73:255–274.
30. van den Berg TJTP, IJspeert JK. Clinical assessment of intraocular straylight. *Appl Opt*. 1992;31:3694–3696.
31. van den Berg TJTP, IJspeert JK. Straylight meter. *OSA Technical Digest.* 1991;1:256–259.
32. IJspeert JK, van den Berg TJTP. Design of a portable straylight meter. Proceedings 14th IEEE-EMBS. Houston: IEEE; 1992:1592–1594.
33. Elliott DB, Mitchell S, Whitaker D. Factors affecting light scatter in contact lens wearers. *Optom Vis Sci*. 1991;68:629–633.
34. Butuner Z, Elliott DB, Gimbel HV, Slimmon S. Visual function one year after excimer laser photorefractive keratectomy. *J Refract Corneal Surg*. 1994;10:625–630.
35. Veraart HG, van den Berg TJ, IJspeert JK, Cardozo OL. Stray light in radial keratotomy and the influence of pupil size and straylight angle. *Am J Ophthalmol*. 1992;114:424–428.
36. Harrison JM, Tennant TB, Gwin MC, et al. Forward light scatter at one month after photorefractive keratectomy. *J Refract Surg*. 1995;11:83–88.
37. Veraart HG, van den Berg TJ, Hennekes R, Adank AM. Stray light in photorefractive keratectomy for myopia. *Doc Ophthalmol*. 1995; 90:35–42.
38. Schallhorn SC, Blanton CL, Kaupp SE, et al. Preliminary results of photorefractive keratectomy in active-duty United States Navy personnel. *Ophthalmology*. 1996;103:5–22.
39. Veraart HG, van den Berg TJ. Ocular lubricants and intraocular stray light. *Doc Ophthalmol*. 1992;82:29–31.
40. Alexander KR, Fishman GA, Derlacki DJ. Intraocular light scatter in patients with retinitis pigmentosa. *Vision Res*. 1996;36:3703–3709.
41. Grover S, Alexander KR, Fishman GA, Ryan J. Comparison of intraocular light scatter in carriers of choroideremia and X-linked retinitis pigmentosa. *Ophthalmology*. 2002;109:159–163.
42. Grover S, Alexander KR, Choi DM, Fishman GA. Intraocular light scatter in patients with choroideremia. *Ophthalmology*. 1998;105: 1641–1645.
43. Fonn D, Du TR, Simpson TL, Vega JA, Situ P, Chalmers RL. Sympathetic swelling response of the control eye to soft lenses in the other eye. *Invest Ophthalmol Vis Sci*. 1999;40:3116–3121.
44. Horn FK, Junemann AG, Korth M. Two methods of lens opacity measurements in glaucomas. *Doc Ophthalmol.* 2001;103:105–117.
45. Meacock WR, Spalton DJ, Boyce J, Marshall J. The effect of posterior capsule opacification on visual function. *Invest Ophthalmol Vis Sci*. 2003;44:4665–4669.
46. van den Berg TJTP, van Rijn LJ. Entoptic straylight measurement using the direct compensation method in relation to driver licensing application. In: Gale A, ed. *Vision in Vehicles*. *X*. Derby: Vision in Vehicles Press. In press.
47. van den Berg TJTP, Coppens JE. Method and device for measuring retinal straylight. Patent number WO2005023103, NL1024232C. 2005.
48. Coppens JE, Franssen L, van den Berg TJTP. Wavelength dependence of retinal straylight. *Exp Eye Res.* In press.
49. de Wit GC, Franssen L, Coppens JE, van den Berg TJTP. Straylight contribution of cataracts and its simulation. *J Cataract Refract Surg.* In press.
50. Harvey LO Jr. Efficient estimation of sensory thresholds. *Behav Res Methods Instr Comp*. 1986;18:623–632.
51. Treutwein B. Adaptive psychophysical procedures. *Vision Res*. 1995;35:2503–2522.
52. Strasburger H. Converting between measures of slope of the psychometric function. *Percept Psychophys*. 2001;63:1348–1355.