1040-5488/06/8305-0290/0 VOL. 83, NO. 5, PP. 290–298
OPTOMETRY AND VISION SCIENCE

ORIGINAL ARTICLE

Measuring Contrast Sensitivity Under Different Lighting Conditions: Comparison of Three Tests

JENS BÜHREN, MD, EVDOXIA TERZI, MD, MICHAEL BACH, PhD, WOLFGANG WESEMANN, PhD, and THOMAS KOHNEN, MD

Department of Ophthalmology, Johann Wolfgang Goethe-University, Frankfurt am Main, Germany (JB, ET, TK), Department of Ophthalmology, University of Rochester Medical Center, Rochester, New York (JB), Department of Ophthalmology, Laboratory for Electrophysiology, University of Freiburg, Freiburg im Breisgau, Germany (MB), and College of Optometry (HFAK), Cologne, Germany (WW)

ABSTRACT

Purpose. The purpose of this study was to evaluate three psychophysical tests for the measurement of contrast sensitivity (CS) and disability glare (DG) at different luminance levels.

Methods. In 60 eyes of 60 individuals (group 1: 20 healthy eyes of young individuals; group 2: 20 healthy eyes of elderly subjects; group 3: 20 eyes with nuclear cataract), CS with best correction was measured twice with the Frankfurt-Freiburg Contrast and Acuity Test System (FF-CATS) and the Functional Acuity Contrast Test (FACT, 1.5 cycles per degree [cpd]) at 167 cd/m^2 and 0.167 cd/m^2, and with the Pelli-Robson Chart (PRC) at 100 cd/m^2 with and without glare. Repeatability of test and retest, and discriminative ability between the different subgroups, were assessed for CS values.

Results. Maximum CS values varied across tests. In all groups, highest CS values were obtained with the photopic FF-CATS. For FACT scores at 1.5 cpd, there was a ceiling effect for young subjects. CS scores obtained with the PRC were the lowest. The PRC had the best test–retest repeatability of all tests. Under mesopic conditions with glare, reliability was generally lower; the FF-CATS had the highest repeatability of the mesopic tests. The FF-CATS discriminated best between the different groups for all conditions.

Conclusions. There are large discrepancies in the test results between CS testing methods, especially under different lighting conditions. Results from different CS tests are not interchangeable.

(Optom Vis Sci 2006;83:290–298)

Key Words: contrast sensitivity; disability glare; psychophysical test; optical quality; mesopic vision

Cataract and refractive surgery have rapidly developed in the last decades. Clinical experience and experimental work has led to the insight that Snellen acuity is inadequate as a sole parameter for describing the quality of vision outcome of refractive surgical procedures.[1–7] Contrast sensitivity (CS) or low-contrast visual acuity (VA) testing will also play a role in determining the quality of vision.[8–13] A comprehensive evaluation of quality of vision requires testing under a range of lighting conditions; mesopic conditions may be more sensitive to optical changes than photopic conditions.[7,12] Patients frequently report symptoms only for mesopic or scotopic conditions, whereas quality of vision under photopic conditions may be unaffected.[5,11] Many tests for CS and disability glare (DG) lack uniformity in test principles or standardization of lighting conditions.[5,14–18] Most tests have problems that limit their use in clinical studies such as low reliability and a high probability of correct guessing.[15,18] The variety of testing methods used in CS studies, in addition to the even greater variety of ambient and glare luminance conditions, makes it difficult to compare results between studies.

For future clinical studies on quality of vision, a single test for measuring CS and DG, with standardized ambient and glare luminance levels that are closer to real-world conditions, is needed. Testing should be performed under a range of lighting conditions to assess the influence of ambient luminance on quality of vision. The lack of a reliable, commercially available test that allows determination of contrast and VA thresholds under different lighting and glare conditions prompted us to construct a new test system. We chose the Freiburg Acuity and Contrast Test (FrACT) to display the stimuli.[19] It allows VA testing at defined contrast levels and CS testing at defined optotype sizes automatically and independent of the observer. Although this program has been in use since 1993[19–22] and one recently published study reported CS and glare data,[23] no critical evaluation has been performed. We constructed a laboratory setup, the Frankfurt-Freiburg Contrast and Acuity Test System (FF-CATS), that provides different and standardized lighting and glare conditions. In a preliminary study, we showed that CS values obtained by the FF-CATS discriminate better between younger and older normal subjects than high-contrast VA values or maximum CS values obtained with the Functional Acuity Contrast Test (FACT). Moreover, there was a higher correlation between FF-CATS values and the ocular higher-order wavefront error than FACT values.[24]

The aim of the present study was to evaluate the FF-CATS in terms of reliability and discrimination ability. To do so, the FF-CATS was compared with two well-established clinical CS tests. Normal subjects and patients with cataract were tested with the FF-CATS with a view-in tester (CST-1800) using the FACT chart[16] for a range of lighting and glare conditions. They were also tested with the Pelli-Robson Chart (PRC)[14] under photopic conditions with and without glare.

SUBJECTS AND METHODS

Subjects

Sixty eyes of 60 volunteers were included in this prospective study and were divided into three groups (Table 1). In group 1 ("young normals"), 20 healthy eyes of 20 volunteers with clear ocular media aged 21 to 47 years (median age, 29 years) were included. Group 2 ("elderly normals") consisted of 20 subjects aged over 50 years (median age, 58 years; range, 50–69 years). Group 3 ("patients with cataract") included 20 eyes of 20 patients with a senile corticonuclear cataract (median age, 68.5 years; range, 50–83 years). The inclusion criteria for groups 1 and 2 were best spectacle-corrected visual acuity (BSCVA) better than 20/25 visual acuity and clear ocular media. The exclusion criteria were any disease or medication that may affect visual function, as well as the mental inability to understand and perform the tests. The inclusion criteria for group 3 were BSCVA better than 20/40 and a senile, nuclear cataract (Lens Opacification Classification System [LOCS] III NO3 and higher). The exclusion criteria were same as for groups 1 and 2 and also the presence of any posterior subcapsular lens opacity. Manifest subjective spherocylindrical refraction was assessed using an AO type phoropter (Rodenstock; Munich, Germany) by the crosscylinder method. The end point of the manifest subjective refraction was determined with a red/green balance test. A randomization scheme (BiAS statistical software package V7.06; Epsilon Verlag; Hofheim, Germany) guaranteed equal numbers for right and left eyes of each group or subgroup (Table 1). In all experiments, informed consent was obtained after the purpose and characteristics of the study were explained using an illustrated information sheet. For all study procedures, the tenets of the declaration of Helsinki were followed.

Testing Environment and Testing Conditions

All psychophysical tests of this study were performed in the same testing laboratory where standardized lighting conditions were ensured by blocking daylight. Fluorescent lamps (Lumilux white,[835] Osram; Munich, Germany) were used to achieve diffuse, glare-free illumination. For all mesopic testing, the lamps were turned off. Luminance and illuminance were measured with a luminance meter (LS-100; Minolta; Tokyo, Japan) and illuminance meter (Illuminance Meter; Minolta), respectively.

All tests were performed monocularly with an undilated pupil and best spectacle correction using carefully cleansed trial glasses (Oculus; Wetzlar, Germany). All CS tests were done twice with an interval of at least 1 hour between the two sessions. The tests took place in two blocks: first at photopic and then, after 2 minutes of dark adaptation, at mesopic luminance conditions (Table 2). To prevent fatigue effects, the sequence of the tests in each block was determined by a randomization scheme (BiAS V7.06). The photopic test sequence included CS measurements with the FF-CATS, the FACT and the PRC, each with and without glare. At mesopic conditions, FF-CATS and FACT testing was performed, also with and without glare.

The Frankfurt-Freiburg Contrast and Acuity Test System

The Freiburg Visual Acuity and Contrast Test Program. The FF-CATS is a psychophysical test based on the FrACT.[19] The FrACT is a program that enables automatic and observer-independent determination of VA at a defined optotype contrast or CS at a specific optotype size. The FrACT uses an eight-alternative forced choice (8-AFC) and the best parameter estimation by sequential testing (PEST) algorithm for threshold determination.[25] This algorithm displays a sequence of stimuli calculating the actual threshold value and each next stimulus from the responses given (Fig. 1).

TABLE 1.
Demographic data

	Group 1	Group 2	Group 3
Age [years]	29	58	68.5
median range	(21–47)	(50–69)	(50–83)
Gender	12 F, 8 M	14 F, 6 M	10 F, 10 M
Eyes	10 R, 10 L	10 R, 10 L	10 R, 10 L
BSCVA [logMAR]	−0.25	−0.1	0.16
median range	(−0.41–0.01)	(−0.28–0.06)	(-0.2–0.29)

[a]Group 1, young normal subjects; Group 2, elderly normal subjects; Group 3, patients with cataracts.
BSCVA, Best spectacle corrected visual acuity.

TABLE 2.
Contrast sensitivity tests and test parameters used in the study

Test	Optotype	Test Strategy	Display Type Glare Source	Maximal Possible Stimulus Range (logCS)	Glare	Luminance (cd/m^2) Photopic	Luminance (cd/m^2) Mesopic	Illuminance (lx) Photopic	Illuminance (lx) Mesopic
FF-ACTS	Landolt C 1.3 logMAR (1.5 cpd)	8-AFC best PEST	High-resolution b/w CRT monitor	0.21–2.50	—	167	0.167	50	0.01
			circle of eight white LEDs in 3.2° to the center			167	0.167	50	0.32
Pelli-Robson	Sloan letters 1.3 logMAR (1.5 cpd)	26-AFC	Directly illuminated wall chart	0–2.25	+	100	ND	50	ND
			handheld BAT wide-angle glare >11.3°			100		320	
FACT (CST 1800)	Sine wave gratings at five spatial frequencies	3-AFC	Indirectly illuminated chart in a view-in test box	1.5 cpd: 0.85–2.00 3 cpd: 1.00–2.20 6 cpd: 1.08–2.26 12 cpd: 0.90–2.08 18 cpd: 0.60–1.81	+	167	0.167	74	0.20
			incandescent lamp 29.6°–48.5° from temporal		+	167	0.167	74	0.32

8-AFC, eight-alternative forced choice; best PEST, best parameter estimation by sequential testing; b/w CRT, black and white cathode ray tube; LED, light-emitting diode; BAT, Brightness Acuity Tester; ND, not done; cpd, cycles per degree.

The Frankfurt-Freiburg Contrast and Acuity Test System. For this study, version 5.6 of the program was run on an Apple computer (iMac, OS V9.1). Landolt rings of 100 in diameter (logarithm of the minimum angle of resolution [LogMAR] equivalent 1.3; dominant spatial frequency 1.5 cycles per degree [cpd]) were displayed on a high-resolution 19-inch black and white cathode ray tube monitor (Richardson/Philips GD402-WBE; Amsterdam, The Netherlands). To prevent the subject from recognizing the gap of the Landolt ring by edge detection, the ring was blurred by a small edge of 50% of the optotype contrast. The background luminance was 167 cd/m^2 and the resolution was set to 1600 × 1280 pixels at a repetition frequency of 72 Hz. The monitor luminance was linearized and the display format was adjusted using the calibration tools provided by the FrACT program. For investigative purposes, the monitor was placed in a metal box (60 × 60 × 60 cm) with a circular aperture of 20 cm in the anterior side (Fig. 2). For mesopic testing (0.167 cd/m^2), a neutral density filter with $\tau = 10^{-3}$ (NG9; Schott; Mainz, Germany) was placed inside the box behind the aperture. A circle of eight white light-emitting diodes (LED, 6400 mcd, 20° full width half maximum [FWHM] angle) served as the glare source. The LEDs were displaced at a visual angle of 3.2° from the center of the Landolt ring at steps of 45° providing an equal distance between each LED in respect to the eight potential positions of the Landolt gap. The illuminance of the glare source at the eye of the subject was regulated to 0.32 lx, a common value used in night vision test devices that allows maintenance of the dark-adapted state of the eye and prevents pupil constriction.[26,27] To prevent the subjects from accidentally looking into the glare source, the LEDs were flashed briefly at a reduced illuminance of 0.05 lx before the glare was switched on. The testing distance was 4 m. The subject indicated the positions of the perceived gap of the Landolt C on the computer key pad and was not given feedback on the correctness of each response. After a learning sequence to familiarize subjects with the task, the results of each sequence, including CS and actual contrast of each trial and the correctness of the responses, were saved and analyzed using custom-programmed MS Excel spreadsheets.

The Pelli-Robson Chart

The PRC (Clement Clarke International Ltd.; Harlow, U.K.) is a wall chart displaying Sloan letters of constant size.[14] Three letters form a triplet of equal contrast and contrast decreases from row to row by a factor of 2 (i.e., 0.15 logCS units per triplet). The test chart was illuminated by room light providing a background luminance of 100 cd/m^2 (minimum, 91 cd/m^2; maximum, 102 cd/m^2). Testing was carried out at 1 m (LogMAR equivalent 1.3) using a forced-choice strategy: 0.05 logCS were credited per letter correctly recognized as suggested earlier.[28] The Brightness Acuity Tester (BAT; Mentor ONO, Norwell, MA) was used as a glare source.[29] The BAT is a handheld device in the form of a white hemisphere (diameter 60 mm) with a central aperture of 12 mm. The subjects held the BAT close to their eye and looked through the aperture at the wall chart. The hemisphere was illuminated indirectly by an incandescent lamp (setting "high" [320 lx], according to a previous CS test comparison study[15]) providing wide-angle glare at least 11.3° from the center of the aperture.

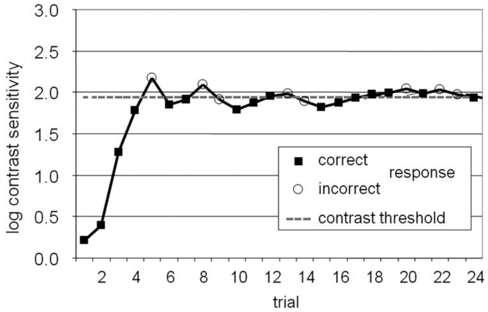

FIGURE 1.

Protocol of one single test sequence illustrating the determination of the contrast threshold (dotted line) by the best parameter estimation by sequential testing algorithm.

The Functional Acuity and Contrast Test

The FACT is a chart displaying sinusoidal gratings of five different spatial frequencies with nine contrast levels for each spatial frequency (Table 2).[16] Gratings are displayed as circular patches with blurred edges either orientated vertically or tilted 15° to the left or right. The decrements of the grating contrast are 0.15 logCS units for each spatial frequency, resulting in a range of 1.2 logCS units (Table 2). In this study, the version integrated in the Contrast Sensitivity Tester 1800 (CST 1800; Vision Science Research Corp.; San Ramon, CA) was used. A miniature version of the chart was mounted in a view-in test box and displayed at optical infinity. The luminance was set at 167 cd/m^2 for photopic and 0.167 cd/m^2 for mesopic testing to ensure comparable test conditions with the FF-CATS. An incandescent bulb providing temporal glare is integrated into the device; it was regulated to an illuminance of 0.32 lx (Table 2). Glare angles ranged from 29.6° when viewing the most temporally located column to 48.5° when viewing the most nasally located column of the FACT chart. This resulted in a smaller glare angle for right eyes viewing gratings with higher spatial frequencies and left eyes viewing gratings with lower spatial frequencies and vice versa. Chart illumination and glare were self-controlled in a closed-loop fashion with an integrated illuminance meter. Trial lenses for distance correction were placed in front of the viewing aperture. The subject had to decide if the grating was directed vertically or tilted to the right or left (3-AFC). The CS at a certain spatial frequency was the last correct response before two successive wrong responses. If a subject did not recognize the lowest CS value, CS was recorded as zero for that spatial frequency. To reduce the amount of data and enable comparison with the other tests, we analyzed CS values acquired at a spatial frequency of 1.5 cpd, and we calculated the area under the approximated contrast sensitivity function (CSF), referred to as *A* FACT that was obtained by linearly connecting the five CS readings of each FACT trial (spatial frequencies on a linear scale).[30,31]

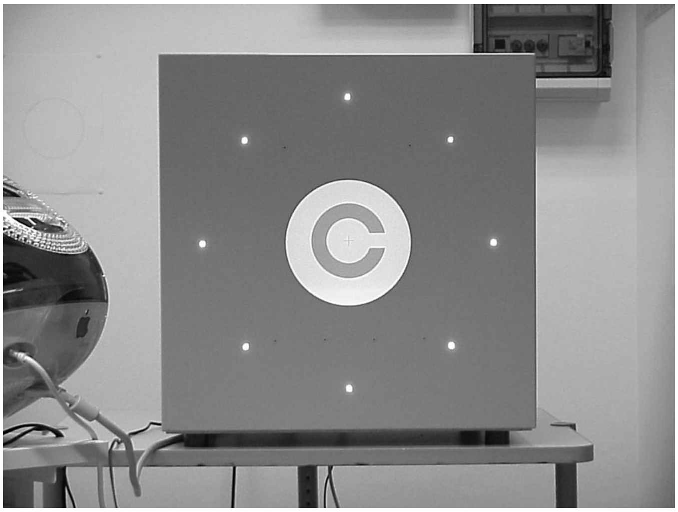

FIGURE 2.

Detail view of the monitor box with the light-emitting diode glare sources.

Statistical Analysis

The main outcome variable of all the tests was the logarithmic value of contrast sensitivity (logCS). All statistical analyses were carried out using the BiAS 7.06 software package.

Mean Values. Mean CS values for each test and condition were calculated from averaged test and retest values. Statistical significance of CS intergroup differences was assessed with the Kruskal-Wallis test and the effect of a glare source was evaluated with the paired Wilcoxon test. A $p < 0.05$ was considered statistically significant. An overall level of < 0.05 for multiple tests was maintained using a Bonferroni correction.

Reliability. Test–retest repeatability was assessed separately for groups 1, 2, and 3 using the method described by Bland and Altman.[32] The coefficient of repeatability (COR), defined as the 1.96-fold value of the standard deviation of the differences between test and retest, was used as a measure of reliability.[32] We also calculated the ratio between the COR and the mean value of test and retest. This was of interest when comparing test repeatability independently of the absolute value (e.g., photopic and mesopic CS, *A* FACT).

Discrimination Ability. Receiver operating characteristic (ROC) curves that plot "sensitivity" against "1-specificity" were computed to analyze the ability of the tests to discriminate between groups of subjects with different ocular conditions (Table 1).[33] The area under the ROC curve (*Az* ROC) served as a measurement of the discriminative power of a test; an ideal test with 100% sensitivity and specificity would reach the maximal Az ROC value of 1. ROC curves were computed from the averaged test and retest measures of CS.

RESULTS

Demographic Data and Mean Values

Statistically significant differences ($p < 0.05$) were found for age and BCVA between all groups. For all tests, CS was higher for normal young than for normal elderly subjects, whereas the patients with cataracts had the lowest CS values. However, median CS values varied among the contrast tests used (Tables 3 and 4; Figs. 3A, B and 4). FF-CATS scores were the highest values, whereas CS values obtained with the PRC were the lowest throughout all groups. There were large differences between FF-CATS and FACT values at 1.5 cpd under mesopic conditions (Fig. 3A, B). For young subjects (group 1), there was a ceiling effect of FACT values at 1.5 cpd shown by the asymmetric distribution in the box plot diagrams (Fig. 3A, B). In contrast, many older subjects

TABLE 3.
Contrast sensitivity (logCS) under photopic conditions: median and range (in parentheses)[a]

Test	Glare	Group 1	Group 2	Group 3
FF-ACTS	—	2.23 (1.98–2.42)	1.82 (1.58–2.29)	1.60 (1.28–2.04)
	+	2.19 (1.73–2.42)	1.92 (1.48–2.29)	1.59 (1.30–2.16)
Pelli-Robson	—	1.85 (1.65–1.95)	1.65 (1.45–1.88)	1.51 (1.20–1.65)
	+	1.73 (1.40–1.90)[b]	1.59 (1.33–1.73)[d]	1.43 (0.70–1.65)[c]
FACT at 1.5 cpd	—	2.00 (1.63–2.00)	1.82 (1.63–2.00)	1.63 (1.40–2.00)
	+	2.00 (1.70–2.00)	1.85 (1.56–2.00)	1.70 (1.48–2.00)
A FACT	—	33.35 (32.25–34.53)	28.84 (13.36–33.20)	20.14 (11.12–31.13)
	+	33.68 (30.03–34.53)	27.72 (14.32–33.14)	21.42 (10.12–29.43)

[a] Group 1, young normal subjects; Group 2, elderly normal subjects; Group 3, patients with cataract.
logCS difference with glare: [b] $p < 0.05$; [c] $p < 0.01$; [d] $p<0.001$ (paired Wilcoxon test with Bonferroni correction).

in groups 2 and 3 failed to recognize the lowest CS value of the mesopic FACT (data not shown).

Under photopic conditions, statistically significant CS loss resulting from glare (DG) could only be observed with the PRC and with the BAT (Table 3), whereas under mesopic conditions, a significant decrease in CS was only found for the FF-CATS (Table 4). Although patients with cataract had the lowest CS values among all tests and had a tendency for higher DG, there were no significant differences in DG values between the groups. As shown above, in the CST 1800, the position of the glare source relative to the sine wave gratings of the FACT was not constant. For the near-threshold gratings, the glare angle was smaller for right than for left eyes. However, no statistically significant difference in DG between right and left eyes was found in any group (data not shown).

Test–Retest Repeatability

The PRC had the best repeatability (COR 0.22 and 0.13 logCS, see Table 5) among photopic CS tests for normal subjects (groups 1 and 2). The FACT at 1.5 cpd and the FF-CATS had higher COR values for normal subjects, ranging from 0.33 logCS (FACT, group 1) to 0.39 logCS (FACT, group 2). In most cases, the introduction of a glare source led to a decrease in repeatability (Table 5). For patients with cataract, tests using letter optotypes (FF-CATS, PRC) had a higher repeatability. Also *A* FACT values were much less repeatable for group 3. Under mesopic conditions table, COR values for groups 1 and 2 were >0.32 logCS (FF-CATS with glare in group 2) for all tests. In group 3, results were similar: repeatability was best for the FF-CATS (0.29 logCS without and 0.15 log CS with glare), whereas FACT testing, especially for the *A* FACT, which included higher spatial frequencies, had larger differences between test and retest.

Discriminative Ability

The discriminative ability of CS was analyzed for three conditions: young vs. elderly normal subjects (group 1 vs. 2), young vs. cataract subjects (group 1 vs. 3), and elderly normal vs. cataract subjects (group 2 vs. 3). Results are shown in Table 7. In general, CS tests showed best discrimination between groups 1 and 3 and a better discrimination between groups 1 and 2 than between groups 2 and 3. Under photopic conditions, the FF-CATS, PRC, and *A* FACT showed a similar ability to distinguish between the three groups. FACT CS values obtained at 1.5 cpd had a markedly lower discrimination ability. Except for the FACT value at 1.5 cpd, testing with glare did not improve discrimination between the groups throughout the tests. Under mesopic conditions, all tests discerned best between young and cataract subjects. Only the FF-CATS was able to sufficiently distinguish between older healthy and cataract subjects (group 2 vs. 3).

DISCUSSION

Comparison of Mean Contrast Sensitivity and Disability Glare Values

Subjects obtained the highest photopic CS scores with the FF-CATS. For the FACT at 1.5 cpd, the highest possible score was 2.0 logCS. Most of the young subjects (group 1, Fig. 3A)

TABLE 4.
Contrast sensitivity (logCS) under mesopic conditions: median and range (in parentheses)[a]

Test	Glare	Group 1	Group 2	Group 3
FF-ACTS	—	1.23 (0.85–1.58)	1.00 (0.65–1.26)	0.78 (0.34–1.17)
	+	0.90 (0.68–1.18)[b]	0.55 (0.21–0.88)[b]	0.21 (0.21–0.58)[b]
FACT at 1.5 cpd	—	1.78 (1.48–2.00)	1.60 (1.26–2.00)	1.40 (0.98–2.00)
	+	1.78 (1.56–2.00)	1.63 (1.26–2.00)	1.42 (0.95–2.00)
A FACT	—	21.45 (14.93–30.11)	11.85 (2.07–23.62)	8.33 (0.74–21.15)
	+	21.03 (14.95–26.43)	11.34 (3.20–23.85)	7.43 (0.71–24.50)

[a] Group 1, young normal subjects; Group 2, elderly normal subjects; Group 3, patients with cataract.
logCS difference with glare: [b] $p < 0.001$ (paired Wilcoxon test with Bonferroni correction).

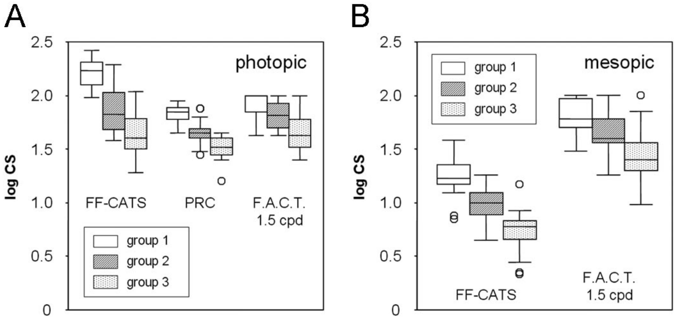

FIGURE 3.

Box plot diagrams showing median contrast sensitivity values without glare. (**A**) Frankfurt-Freiburg Contrast and Acuity Test System (FF-CATS), Pelli-Robson chart, and Functional Acuity Contrast Test (FACT) at 1.5 cpd (photopic conditions). (**B**) FF-CATS and FACT at 1.5 cpd (mesopic conditions).

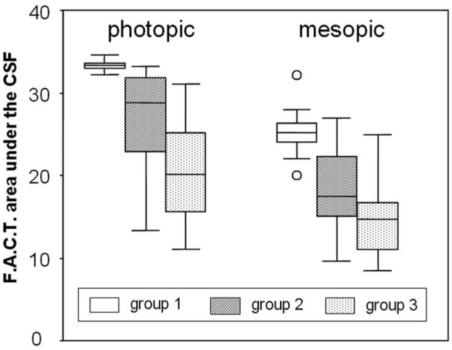

FIGURE 4.

Box plot diagrams showing median contrast sensitivity values for the area under the contrast sensitivity function obtained with the Functional Acuity Contrast Test (*A* FACT) without glare.

reached the 2.0 log CS score, which was reflected by the asymmetric distribution in the box plot diagrams (Fig. 3A, B). Most likely, their real CS is higher than the highest possible FACT score (ceiling effect). This effect has been described in detail in a previous study and is attributed to the limited range of 1.2 logCS of the FACT chart.[18] In contrast to the FF-CATS, the reduction of luminance by a factor of 1000 only marginally affected FACT scores (Fig. 3A, B). Several factors may be responsible for the discrepancy in mesopic CS scores obtained with FACT and FF-CATS. First, although luminance levels were identical for both tests (0.167 cd/m^2), the effective illuminance was higher (0.2 lx) for the FACT because the area of the illuminated background chart was much larger. Second, the fact that grating CS has been reported to be higher than letter CS might be another reason for the discrepancy.[34] Finally, it is possible that a drop in the contrast threshold from values above the highest score into the range of the FACT when luminance was reduced could not be recognized as a result of the ceiling effect. The PRC did not show a definite ceiling effect, although values were distributed much narrower around the medians. The discrimination task was more complex for the PRC because the letters had to be recognized completely compared with simple gap recognition for the FF-CATS. Besides this fact, a lower probability of correct guessing (26- vs. 8-AFC) and the lower chart luminance (100 cd/m^2) could be responsible for lower CS values measured with the PRC throughout all groups.

The glare sources influenced CS scores differently in each test. The additional low-illuminance glare sources designed to simulate night glare that were used for FACT and FF-CATS did not affect photopic DG significantly (Tables 3 and 4) because photopic CS scores were already affected by wide-angle scatter from the background of the charts and surroundings.[15,35] Only the wide-angle glare provided by the BAT, which was 1000-fold higher (320 lx) than the glare of FACT and FF-CATS, was able to reduce CS significantly. However, because of its high illuminance, the BAT is inappropriate for mesopic testing because it is likely to change the adaptation state of the eye and provoke pupil constriction. This is especially relevant in refractive surgery trials in which pupil constriction would mask possible effects of induced aberrations on visual performance.[36,37] Although glare illuminance was the same for FACT and FF-CATS, their glare angle (3.2° for FF-CATS, 29.6°–48.5° for FACT) and their configuration (single incandescent lamp vs. circle of white LEDs) were different, which could explain the marginal effect of FACT glare.[38] Therefore, it would be reasonable to test only under mesopic conditions with glare or to adapt the configuration and illuminance of the glare source to the specific situation to be simulated by the test.

TABLE 5.

Reliability of photopic contrast sensitivity (logCS) measurements: coefficient of repeatability (COR) in logCS and ratio between COR and mean between test and retest (COR ratio)[a]

		COR (log CS)			COR Ratio		
Test	Glare	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
FF-ACTS	—	0.38	0.35	0.22	0.18	0.20	0.13
	+	0.34	0.35	0.22	0.16	0.18	0.14
Pelli-Robson	—	0.22	0.13	0.15	0.12	0.08	0.11
	+	0.46	0.17	0.16	0.26	0.11	0.11
FACT at 1.5 cpd	—	0.33	0.39	0.30	0.17	0.21	0.18
	+	0.24	0.33	0.37	0.12	0.18	0.22
A FACT	—	3.10	4.98	7.54	0.09	0.19	0.37
	+	2.41	8.09	4.42	0.07	0.31	0.22

[a] Group 1, young normal subjects; Group 2, elderly normal subjects; Group 3, patients with cataract.

TABLE 6.
Reliability of mesopic contrast sensitivity (logCS) measurements: coefficient of repeatability (COR) in logCS and ratio between COR and mean between test and retest (COR ratio)[a]

		COR (log CS)			COR Ratio		
Test	Glare	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
FF-ACTS	—	0.46	0.55	0.29	0.37	0.55	0.39
	+	0.36	0.32	0.15	0.40	0.55	0.51
FACT at 1.5 cpd	—	0.55	0.35	0.37	0.30	0.21	0.25
	+	0.57	0.36	0.31	0.32	0.22	0.22
A FACT	—	10.64	8.43	11.10	0.50	0.67	1.27
	+	10.16	4.58	12.47	0.47	0.37	1.60

[a] Group 1, young normal subjects; Group 2, elderly normal subjects; Group 3, patients with cataract.

Test–Retest Repeatability

Repeatability results for the PRC and for the FACT are consistent with earlier studies.[15,18,28] For subjects in group 1, the reliability of FACT scores was significantly lower under mesopic conditions because the ceiling effect led to the apparently higher repeatability of photopic CS scores. The effect of randomly occurring variance of CS at different spatial frequencies could be mitigated by calculating the area under the CSF (*A* FACT). For healthy subjects (group 1 and 2), *A* FACT scores showed good reliability that decreased with introduction of a glare source. Repeatability was also lower in patients with cataracts, which could be explained by a poor repeatability at higher spatial frequencies (data not shown). COR values for the FF-CATS were higher as expected for a PEST-based forced-choice test with 24 trials. One possible reason is that the smaller step size compared with the PRC uncovers a random variance of CS values. Moreover, training effects could also play a role because there was a tendency toward higher values for the repeat trial (data not shown). It remains to be determined whether the best PEST algorithm has to be modified for CS measurements. Incorrect responses, e.g., as a result of fatigue during the initial phase of the trial sequence (trials 4–8), might lead to underestimation of CS because the step size remains small later throughout the sequence.

Discrimination Ability

One basic property of psychophysical tests is the ability to distinguish between different conditions. It is known that retinal image quality decreases with age as a result of increasing light scatter and optical aberrations.[39] Correspondingly, CS decreases as a function of age.[40,41] To evaluate discrimination ability of the three tests used in this study, we used a setup of three groups of 20 eyes in each group. Older cataract subjects (group 3) served as positive controls, whereas the other two groups (groups 1 and 2) consisted of healthy eyes. All tests performed satisfactory when separating young normal subjects from patients with cataracts (1 vs. 3). Interestingly, addition of a glare source did not improve selectivity in general because wide-angle scatter from the chart background might already have decreased CS in patients with cataracts. The FACT scores (1.5 cpd and *A* FACT) failed to discriminate between groups 2 and 3 in most cases. This was because of the limited range of the FACT chart contrast thresholds, because most older normal subjects and patients with cataracts were below the lowest CS step (floor effect, data not shown). Consequently, use of a glare source did not improve discriminative ability.

Overall Rating of the Tests

The Frankfurt-Freiburg Contrast and Acuity Test System. The present study has shown that the FrACT with its extension, the

TABLE 7.
Discriminative ability: area under the receiver operating curve (A_zROC) for contrast sensitivity (CS) measurements[a]

Test		A_z ROC Photopic			A_z ROC Mesopic		
	Glare	1 vs. 2	1 vs. 3	2 vs. 3	1 vs. 2	1 vs. 3	2 vs. 3
FF-ACTS	—	0.90[d]	0.99[d]	0.80	0.82	0.97[d]	0.83
	+	0.80	0.97[b]	0.84[c]	0.90[b]	1.00[d]	0.87[b]
Pelli-Robson	—	0.89[b]	1.00[d]	0.83		ND	
	+	0.89[b]	0.96[d]	0.80		ND	
FACT at 1.5 cpd	—	0.74	0.85[d]	0.75	0.72	0.90[d]	0.80
	+	0.74	0.91[d]	0.75	0.76	0.88[d]	0.74
A FACT	—	0.95[d]	1.00	0.77[d]	0.91[d]	0.97[d]	0.71
	+	0.96[d]	1.00	0.77[d]	0.91[d]	0.96[d]	0.74

[a] For each CS test, p values of intergroup comparison of CS values as obtained by the Kruskal-Wallis test are shown. Group 1, young normal subjects; Group 2, elderly normal subjects; Group 3, patients with cataract.

Kruskal-Wallis test (intergroup comparison of CS values): [b]$p < 0.05$; [c]$p < 0.01$; [d]$p < 0.001$ (with Bonferroni correction).

ND, not done.

FF-CATS, is a valuable psychophysical test for determining CS at different luminance levels with and without glare. Because for all luminance and glare conditions the same test procedure is used, the influence of luminance and glare on CS could be studied. Its test strategy follows the criteria that have been deemed necessary for psychophysical tests[15,18]: a wide range of stimuli, forced-choice testing, a low probability of guessing, and prevention of learning effects by randomized computer presentation. The latter is an important requirement for repeated testing under different luminance and glare conditions. Recognizing the gap of a Landolt C is a simple task that most patients are familiar with. However, because the PEST algorithm presents most of the stimuli around the assumed threshold[19,25] (Fig. 1), testing may be exhausting. The typing task requires some dexterity, especially in elderly patients. To avoid change in accommodative state, patients should be advised not to look at the keypad but to localize the correct key digitally. Therefore, Wesemann recommended the subject give the answer verbally with the examiner typing it into the computer.[21] Typing errors in the initial phase of the sequence may be one cause of the somewhat lower reliability of the FF-CATS compared with the PRC. Larger CS steps, as presented by PRC and FACT, helped to smooth out this noise but also contained a risk of lower sensitivity. In fact, the FF-CATS showed the highest ability of all the tests to discriminate between the groups for all combinations and all conditions of luminance and glare.

The Landolt ring used has a dominant spatial frequency of 1.5 cpd. For a comprehensive determination of functional parameters of quality of vision, a test at a higher spatial frequency would also be required. By additional testing of high-contrast VA with the FF-CATS, the two most relevant points of the CSF could be measured.[18,24,28] In its present form, the FF-CATS is a test designed for clinical studies in a laboratory environment. Changes of setup or testing modes (size of the Landolt ring, VA testing at defined contrast levels, modifications of background luminance by neutral filters, changes of glare illuminance, recording answering time) could easily be performed if required for special investigations.

The Pelli-Robson Chart. Although for the PRC, luminance and glare conditions were significantly different from the two other tests, we used it as a reference method because it had been proven to be a test with good reliability.[15,28,42] It is an optotype-based test; the letter "C" of the PRC equals the Landolt ring. It is easy to administer because patients are familiar to Sloan letter optotypes. With a 26-AFC, guessing probability is low and smaller steps than 0.15 logCS could be simulated using a by-letter scoring system.[28] A drawback that motivated us to build a modification of the FrACT for CS testing was the lack of a glare source that provided equal conditions for all letter triplets. The BAT is not useful for mesopic glare testing because it both induces pupil constriction and changes the adaptive state of the eye. Moreover, for repetitive trials at different luminance levels, more than two charts are required to prevent learning effects.

The Functional Acuity Contrast Test. The FACT chart has become a popular CS test with widespread use because of its simple and clear testing procedure. Yet several studies, including this study, show that the FACT and its predecessor, the Vistech Chart, have problems such as low repeatability, high probability of correct guessing, and limited range of presented contrasts all limit their usefulness.[15,18,43] The ceiling effect in young healthy subjects (group 1) makes the FACT unsuitable for use in refractive surgery outcomes research, whereas the floor effect in older subjects (groups 2 and 3, data not shown) makes the test unsuitable for use in studies of cataract surgery. The data from five spatial frequencies could be easier to interpret and more reliable if, for example, the area under the CSF (*A* FACT) was specified; that has the potential advantage of containing information from all spatial frequencies.[30,31] However, the poor reliability of the mesopic FACT at higher spatial frequencies (data not shown) may affect repeatability of a FACT scored. As mentioned previously and demonstrated by factor analysis in a recent study, the Vistech and FACT charts need only two components, one representing low and the other representing high spatial frequencies to determine quality of vision comprehensively.[18] It should be mentioned that the luminance conditions in the present study were different from those recommended by the manufacturer; they were lower in our study. The 3 cd/m^2, recommended by the manufacturer as "scotopic" lighting conditions, is not actually scotopic. Under higher luminance levels for mesopic testing (3 or 6 cd/m^2), a more pronounced ceiling effect is likely to occur as well. It remains to be determined if a higher glare illuminance would have been more useful for discrimination between normal and cataract subjects. However, the variability of the glare angle (29.6°–48.5°) in the CST 1800 and the asymmetry between right and left eyes are more likely to have an effect if the glare source induced significant DG. Although the CST 1800 with the FACT allows approximation of the CSF under variable and controlled lighting and glare conditions without a complex laboratory setup, it contains certain flaws that make it of questionable use for clinical studies.

In summary, the perfect contrast test has not been found yet. The present study showed that different tests show different characteristics and results may not be interchangeable. For testing at different luminance levels, the FF-CATS, a laboratory application of the FrACT, fulfilled the requirements to comprehensively assess quality of vision. Further studies will be needed to show its usefulness for clinical trials in the field of cataract and refractive surgery.

ACKNOWLEDGMENTS

This study was supported by a grant from the German Ophthalmological Society (Forschungsförderungspreis der DOG 2001 [JB]).

Presented in part at the 2003 annual meeting of the Association for Research in vision and Ophthalmology (ARVO), Fort Lauderdale, Florida; at the 2003 annual meeting of the German Society for Intraocular Lens Implantation and Refractive Surgery (DGII), Ludwigshafen, Germany; at the 2003 annual meeting of the German Ophthalmological Society (DOG) Berlin, Germany; at the 2004 Refractive Surgery Winter Meeting of the European Society of Cataract and Refractive Surgeons (ESCRS), Barcelona, Spain.

The authors have no proprietary interest in any programs or of the devices used in this study.

Received June 14, 2005; accepted February 14, 2006.

REFERENCES

1. Brown B, Lovie-Kitchin JE. High and low contrast acuity and clinical contrast sensitivity tested in a normal population. Optom Vis Sci 1989;66:467–73.

2. Haegerstrom-Portnoy G, Schneck ME, Lott LA, Brabyn JA. The relation between visual acuity and other spatial vision measures. Optom Vis Sci 2000;77:653–62.
3. Kohnen T. Measuring vision in refractive surgery. J Cataract Refract Surg 2001;27:1897–8.
4. McLeod SD. Beyond Snellen acuity: the assessment of visual function after refractive surgery. Arch Ophthalmol 2001;119:1371–3.
5. Fan-Paul NI, Li J, Miller JS, Florakis GJ. Night vision disturbances after corneal refractive surgery. Surv Ophthalmol 2002;47:533–46.
6. Wilson SE. Wave-front analysis: are we missing something? Am J Ophthalmol 2003;136:340–2.
7. Pesudovs K, Marsack JD, Donnelly WJ, Thibos LN, Applegate RA. Measuring visual acuity—mesopic or photopic conditions, and high or low contrast letters? J Refract Surg 2004;20:S508–14.
8. Schallhorn SC, Blanton CL, Kaupp SE, Sutphin J, Gordon M, Goforth H Jr, Butler FK Jr. Preliminary results of photorefractive keratectomy in active-duty United States Navy personnel. Ophthalmology 1996;103:5–22.
9. Lee YC, Hu FR, Wang IJ. Quality of vision after laser in situ keratomileusis: influence of dioptric correction and pupil size on visual function. J Cataract Refract Surg 2003;29:769–77.
10. Yamane N, Miyata K, Samejima T, Hiraoka T, Kiuchi T, Okamoto F, Hirohara Y, Mihashi T, Oshika T. Ocular higher-order aberrations and contrast sensitivity after conventional laser in situ keratomileusis. Invest Ophthalmol Vis Sci 2004;45:3986–90.
11. Kohnen T, Bühren J, Kasper T, Terzi E. Quality of vision after refractive surgery. In: Kohnen T, Koch DD, eds. Cataract and Refractive Surgery. Berlin: Springer;2005:303–22.
12. Rabin J. Luminance effects on visual acuity and small letter contrast sensitivity. Optom Vis Sci 1994;71:685–8.
13. Pesudovs K, Schoneveld P, Seto RJ, Coster DJ. Contrast and glare testing in keratoconus and after penetrating keratoplasty. Br J Ophthalmol 2004;88:653–7.
14. Pelli DG, Robson JG, Wilkins AJ. Designing a new letter chart for measuring contrast sensitivity. Clin Vision Sci 1988;2:187–99.
15. Elliott DB, Bullimore MA. Assessing the reliability, discriminative ability, and validity of disability glare tests. Invest Ophthalmol Vis Sci 1993;34:108–19.
16. Ginsburg AP. Next generation contrast sensitivity testing. In: Rosenthal BP, Cole RG, eds. Functional assessment of Low Vision. St Louis: Mosby-Year Book Inc;1996:77–88.
17. Rabin J, Wicks J. Measuring resolution in the contrast domain: the small letter contrast test. Optom Vis Sci 1996;73:398–403.
18. Pesudovs K, Hazel CA, Doran RM, Elliott DB. The usefulness of Vistech and FACT contrast sensitivity charts for cataract and refractive surgery outcomes research. Br J Ophthalmol 2004;88:11–6.
19. Bach M. The Freiburg Visual Acuity test—automatic measurement of visual acuity. Optom Vis Sci 1996;73:49–53.
20. Bach M, Kommerell G. [Determining visual acuity using European normal values: scientific principles and possibilities for automatic measurement.] Klin Monatsbl Augenheilkd 1998;212:190–5.
21. Wesemann W. [Visual acuity measured via the Freiburg visual acuity test (FVT), Bailey Lovie chart and Landolt Ring chart.] Klin Monatsbl Augenheilkd 2002;219:660–7.
22. Loumann Knudsen L. Visual acuity testing in diabetic subjects: the decimal progression chart versus the Freiburg visual acuity test. Graefes Arch Clin Exp Ophthalmol 2003;241:615–8.
23. Dennis RJ, Beer JM, Baldwin JB, Ivan DJ, Lorusso FJ, Thompson WT. Using the Freiburg Acuity and Contrast Test to measure visual performance in USAF personnel after PRK. Optom Vis Sci 2004;81:516–24.
24. Terzi E, Bühren J, Wesemann W, Kohnen T. [Frankfurt-Freiburg Contrast and Acuity Test System (FF-CATS). A new test to determine contrast sensitivity under variable ambient and glare luminance levels.] Ophthalmologe 2005;102:507–13.
25. Lieberman HR, Pentland A. Microcomputer-based estimation of psychophysical thresholds: the best PEST. Behav Res Methods Instrum Comput 1982;14:21–5.
26. Rassow B. [Effect of luminance on contrast sensitivity and glare in the mesopic range.] Klin Monatsbl Augenheilkd 1999;214:401–6.
27. Kohnen T, Terzi E, Bühren J, Kohnen EM. Comparison of a digital and a handheld infrared pupillometer for determining scotopic pupil diameter. J Cataract Refract Surg 2003;29:112–7.
28. Elliott DB, Bullimore MA, Bailey IL. Improving the reliability of the Pelli-Robson contrast sensitivity test. Clin Vis Sci1991;6: 471–5.
29. Smith PW, Pratzer KA, Webster N, Fenton J, Bonham RD. A clinical comparison of two methods of glare testing. Ophthalmic Surg 1987; 18:680–2.
30. Marcos S. Aberrations and visual performance following standard laser vision correction. J Refract Surg 2001;17:S596–601.
31. Applegate RA, Hilmantel G, Howland HC, Tu EY, Starck T, Zayac EJ. Corneal first surface optical aberrations and visual performance. J Refract Surg 2000;16:507–14.
32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
33. Swets JA. Sensitivities and specificities of diagnostic tests. JAMA 1982;248:548–50.
34. Ginsburg AP. Visual information processing based on spatial filters constrained by biological data. PhD dissertation; Cambridge University; 1978.
35. van den Berg TJ. On the relation between glare and straylight. Doc Ophthalmol 1991;78:177–81.
36. Martinez CE, Applegate RA, Klyce SD, McDonald MB, Medina JP, Howland HC. Effect of pupillary dilation on corneal optical aberrations after photorefractive keratectomy. Arch Ophthalmol 1998; 116:1053–62.
37. Bühren J, Kühne C, Kohnen T. Influence of pupil and optical zone diameter on higher-order aberrations after wavefront-guided myopic LASIK. J Cataract Refract Surg 2005;31:2272–80.
38. Paulsson LE, Sjostrand J. Contrast sensitivity in the presence of a glare light. Theoretical concepts and preliminary clinical studies. Invest Ophthalmol Vis Sci 1980;19:401–6.
39. Artal P, Ferro M, Miranda I, Navarro R. Effects of aging in retinal image quality. J Opt Soc Am A Opt Image Sci Vis 1993;10:1656–62.
40. Mantyjarvi M, Laitinen T. Normal values for the Pelli-Robson contrast sensitivity test. J Cataract Refract Surg 2001;27:261–6.
41. Puell MC, Palomo C, Sanchez-Ramos C, Villena C. Normal values for photopic and mesopic letter contrast sensitivity. J Refract Surg 2004;20:484–8.
42. Reeves BC, Wood JM, Hill AR. Reliability of high- and low-contrast letter charts. Ophthal Physiol Opt 1993;13:17–26.
43. Reeves BC, Wood JM, Hill AR. Vistech VCTS 6500 charts—within- and between-session reliability. Optom Vis Sci 1991;68:728–37.

Thomas Kohnen
Johann Wolfgang Goethe-University
Department of Ophthalmology
Theodor-Stern-Kai 7 60590
Frankfurt am Main, Germany
e-mail: kohnen@em.uni-frankfurt.de