

Comparing Spatial Contrast Sensitivity Functions Measured With Digit and Grating Stimuli

Haiyan Zheng¹, Menglu Shen¹, Xianghang He¹, Rong Cui¹, Luis Andres Lesmes², Zhong-Lin Lu^{3,4}, and Fang Hou¹

¹ School of Ophthalmology & Optometry and Eye Hospital, Wenzhou Medical University, Wenzhou, Zhejiang, China

² Adaptive Sensory Technology, Inc, San Diego, CA, USA

³ Division of Arts and Sciences, and NYU-ECNU Institute of Cognitive Neuroscience, NYU Shanghai, Shanghai, China

⁴ Center for Neural Science and Department of Psychology, New York University, New York, NY, USA

Correspondence: Fang Hou, School of Ophthalmology & Optometry and Eye Hospital, Wenzhou Medical University, 270 Xue Yuan Xi Rd, 315 Teaching Building, Wenzhou, Zhejiang 325027, China. e-mail: houf@mail.eye.ac.cn

Received: 18 February 2019

Accepted: 1 September 2019

Published: 15 November 2019

Keywords: contrast sensitivity function; Sloan digit; grating; precision; confusion matrix

Citation: Zheng H, Shen M, He X, Cui R, Lesmes LA, Lu Z-L, Hou F.

Comparing spatial contrast sensitivity functions measured with digit and grating stimuli. *Trans Vis Sci Tech.* 2019;8(6):16, <https://doi.org/10.1167/tvst.8.6.16>

Copyright 2019 The Authors

Purpose: The contrast sensitivity function (CSF) is measured traditionally with grating stimuli. Recently, we introduced a new set of digit stimuli to improve the efficiency of CSF tests for people unfamiliar with the Latin alphabet. Given the significant differences between grating and digit stimuli, we conducted this study to evaluate whether the estimated CSFs from the digit test are equivalent to those from the grating test.

Methods: The CSFs of five young (with Psi) and five older (with quick CSF [qCSF]) participants were measured with a two-alternative forced choice (2AFC) grating orientation identification task and a 10-digit identification task. The CSFs obtained from the two tasks were compared.

Results: The estimated CSFs from the two tasks matched well after controlling for stimulus types and performance levels. The root mean square error (RMSE) between the CSFs from the two tasks was 0.093 ± 0.029 (300 trials) and 0.131 ± 0.016 (100 trials) log₁₀ units for young and older observers, respectively. To reach the same standard deviation (0.1 log₁₀ units), the digit CSF test required fewer trials/less time than the classic grating CSF for young (60 vs. 90 trials) and older (15 vs. 21 trials) observers. The complicated behavioral responses of the observer in the 10-AFC digit identification task can be accounted by a model that consists of digit similarity and one single parameter of sensory noise ($\chi^2[99] = 3.42, P = 0.999$).

Conclusions: The estimated CSFs from the digit test highly matched those obtained from the grating test; however, the digit test is much more efficient.

Translational Relevance: The digit CSF test provides a compatible assessment of the CSF as the traditional grating CSF test with more efficiency.

Introduction

The contrast sensitivity function (CSF), which delineates how visual sensitivity varies with spatial frequency, provides a comprehensive characterization of spatial vision in basic and clinical applications.^{1–5} Since the discovery of multiple spatial frequency channels in vision and the introduction of Fourier analysis into vision science by the pioneering work of Campbell and Robson,⁶ sinewave gratings have been widely used in vision research^{7–11} and especially in contrast sensitivity tests, including the various CSF

charts, such as the Functional Acuity Contrast Test (FACT),^{1,12} CSV1000,¹³ and Vistech Contrast Test,¹⁴ and computer-based procedures.¹⁵

CSF tests with grating stimuli are based on either a Yes–No detection or a two-alternative/interval forced-choice paradigm (2AFC) with two response alternatives. There is ample evidence that the test efficiency of psychophysical procedures increases with the number of task alternatives.^{16–22} The efficiency of CSF tests with grating stimuli is limited by the number of response alternatives. One possible solution is using an *N*-alternative forced choice orienta-

tion discrimination task with $N > 2$. The problem is that gratings do not have a lot of different and readily named orientations. When more than four orientations are involved, grating orientation identification may introduce some cognitive load for the participants,²⁰ which makes it very difficult to improve the efficiency of the grating test.²³

Another solution is to use characters, such as letters or digits, as stimuli in N -alternative forced-choice tasks to improve the efficiency of CSF tests. There is an added advantage that characters have distinct spatial structures and can be readily recognized without much effort. Indeed, a number of recent developments have shown that CSF tests with 10-alternative forced choice tasks using either Sloan letters²⁴ or digits²⁵ are much more efficient than those with 2AFC grating tests. Because of the presence of multiple spatial channels in the visual system,⁶ the contrast sensitivity function; that is, contrast sensitivity as a function of spatial frequency, measured with narrow-band stimuli provides a more comprehensive assessment of vision.^{15,24,26} Many studies have used bandpass filtered letter or digit stimuli to assess the CSF.^{24,25,27,28}

The CSF measured with grating stimuli has been used as the reference standard in spatial vision testing.^{29,30} For the newly developed CSF tests, a very important question is: Are the estimated CSFs from the new tests with bandpass-filtered characters equivalent to those from the grating tests? We attempt to answer the question for the CSF measured with the digit stimuli.

The 10AFC digit identification paradigm differs from the 2AFC grating task in many ways. First, digit stimuli, unlike gratings, are nonperiodic, which means the Michelson contrast varies from location to location in a digit image. The threshold of a spatial frequency is determined by the contrast energy,³¹ which is the sum of squared contrast at each spatial location of the stimulus and accounts for effects of luminance distribution and stimulus area.^{5,32,33} Thus the grating and digit stimuli with the same Michelson contrast value have different total contrast energy and result in drastically different threshold/sensitivity measures. Second, CSF tests in the literature use gratings of either equal size or equal number of cycles across different spatial frequencies.³⁴ Digit stimuli are by definition of equal number of cycles. Third, the contrast thresholds measured in the 10AFC and 2AFC tasks may correspond to different performance levels. Fourth, while most 2AFC grating CSF tests use equivalent and orthogonal stimuli, letter or digit

stimuli have different legibilities.^{35,36} The uneven similarities can make some alternatives more favored than others and, therefore, affect the test precision. Although the bandpass-filtered digit stimuli used in our previous study had a greatly improved similarity structure,²⁵ the effect of nonorthogonal stimuli on the estimated contrast sensitivities has not been systematically investigated to our knowledge.

We evaluated the similarity/difference between the CSFs obtained with grating and digit stimuli in young (Experiment 1) and older (Experiment 2) observers. To analyze the data using psychophysical models and account for potential differences among digits, an enormous amount of data were collected from each observer in Experiment 1. A number of analyses were done to evaluate the similarity/difference between CSFs obtained from the two tasks and to test the underlying assumptions in the digit CSF test.

Experiment 1

Method

Observers

Five graduate students (mean age 26 ± 1.58 years; age range, 24–28 years) at Wenzhou Medical University participated in Experiment 1. All participants underwent detailed ophthalmologic and optometric examinations performed by the first and second authors. All five observers (Y1–Y5) had normal or corrected-to-normal vision ($\log\text{MAR} \leq 0$) and no history of mental diseases. The observers showed no sign of any eye disease. They were free from any systemic diseases. Detailed characteristics of the participants are listed in Table 1. All participants except one author (Y5) were naive to the purpose of the study. The study protocol adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board of human subject research of the Eye Hospital at Wenzhou Medical University. Written informed consent was obtained before the experiment for all observers.

Apparatus

The programs used in the study were written in MATLAB (MathWorks, Natick, MA) with Psychtoolbox extensions³⁷ and run on an Apple Mac mini computer (Model No. A1347; Apple Inc., Cupertino, CA). Stimuli were displayed on a 27-inch ASUS PG279Q monitor (Asus Corp. Taipei, Taiwan) with a 2560×1440 pixel resolution and a 60 HZ refresh rate. The display was gamma-corrected

Table 1. Clinical Characteristics of All Observers

Subject	Age/Sex	Refractive Correction (1.34 m, OD/OS)	LogMAR VA (OD/OS)	Correction Method	MMSE (point)
Y1	24/M	−6.00 −5.50	0 0	Spectacles	–
Y2	26/M	−6.75/−1.00 × 25 −6.00/−0.75 × 85	0 −0.1	Spectacles	–
Y3	28/F	−7.50 −7.50	0 0	Spectacles	–
Y4	27/F	−7.50 −7.50	−0.1 −0.1	Spectacles	–
Y5	25/F	−7.25/−1.00 × 35 −6.00/−1.25 × 160	0 0	Spectacles	–
O1	66/M	+1.50/−1.50 × 75 +1.50/−1.00 × 95	0 0	Spectacles	27
O2	61/F	+2.50/−0.50 × 70 +2.75/−0.50 × 115	−0.1 −0.1	Spectacles	28
O3	65/F	+0.50/−0.75 × 90 −1.25	−0.1 −0.1	Spectacles	27
O4	60/M	+1.50/−1.50 × 80 +1.50/−1.50 × 95	−0.1 −0.1	Spectacles	28
O5	73/M	+3.25/−1.50 × 70 +2.50/−1.25 × 90	0 0	Spectacles	30

Y1–Y5, represent observers who participated in Experiment 1; O1–O5, represent observers who participated in Experiment 2; F, female; M, male; OD, right eye; OS, left eye; LogMAR VA, LogMAR Visual Acuity; MMSE, Mini-Mental State Examination.

with a photometer (ST-86A; Photoelectric Instrument Factory of Beijing Normal University, Beijing, China) and had a mean luminance of 91.2 cd/m². A bit-stealing algorithm was used to achieve 9-bit gray-scale resolution.³⁸ Observers viewed the display binocularly with their best refractive corrections by wearing glasses in a dark room at a distance of 1.34 m. A chin/forehead rest was used to minimize head movement during the experiment.

Stimuli

Sinewave gratings and bandpass-filtered digits were used as stimuli in the experiment. We generated Sloan digits following the specifications for Sloan letters.³⁹ The detailed specifications of the digits have been reported in our parallel study.²⁵ Because the unfiltered digits contains a range of spatial frequencies,²⁶ all digits were filtered with a raised cosine filter^{24,26} to create bandpass-filtered stimuli (Fig. 1a). The center frequency of the filter is 3 cycles per object (cpo) and the full bandwidth at half height is one octave. The filtered digits, resized to 12°, 6°, 3°, 1.5°,

0.75°, and 0.38°, corresponding to central spatial frequencies of 0.5, 1, 2, 4, 8, and 15.8 cycles per degree (Fig. 1c), were used as signal stimuli in the 10-digit identification task.

The grating stimuli were static sinusoidal gratings oriented either +45° or −45° from vertical (Fig. 1b). Each stimulus contained six cycles of a sinewave. A half-Gaussian envelope ($\sigma = \lambda/3$) was used to blend the grating into the background. The spatial envelope resulted in a four-cycle full contrast circular window. By rescaling the grating images to different sizes, we generated grating stimuli with center spatial frequencies of 0.5, 1, 2, 4, 8, and 15.8 cycles per degree. The grating stimulus had a random selected phase from 0 to 2π in each trial.

Design

In the experiment, the CSF of each participant was measured with a 2AFC grating orientation identification task and a 10-digit identification task in separate blocks. Contrast thresholds at 0.50, 1, 2, 4, 8, and 15.8 cpd were measured using the Psi

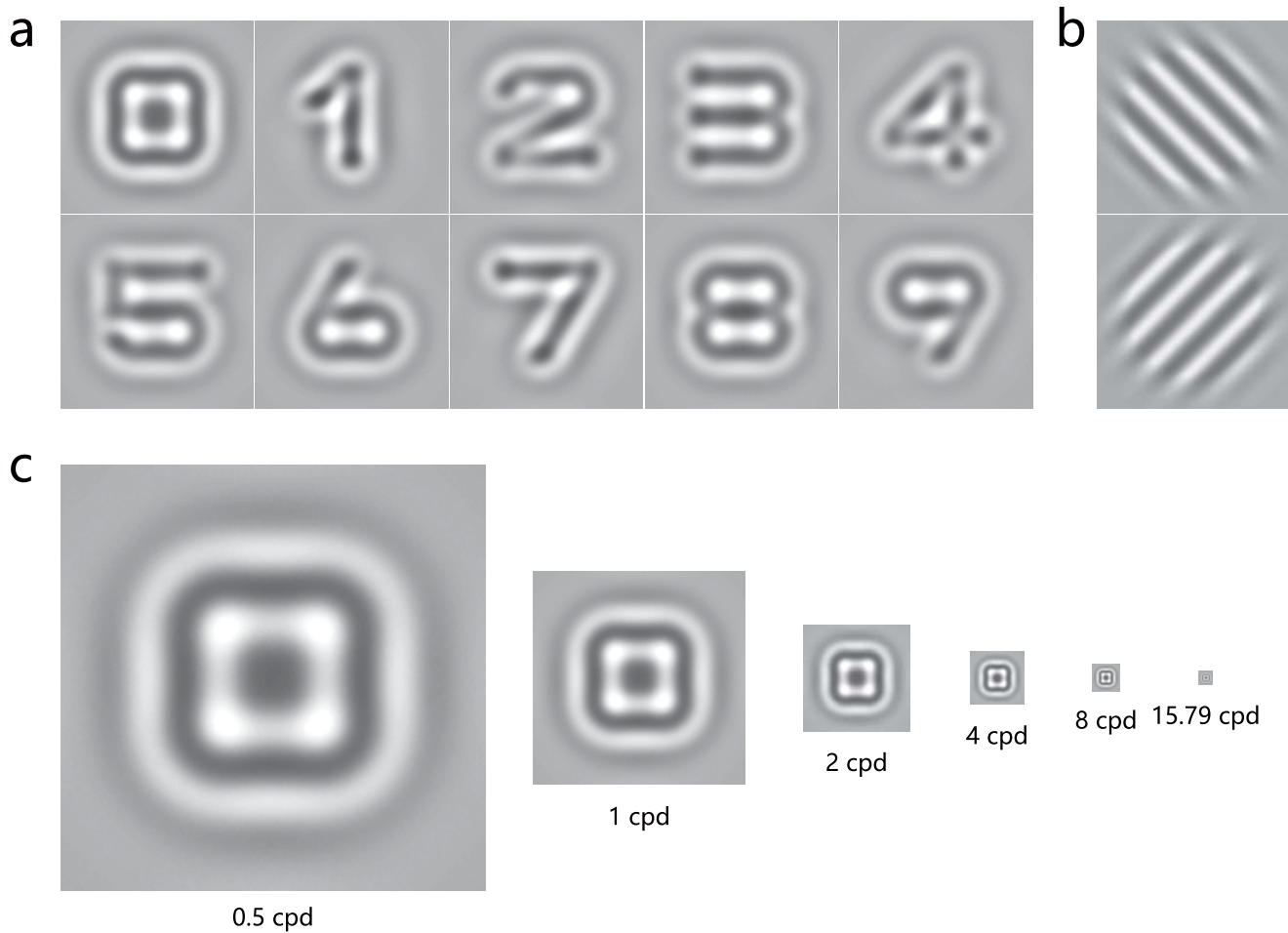


Figure 1. (a) Ten filtered digits. (b) Two Gaussian-windowed gratings. (c) Filtered digit “0” in different spatial frequency conditions.

method.⁴⁰ The Psi method was programmed to estimate the thresholds using Weibull psychometric functions with a fixed slope of 3 in the grating task and a fixed slope of 2.74 in the digit task based on data from pilot studies. Each experimental block lasted approximately 20 minutes and consisted of 300 trials (50 trials \times 6 spatial frequencies). Every block was divided into six mini-blocks. The observers could take a break after each mini-block. The spatial frequency of the test stimulus in each trial was selected randomly. Observers completed 10 grating blocks and 10 digit blocks over a total of five days. In each day, the order of the sessions was either “grating, digit, digit, grating” or “digit, grating, grating, digit”, counter-balanced over days for each observer.

In the beginning of the experiment, observers were prescribed with approximately 200 practice trials in each task to become familiar with the experimental setting and procedure. In the beginning of the first

block in each day, observers were given 5 minutes to adapt to the dim test environment.

Procedure

Each trial began with a 250-ms crosshair fixation in the center of the screen, and a brief tone signaling its onset, followed by a short 125-ms blank screen with mean luminance. The test stimulus then was presented for 133 ms. To facilitate response collection, a response screen was shown 500 ms after stimulus presentation. The response screen displayed 10 digits during the digit identification task and two thick oblique lines, oriented 45° clockwise or 45° counter-clockwise from vertical during the grating task. Digits and oblique lines in the response screen were arranged as a 2 \times 5 or 1 \times 2 matrix, and presented in the center of the display.

Observers were instructed to use the keyboard to type or mouse to select the stimulus they saw. Allowing different participants to use their more

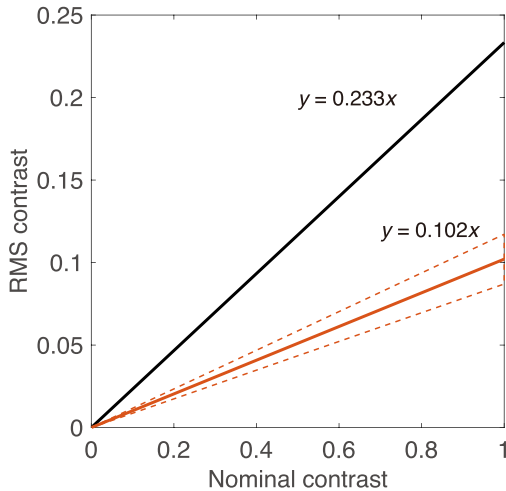


Figure 2. The relationship between nominal and RMS contrasts for grating and digit stimuli. *Black line:* grating; *Red lines:* digit; *Red dashed lines* indicate ± 1 SD of the slopes of 10 digits.

familiar input devices could reduce test variability. An auditory beep immediately followed each response, independent of response accuracy. A new trial started 500 ms after each response.

Results

Equating Stimulus Energy and d' Performance Level

As we mentioned previously, the contrast energy is a better contrast metric for stimuli with nonperiodic luminance distributions.^{32,41} When the stimulus area and duration are known, the square root of contrast energy can be simplified to the root mean square (RMS) contrast, which is defined as the standard deviation of the pixel contrast in an image. To equate the contrast energy of grating and digit stimuli, we converted the nominal contrast values of both types of stimulus into RMS contrast. We generated grating images and digit images at 100 different contrast levels and six spatial frequencies and calculated the RMS contrast for every image. By averaging the RMS contrast across orientations/digits and spatial frequencies, we obtained the RMS contrast as a function of nominal contrast for grating and digit stimuli, respectively. Figure 2 shows the relationship between the nominal and RMS contrasts for grating and digit stimuli, which can be described by straight lines with different slopes:

$$\text{RMS}_{\text{grating}} = 0.233 \times C_{\text{grating}} \quad (1a)$$

$$\text{RMS}_{\text{digit}} = 0.102 \times C_{\text{digit}} \quad (1b)$$

Second, contrast thresholds measured with the two

tasks may correspond to different d' performance levels, which represent the distance between the means of the signal and the noise distributions, divided by their common standard deviation.⁴² The contrast threshold measured by the 2AFC task with the Psi method corresponds to 81.6% correct,⁴³ or a d' of 1.27. On the other hand, the contrast threshold measured by the 10AFC task corresponds to 66.9% correct, or a d' of 1.99.²⁴ To equate the d' performance level of the two tasks, we converted the measured thresholds in the digit task at 66.9% correct ($d' = 1.99$) into thresholds at 43% correct ($d' = 1.27$) using a Weibull psychometric function:

$$\text{CSF}_{d'=1.27} = \text{CSF}_{d'=1.99} + \log_{10}(\log[(1 - \gamma)/(1 - 0.43)])/\beta, \quad (2)$$

where $\gamma = 0.1$, and $\beta = 2.74$. (We used a fixed slope of 2.74 for the Psi method in the experiment.) With Equations 1 and 2, we converted the estimated contrast thresholds in both tasks into RMS contrast thresholds at the same d' performance level of 1.27.

Agreement Between the CSFs From the Two Tasks

To evaluate the equivalency of the estimated CSFs, we compared the CSFs obtained from the digit and grating tests. The estimated contrast sensitivities at the six spatial frequencies from the Psi method over 10 blocks of the digit and grating tasks for each young observer are plotted in Figure 3. As we can see, after equating the contrast energy and d' performance levels, the CSFs obtained from the digit and grating tasks were very close for all observers.

The repeated measure analysis of variance (ANOVA) showed that, for the young observers, there was a significant effect of stimulus type, $F(1, 4) = 12.99$, $P = 0.023$, with no significant interaction between the stimulus type and spatial frequency, $F(5, 20) = 0.358$, $P = 0.871$, which suggests there was a constant difference between the CSFs measured with grating and digit. To quantify the amount of difference between the CSFs obtained from the two tasks, we calculated the RMS error (RMSE) between them. As the CSF using the Psi method was measured with six parallel adaptive procedures in six spatial frequencies, the RMSE was evaluated every six trials, one for each spatial frequency. Figure 4a shows the RMSE as a function of trial number in Experiment 1. The RMSE decreased drastically in the first 30 trials. The average RMSE across all five young observers was 0.131 ± 0.042 , 0.112 ± 0.038 , 0.095 ± 0.040 , and 0.093 ± 0.029 log₁₀ units after 30, 60, 120, and 300 trials, respectively. We also performed a correlation

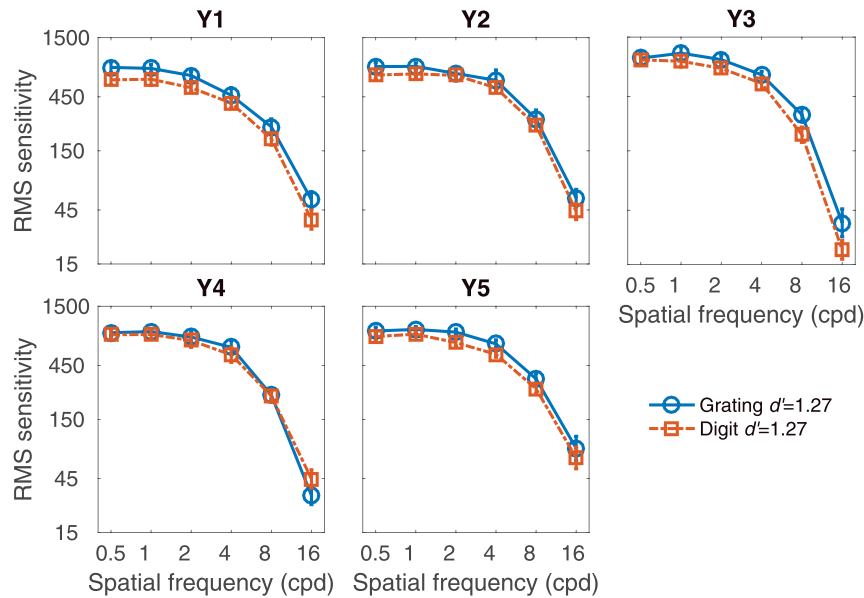


Figure 3. The CSFs from 10 blocks of digit and grating tasks for each young observer in Experiment 1. The blue and red curves represent CSFs obtained from the grating ($d' = 1.27$) and digit ($d' = 1.27$, after conversion) tasks, respectively. Error bars: ± 1 SD.

analysis on the estimated contrast sensitivities from the two tasks (Fig. 4b). The Pearson correlation coefficient was 0.991 ($P = 5.71 \times 10^{-26}$).

Moreover, we performed a Bland-Altman analysis⁴⁴⁻⁴⁶ to quantify the agreement between the estimated CSFs from the digit and grating tasks (Fig. 4c). The mean difference was -0.072 log₁₀ units. The limits of agreement were ± 0.128 log₁₀ units. Taken together, the results indicated that the estimated CSFs with the grating and digit stimuli were highly

matched with a small (approximately 0.1 log₁₀ units) constant difference across spatial frequencies.

Precision of the Two CSF Tests

The precision of a test can be gauged by the standard deviation of test results from repeated measures. We calculated the standard deviation of the CSFs of 10 blocks in every six trials, and averaged it across six spatial frequencies for each young observer. The mean standard deviations over five

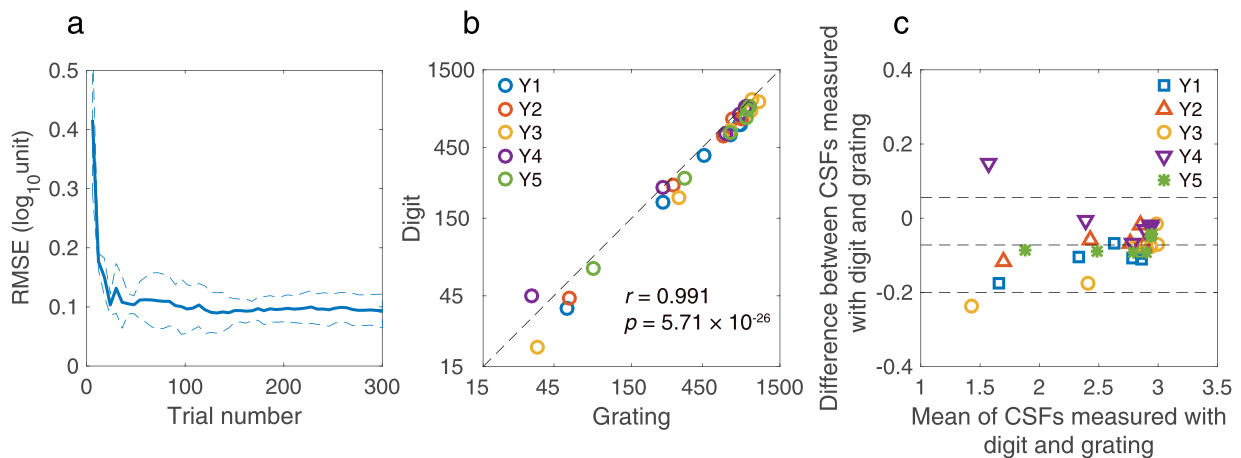


Figure 4. (a) The average RMSE between the CSFs obtained from the digit and grating tasks as a function of trial number in Experiment 1. The dashed lines represent ± 1 SD. (b) Scatter plot of the CSFs obtained from the two tasks for all young observers in Experiment 1. (c) The Bland-Altman plot of the estimated CSFs from the digit and grating tasks in Experiment 1: the difference between the CSFs from the two tasks against the mean of the CSFs from the two tasks. Different symbols represent different observers.

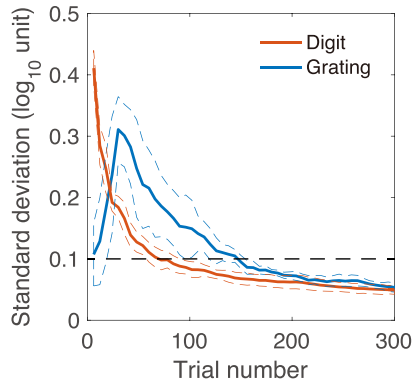


Figure 5. The average standard deviation of the estimated CSFs from the two tasks as functions of trial number in Experiment 1. The red and blue curves represent results from the digit and grating tests, respectively. The dashed lines represent ± 1 SD.

young participants for the two tests are plotted as functions of trial number in Figure 5. The average standard deviation of the CSFs obtained from the digit task decreased rapidly in the first 50 trials. After the first 24 trials, the standard deviation of the CSFs from the digit task was lower than that of grating CSF. The standard deviation of the CSFs was 0.184 ± 0.018 , 0.111 ± 0.016 , 0.078 ± 0.009 , and 0.049 ± 0.007 log₁₀ units after 30, 60, 120, and 300 trials, respectively, from the digit task and 0.311 ± 0.054 , 0.215 ± 0.062 , 0.124 ± 0.026 , and 0.054 ± 0.008 log₁₀ units after 30, 60, 120, and 300 trials, respectively, from the grating task. The *t*-tests have been applied to compare the standard deviations of the CSFs at each trial with 0.1 log₁₀ units for two tasks. The standard deviation of CSFs from the digit task decreased to 0.1 log₁₀ units after 60 trials ($t[4] = 1.508$, $P = 0.206$) and became significantly lower than 0.1 log₁₀ units after 96 trials ($t[4] = 2.867$, $P = 0.046$). In contrast, the standard deviation of CSFs from the grating task required at least 90 trials to reach 0.1 log₁₀ units ($t[4] = 2.556$, $P = 0.063$) and 162 trials to get below 0.1 log₁₀ units ($t[4] = 3.007$, $P = 0.040$). The digit test exhibited higher precision.

Test–Retest Reliability of the Digit CSF Test

To evaluate the test–retest reliability of the digit test with the Psi method in young group, we analyzed the overall concordance correlation coefficient (OCCC)^{47,48} for measuring agreement among 10 digit assessments. The OCCC is the weighted average of the pairwise concordance correlation coefficient (CCC) between any two assessments, which evaluates the agreement between two tests by computing the

weighted average of the pairwise CCC between any two assessments and has desirable characteristics. Concretely, the OCCC is estimated by the following equation⁴⁷:

$$\widehat{\rho}_O = \frac{2 \sum_{j=1}^{J-1} \sum_{k=j+1}^J S_{jk}}{(J-1) \sum_{j=1}^J S_j^2 + J \sum_{j=1}^J (Y_{\bullet j} - Y_{\bullet\bullet})^2} \quad (3)$$

where each of J observers assesses each of N subjects (a random sample from the population of interest) with a continuous scale Y . And Y_{ij} is the estimate from observer j for subject i ($i = 1, \dots, N$). $Y_{\bullet j}$, S_j , and S_{jk} represent as sample means, variances, and covariances, respectively. In Experiment 1, the mean OCCC of the estimated digit CSF across five young observers was 0.956 ± 0.157 .

Properties of the Psychometric Functions in the Digit CSF Test

In the procedure, we treated all 10 digits as equivalent stimuli and used a single psychometric function to model observer’s performance in each spatial frequency condition. Although the similarities structure among filtered digits were improved with our design,²⁵ the RMS contrasts of different digits with the same nominal contrast were different (as evidenced by the SD of RMS contrast function for digit stimuli in Fig. 2). This would undermine the assumption of the single psychometric function. To evaluate if the difference in RMS contrast impaired the fidelity of the estimated CSF, we examined the psychometric function for each digit at each spatial frequency by collecting a large amount of data in Experiment 1.

For each observer in Experiment 1, there were a total of 6 (spatial frequency) \times 10 (digit) \times 50 (trials per digit). This allowed us to construct 6 (spatial frequency) \times 10 (digit) psychometric functions for each observer and fit them with Weibull functions using a maximum likelihood procedure.⁴⁹

Two models were used in the fitting procedure. In Model I, the threshold and slope of the Weibull function are independent across all the digit and spatial frequency conditions. In Model II, we constrained the slope of the Weibull function to be the same across all the digit and spatial frequency conditions. A χ^2 test was used to test the goodness of fit of/between the nested models.⁴⁹ Both models showed significant good fit to the raw psychometric function for all observers (χ^2 test, all $P > 0.4$, Table 2). Model comparison showed that the two models were equivalent ($\chi^2[59]$, all $P > 0.10$, Table 2),

Table 2. Results of the Model Fit for All Observers

Observer	Y1	Y2	Y3	Y4	Y5
Model I					
$\chi^2(2880)$	2.75×10^3	2.87×10^3	2.59×10^3	2.88×10^3	2.62×10^3
P	0.963	0.550	1.00	0.498	1.00
Model II					
$\chi^2(2939)$	2.81×10^3	2.94×10^3	2.65×10^3	2.95×10^3	2.69×10^3
P	0.950	0.476	1.00	0.421	1.00
Model II vs. Model I					
$\chi^2(59)$	67.8	73.1	64.1	73.9	64.7
P	0.202	0.103	0.301	0.092	0.285

suggesting that the slope of the psychometric functions was the same across all the conditions. The best fitting model for Y1 is showed in Figure 6.

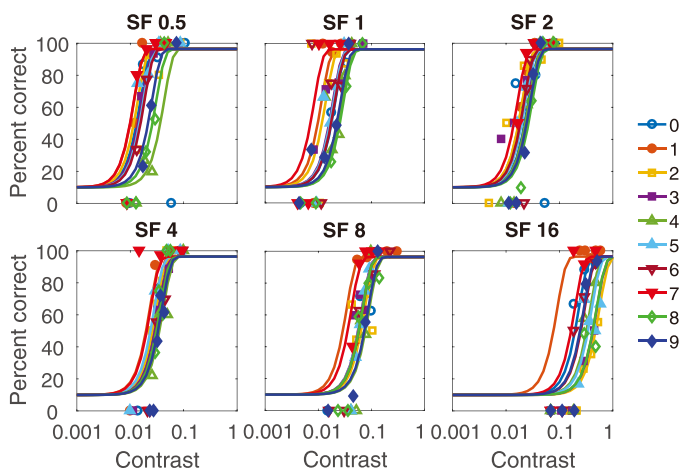
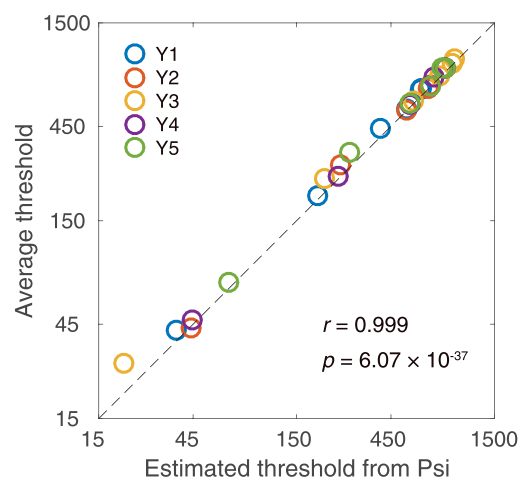
Based on the best fitting model for each observer, we calculated the threshold and slope of the psychometric function in each digit and spatial frequency condition. The average slope across observers was 2.93 ± 0.45 , which was not significantly different from the fixed slope of 2.74 used in the experiment ($t[4] = 0.923$, $P = 0.408$). In Figure 7, we plot the average threshold across all digit conditions against the estimated threshold from the Psi method for all the observers and all the spatial frequency conditions. The two measures matched extremely well, with a RMSE of only 0.038 log10 unit. The diagonal line accounted for 99.4% of the variance.

Confusion Matrix

To gain a more complete picture of observer behavior in the digit identification task, we combined the data from all the observers and computed a

confusion matrix (Fig. 8a). The i th column of the confusion matrix represents the frequency of reporting the j th ($j = 1, 2, \dots, 10$) digit when the i th digit was presented to the observer. The diagonal entries of the confusion matrix indicated the frequency of correct responses. Across all digits, 70.2% responses were correct, with digit 7 being the least confusable digit with a 92% correct response rate. The off-diagonal entries indicated “confusion.” Those with more than 10% response rate were highlighted in red. Digit 8 was the most confusing digit, and was confused with digit 3 and 5 quite often.

Given the complicated behavioral responses of the observer in the 10AFC digit identification task, one would ask whether the CSF measured with the complicated task can faithfully reflect the sensory limitation of our visual system. To answer the question, we constructed a simple observer model, in which:

**Figure 6.** The best fitting model (Model II) for Y1.**Figure 7.** The comparison between the average threshold across 10 digits and the estimated threshold from the Psi method.

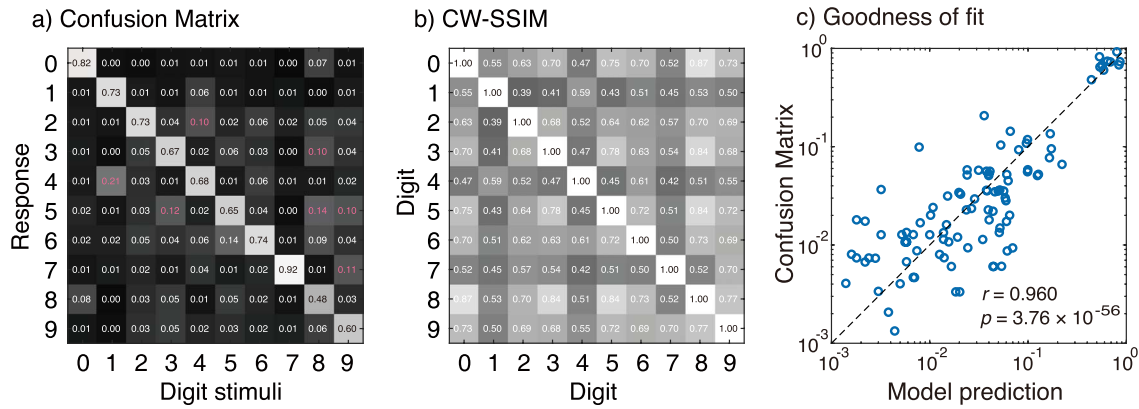


Figure 8. (a) The digit response confusion matrix. The i th column of the confusion matrix represents the frequency of reporting the j th digit when the i th digit was presented to the observer. (b) The complex wavelet structural similarity indexes (CW-SSIM) of the filtered digit stimuli. (c) Scatter plot of the entries of the observed confusion matrix against the CW-SSIM based model predictions.

1. The observer has full knowledge of all filtered digits when they are at the highest contrast;
2. The input of an i th digit input is represented in the j th digit template with an mean activation that is proportional to the complex wavelet structural similarity indexes (CWSSIM)⁵⁰ of the filtered digits²⁵ (Fig. 10b), between the input and the template.
3. Because we did not use external noise and have collapsed all the spatial frequency conditions in

this analysis, we made another simplifying assumption that the gain of the digit templates is 1, and the internal noise in each digit channel is a Gaussian random variable. The internal noise in all channels is independent and identically-distributed with the same standard deviation, Na .

4. The observer reports the digit associated with the template that has the maximum activation.

The model can be expressed with the following

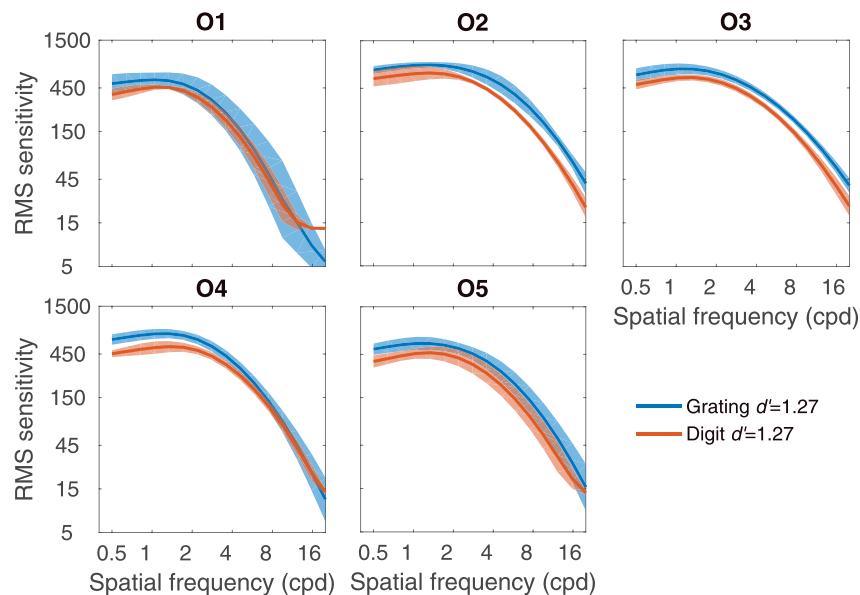


Figure 9. The averaged CSFs from the digit and grating tasks with the qCSF method in older group. The blue and red curves represent CSFs obtained from the grating ($d' = 1.27$) and digit ($d' = 1.27$, after conversion) tasks, respectively. The shaded regions represent ± 1 inter-run SD.

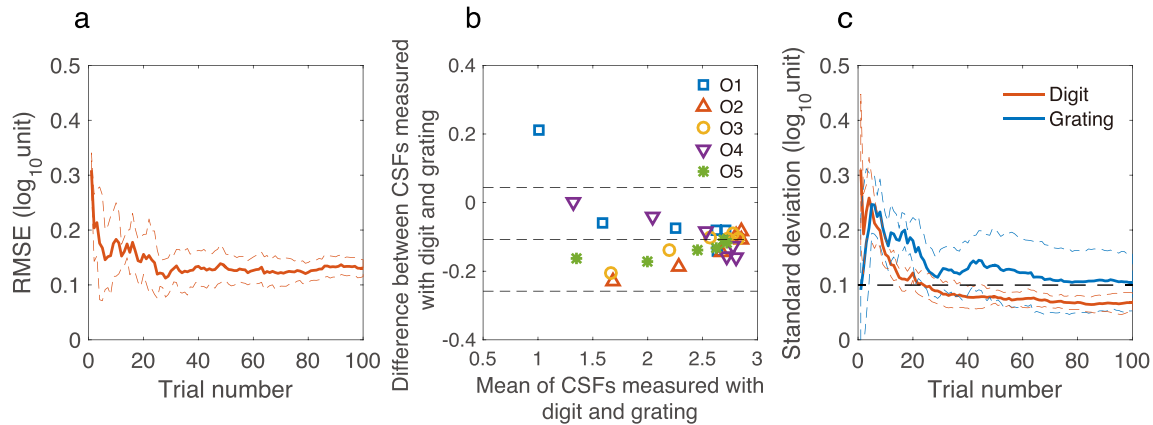


Figure 10. (a) The average RMSE between the CSFs obtained from the digit and grating tasks as a function of trial number for older observers in Experiment 2. The *dashed lines* represents ± 1 SD. (b) The Bland–Altman plot of the estimated contrast sensitivities from the digit and grating tasks in Experiment 2. Different symbols represent different observers. (c) The average standard deviation of the estimated CSFs from the two tasks as functions of trial number in Experiment 2. The *red* and *blue* curves represent results from the digit and grating tests, respectively. The dotted lines represent ± 1 SD.

mathematical expression:

$$R_i = \arg \max_j (\text{CWSSIM}_{i,j} + Na \cdot G[0, 1]), \quad (4)$$

where $G(0,1)$ is a Gaussian random variable with mean of zero and standard deviation of one. After considering the image similarity, the model with one additive noise parameter provided an excellent fit to the confusion matrix (the maximum likelihood procedure, χ^2 test, $\chi^2[99] = 3.42$, $P = 0.999$). The observed confusion matrix is plotted against the model predictions in Figure 8c. The correlation between them is 0.960 ($P < 0.001$). The results suggested the CSF measured with the 10-AFC digit identification task can faithfully reflect the sensory constraint.

Experiment 2

In Experiment 1, we showed that the CSF measured with the digit and grating stimuli matched well in young observers. Because the data from a diverse population is more convincing, we further compared the CSFs measured with digit and grating in a group of older participants.

Method

In Experiment 2, the quick CSF procedure (qCSF) used in our previous study²⁵ was used to save the testing time. Older participants who were not familiar with computers were asked to verbally report the identity of the digit stimulus, and their responses were

collected by the experimenter via the computer keyboard. The rest of the experiment setup, such as apparatus and stimuli, were exactly the same as those used in Experiment 1.

Observers

Five older participants (mean age, 65 ± 5.15 years; age range, 60–73 years) from the local communities in Wenzhou, Zhejiang, China, took part in Experiment 2. The first and second authors performed the ophthalmologic and optometric examinations for them. All observers were phakic and had normal or corrected-to-normal vision ($\log\text{MAR} \leq 0$). The observers showed no sign of eye disease other than cataract (O1–O5) and pterygium (O1, right eye). They also were free from diabetes, hypertension, and cognitive deficits (Mini-Mental State Examination [MMSE] > 26 points), and no history of mental diseases. All participants were naive to the purpose of the study. The study protocol adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board of human subject research of the Eye Hospital at Wenzhou Medical University. Written informed consent was obtained before the experiment for all observers.

Procedure and Design

Contrast sensitivities at 0.50, 1, 2, 4, 8, and 15.8 cpd were measured using the qCSF method with grating²⁴ and digit stimuli.²⁵ Each participant completed six qCSF runs with grating task and six qCSF runs with digit task over a total of three days. Each qCSF run consisted of 100 trials and lasted approx-

imately 10 minutes. Participants were given five minutes to adapt to the dim test environment before the test in each day. Some participants who usually did not wear spectacles were given an additional 25 minutes to get used to their prescribed optical correction.

Results

Agreement Between the CSFs From the Two Tasks

We compared the CSFs obtained from the grating and digit qCSF in the older group using same correction in Experiment 1. The CSFs from six qCSF runs with grating and digit stimuli are plotted in Figure 9. After equating the contrast energy and d' performance levels, the CSFs obtained from the digit task were close to those from the grating task for all older observers.

The result of the repeated measure ANOVA showed a similar pattern to that in the young group. For the older observers, there was a significant effect of stimulus type ($F[1, 4] = 32.2, P = 0.005$), with no significant interaction between stimulus type and spatial frequency ($F[5, 20] = 0.231, P = 0.945$). We also calculated the RMSE between the CSFs obtained from the two tasks to further quantify the agreement between them. From Figure 10a, we can see that the RMSE decreased rapidly in the first 10 trials. The average RMSE across all older observers was $0.159 \pm 0.041, 0.125 \pm 0.034, 0.126 \pm 0.024,$ and 0.131 ± 0.016 log₁₀ units after 15, 30, 60, and 100 trials, respectively. We also performed a Bland-Altman analysis⁴⁴⁻⁴⁶ to quantify the agreement between the estimated contrast sensitivities from the digit and grating tasks in Experiment 2 (Fig. 10b). The mean difference was -0.107 log₁₀ units. The limits of agreement were ± 0.151 log₁₀ units.

Precision of the Two CSF Tests

The similar pattern of the standard deviation as functions of trial number also was shown for older observers (Fig. 10c). The average standard deviation of the CSFs among six blocks obtained from the digit task decreased rapidly in the first 20 trials. After the first five trials, the standard deviation of the CSFs from the digit task was lower than that of grating CSF. The standard deviation of the CSFs was $0.114 \pm 0.016, 0.087 \pm 0.023, 0.074 \pm 0.015,$ and 0.068 ± 0.018 log₁₀ units after 15, 30, 60, and 100 trials, respectively, for the digit task and $0.181 \pm 0.022, 0.122 \pm 0.042, 0.122 \pm 0.061,$ and 0.105 ± 0.052 log₁₀ units after 15, 30, 60, and 100 trials, respectively, for the grating task. For the digit task, the

standard deviation of the CSFs reached 0.1 log₁₀ units after 15 trials ($t[4] = 1.867, P = 0.135$) and became significantly lower than 0.1 log₁₀ units after 47 trials ($t[4] = 3.225, P = 0.032$). For the grating task, the standard deviation of the CSFs reached 0.1 log₁₀ units in 21 trials ($t[4] = 2.314, P = 0.082$) but never got below 0.1 log₁₀ units in the entire run ($P_s > 0.05$ for trials after 21). Again, the digit test was more precise.

Test-Retest Reliability of the Digit CSF Test

To evaluate the test-retest reliability of the digit test with qCSF method for older observers, we analyzed the OCCC with Equation 3 for assessing agreement among six digit measurements. The mean OCCC of the estimated CSF across five older observers in Experiment 2 was 0.913 ± 0.042 .

Discussion

In this study, we compared the CSFs obtained from a digit identification task with those from a more traditional grating orientation identification task. We found that, after controlling for stimulus type and performance level differences, the CSFs obtained from the two tasks highly matched for young and older observers. However, if we defined the efficiency as the standard deviation change per unit time or trial,²⁴ the digit CSF test was more efficient than the grating CSF test. For young observers, with 120 trials, the standard deviation of the estimated contrast sensitivities was 0.078 ± 0.009 and 0.124 ± 0.026 log₁₀ units for the digit and grating tests, respectively. For older observers, after 60 trials the standard deviation of the CSFs was 0.074 ± 0.015 and 0.122 ± 0.061 log₁₀ units for the digit and grating tests, respectively. To reach the same standard deviation, the digit CSF test required fewer trials/less time than the grating CSF. In addition, we found, through a detailed analysis of the psychometric functions for individual digits, that the average thresholds over all digits matched the estimated thresholds in the experiment with a method that assumed a single psychometric function for all the digits and the complicated behavioral responses of the observer in the 10AFC digit identification task can faithfully reflect the sensory constraint of the visual system. The results suggested that the digit CSF test provides a compatible assessment of the CSF as the traditional grating CSF test with more efficiency.

From a first principle point of view, the CSFs should be independent of the task used in measurement, because the contrast threshold is determined by

the sensory constraint of the same visual system. Previously, we have shown that the CSF could be decomposed into several key system limiting factors: internal additive noise, template gain, multiplicative noise, and transducer non-linearity.^{28,51} The only difference between the grating and digit CSF tests are the input stimuli and response structure; the system parameters do not change. That is why it is not surprising to find the equivalency of the two tests after we matched the stimulus energy and performance level. Because we have only collected data in the zero external noise condition in this study, we cannot fully determine the system parameters.^{52–54} It would be interesting to compare these system parameters between grating and digit tasks in future studies with data in a range of external noise conditions.

One potential complication with the use of digit stimuli is that they are not all equivalent and nonorthogonal. This may introduce response bias in multialternative forced choice tasks.^{55,56} The digit stimuli used in the current study were developed in a parallel study and had an improved similarity structure compared to those used in the chart reported by Khambhiphanta et al.²⁵ Bandpass filtering^{26,28} further reduced the standard deviation of the pairwise similarities (CW-SSIM) between digits to 0.126. In this study, we analyzed the confusion matrix by aggregating the data from all observers, and found that it can be well predicted by the CW-SSIM matrix of the stimuli. The result suggested that we can incorporate CW-SSIM into future development of the CSF test to take into account the nonorthogonal nature of the stimuli.

Consistent with the results in the literature for Sloan letters,³⁵ we found that the slopes of the psychometric functions were the same across the 10 digits and all the spatial frequencies. The finding is consistent with our previous observation that the slope of the psychometric function is constant across different spatial frequency and external noise conditions.^{28,43,51} The observed slope invariance supports the “homogeneity assumption” of slope for all pattern-detecting mechanisms.⁵⁷ It also is an important regularity that we can exploit to model human performance in multiple conditions. For example, the qCSF method makes use of the fixed slope across spatial frequencies to gain testing efficiency.^{15,24,43}

CSF tests with grating stimuli traditionally have been performed either with stimuli of a fixed size (and, therefore, different numbers of sinewave cycles),^{34,58} or Gabors with a fixed number of cycles

(and therefore different sizes).⁵⁸ Pelli and Bex⁵ suggest that the use of a fixed number of sinewave cycles is better because neurons in the visual cortex are roughly spatial scale-invariant, and fixing stimulus size would introduce additional processes, such as spatial summation as spatial frequency increase.⁵⁹ The grating stimuli with a fixed number of sinewave cycles was used in the current experiment. In digit or letter CSF tests, the stimuli are scaled to generate tests at different spatial frequencies so that the size of grating stimuli was identical to that of the digit stimuli at same spatial frequency.

In summary, our results suggested that the digit CSF test provides a compatible assessment of the CSF as the traditional grating CSF test with more efficiency. It can be incorporated into vision assessment instruments to test people unfamiliar with the Latin alphabet.

Acknowledgments

Supported by the National Key R&D Program of China (2016YFB0401203; FH), the National Natural Science Foundation of China (NSFC81600764; FH), and Wenzhou Medical University (QTJ16006; FH), and by the National Eye Institute (EY021553 and REY025658; ZLL).

Disclosure: **H. Zheng**, None; **M. Shen**, None; **X. He**, None; **R. Cui**, None; **L.A. Lesmes**, qCSF technology (P), Adaptive Sensory Technology, Inc. (I), AST (E); **Z.-L. Lu**, qCSF technology (P), Adaptive Sensory Technology, Inc. (I); **F. Hou**, None

References

1. Ginsburg AP. Contrast sensitivity and functional vision. *Int Ophthalmol Clin* 2003;43:5–15.
2. Frisby JP, Mayhew JE. Contrast sensitivity function for stereopsis. *Perception*. 1978;7:423–429.
3. Bodis-Wollner I. Detection of visual defects using the contrast sensitivity function. *Int Ophthalmol Clin*. 1980;20:135–153.
4. Arden GB. The importance of measuring contrast sensitivity in cases of visual disturbance. *Br J Ophthalmol*. 1978;62:198–209.
5. Pelli DG, Bex P. Measuring contrast sensitivity. *Vision Res*. 2013;90:10–14.

6. Campbell FW, Robson JG. Application of Fourier analysis to the visibility of gratings. *J Physiol.* 1968;197:551–566.
7. Blakemore C, Campbell FW. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J Physiol.* 1969;203:237–260.
8. Legge GE, Foley JM. Contrast masking in human vision. *J Opt Soc Am.* 1980;70:1458–1471.
9. Lundh BL, Lennerstrand G, Derefeldt G. Central and peripheral normal contrast sensitivity for static and dynamic sinusoidal gratings. *Acta Ophthalmol (Copenh).* 1983;61:171–182.
10. Mullen KT. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *J Physiol.* 1985;359:381–400.
11. Hou F, Huang CB, Tao L, Feng L, Zhou Y, Lu ZL. Training in contrast detection improves motion perception of sinewave gratings in amblyopia. *Invest Ophthalmol Vis Sci.* 2011;52:6501–6510.
12. Ginsburg AP. Next generation contrast sensitivity testing. In: Rosenthal BCR, ed. *Functional Assessment of Low Vision.* St Louis: Mosby Year Book Inc; 1996:77–88.
13. Pomerance GN, Evans DW. Test-retest reliability of the CSV-1000 contrast test and its relationship to glaucoma therapy. *Invest Ophthalmol Vis Sci.* 1994;35:3357–3361.
14. Ginsburg AP. A new contrast sensitivity vision test chart. *Am J Optom Physiol Opt.* 1984;61:403–407.
15. Lesmes LA, Lu ZL, Baek J, Albright TD. Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J Vis.* 2010;10:1711–21.
16. Alcaláquintana R, Garcíaapérez MA. The role of parametric assumptions in adaptive Bayesian estimation. *Psychol Methods.* 2004;9:250.
17. Patterson VH, Foster DH, Heron JR. Variability of visual threshold in multiple sclerosis: effect of background luminance on frequency of seeing. *Brain.* 1980;103:139–147.
18. Taylor MM. On the efficiency of psychophysical measurement. *J Acoust Soc Am.* 1971;49(suppl 2): 505.
19. Leek MR, Hanna TE, Marshall L. Estimation of psychometric functions from adaptive tracking procedures. *Percept Psychophys.* 1992;51:247–256.
20. Jakel F, Wichmann FA. Spatial four-alternative forced-choice method is the preferred psychophysical method for naive observers. *J Vis.* 2006; 6:1307–1322.
21. Shelton BR, Scarrow I. Two-alternative versus three-alternative procedures for threshold estimation. *Percept Psychophys.* 1984;35:385–392.
22. Hall JL. A procedure for detecting variability of psychophysical thresholds. *J Acoust Soc Am.* 1983;73:663–667.
23. Pelli DG, Robson J, Wilkins A. The design of a new letter chart for measuring contrast sensitivity. *Clin Vision Sci.* 1988;2:187–199.
24. Hou F, Lesmes L, Bex P, Dorr M, Lu ZL. Using 10AFC to further improve the efficiency of the quick CSF method. *J Vis.* 2015;15.
25. Zheng H, Wang C, Cui R, et al. Measuring the contrast sensitivity function using the qCSF method with 10 Digits. *Transl Vis Sci Technol.* 2018;7:9.
26. Chung ST, Legge GE, Tjan BS. Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vision Res.* 2002;42:2137–2152.
27. Hou F, Lesmes LA, Kim W, et al. Evaluating the performance of the quick CSF method in detecting contrast sensitivity function changes. *J Vis.* 2016;16:18.
28. Hou F, Lu ZL, Huang CB. The external noise normalized gain profile of spatial vision. *J Vis.* 2014;14:9.
29. Murray IJ, Parry NR, Ritchie SI, et al. Importance of grating orientation when monitoring contrast sensitivity before and after refractive surgery. *J Cataract Refract Surg.* 2008;34:551–556.
30. Logi F, Pellegrinetti A, Bonfiglio L, et al. Effects of grating spatial orientation on visual evoked potentials and contrast sensitivity in multiple sclerosis. *Acta Neurol Scand.* 2001;103: 97–104.
31. Watson AB. Visual detection of spatial contrast patterns: evaluation of five simple models. *Opt Exp.* 2000;6:12–33.
32. Kukkonen H, Rovamo J, Tiippana K, Nasanen R. Michelson contrast, RMS contrast and energy of various spatial stimuli at threshold. *Vision Res.* 1993;33:1431–1436.
33. Ahumada AJ, Watson AB. A spatial standard observer based on contrast energy. *Invest Ophthalmol Vis Sci.* 2004;45:U492–U492.
34. Kelly DH. Motion and vision. II. Stabilized spatio-temporal threshold surface. *J Opt Soc Am.* 1979;69:1340–1349.
35. Alexander KR, Xie WEI, Derlacki DJ. Visual Acuity and Contrast Sensitivity for Individual Sloan Letters. *Vision Res.* 1997;37:813–819.

36. Khambhaphanta B, Tulvatanab W, Busayaratb M. The new numbers contrast sensitivity chart for contrast sensitivity measurement. *J Optom.* 2011; 4:128–133.
37. Kleiner M, Brainard D, Pelli D. What's new in Psychtoolbox-3? *Perception.* 2007;36:14–14.
38. Tyler CW. Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spat Vis.* 1997;10:369–377.
39. Committee on Vision NRC, National Academy of Sciences, Washington, D.C. Recommended standard procedures for the clinical measurement and specification of visual acuity. *Adv Ophthalmol.* 1980;41:103–148.
40. Kontsevich LL, Tyler CW. Bayesian adaptive estimation of psychometric slope and threshold. *Vision Res.* 1999;39:2729–2737.
41. Moulden B, Kingdom F, Gatley LF. The standard deviation of luminance as a metric for contrast in random-dot images. *Perception.* 1990; 19:79–101.
42. Kingdom FAA, Prins N. *Psychophysics: a practical introduction.* 2nd ed. London, UK: Academic Press; 2016.
43. Hou F, Huang CB, Lesmes L, et al. qCSF in clinical application: efficient characterization and classification of contrast sensitivity functions in amblyopia. *Invest Ophthalmol Vis Sci.* 2010;51: 5365–5377.
44. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307–310.
45. Bland JM, Altman DG. Measuring agreement in method comparison studies. *STAT Method Medical Res.* 1999;8:135–160.
46. Giavarina D. Understanding Bland Altman analysis. *Biochem Med.* 2015;25:141–151.
47. Barnhart HX, Haber M, Song J. Overall concordance correlation coefficient for evaluating agreement among multiple observers. *Biometrics.* 2002; 58:1020–1027.
48. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989;45:255–268.
49. Watson AB. Probability summation over time. *Vision Res.* 1979;19:515–522.
50. Sampat MP, Wang Z, Gupta S, Bovik AC, Markey MK. Complex wavelet structural similarity: a new image similarity index. *IEEE T Image Process.* 2009;18:2385–2401.
51. Chen G, Hou F, Yan FF, et al. Noise provides new insights on contrast sensitivity function. *PLoS One.* 2014;9:e90579.
52. Lu ZL, Doshier BA. Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychol Rev.* 2008;115:44–82.
53. Lu ZL, Doshier BA. External noise distinguishes attention mechanisms. *Vision Res.* 1998;38:1183–1198.
54. Lu ZL, Doshier BA. *Visual Psychophysics: From Laboratory to Theory.* Cambridge, MA: The MIT Press; 2014.
55. Hacker MJ, Ratcliff R. A revised table of d' for M-alternative forced choice. *Percept Psychophys.* 1979;26:168–170.
56. Klein SA. Measuring, estimating, and understanding the psychometric function: a commentary. *Percept Psychophys.* 2001;63:1421–1455.
57. Nachmias J. On the psychometric function for contrast detection. *Vision Res.* 1981;21:215–223.
58. Carney T, Klein S, Tyler C, et al. *Development of an image/threshold database for designing and testing human vision models.* Taiwan: National Taiwan University; 1999. Thesis.
59. Jamar JH, Koenderink JJ. Sine-wave gratings: scale invariance and spatial integration at supra-threshold contrast. *Vision Res.* 1983;23:805–810.