qCSF in Clinical Application: Efficient Characterization and Classification of Contrast Sensitivity Functions in Amblyopia

Fang Hou,1 Chang-Bing Huang,2 Luis Lesmes,3 Li-Xia Feng,4 Liming Tao,4 Yi-Feng Zhou,*,1 and Zhong-Lin Lu*,2

PURPOSE. The qCSF method is a novel procedure for rapid measurement of spatial contrast sensitivity functions (CSFs). It combines Bayesian adaptive inference with a trial-to-trial information gain strategy, to directly estimate four parameters defining the observer’s CSF. In the present study, the suitability of the qCSF method for clinical application was examined.

METHODS. The qCSF method was applied to rapidly assess spatial CSFs in normal 10 and 8 amblyopic participants. The qCSF was evaluated for accuracy, precision, test–retest reliability, suitability of CSF model assumptions, and accuracy of amblyopia screening.

RESULTS. qCSF estimates obtained with as few as 50 trials matched those obtained with 300 trials. The precision of qCSF estimates obtained with 120 and 130 trials, in normal subjects and amblyopes, matched the precision of 300 trials. For both groups and both methods, test–retest sensitivity estimates were well matched (all R > 0.94). The qCSF model assumptions were valid for 8 of 10 normal participants and all amblyopic participants. Measures of the area under log CSF (AU LCSF) and the cutoff spatial frequency (cutSF) were lower in the amblyopia group; these differences were captured within 50 qCSF trials. Amblyopia was detected at an approximately 80% correct rate in 50 trials, when a logistic regression model was used with AU LCSF and cutSF as predictors.

CONCLUSIONS. The qCSF method is sufficiently rapid, accurate, and precise in measuring CSFs in normal and amblyopic persons. It has great potential for clinical practice. (Invest Ophthalmol Vis Sci. 2010;51:5365–5377) DOI:10.1167/iovs.10-5468

Amblyopia is a developmental visual disorder characterized by poor spatial vision without detectable structural or pathologic abnormalities.1,2 It affects 2% to 4% of the general population3 and is the most common cause of uniocular blindness in adults.4 In the United States, 750,000 preschoolers are at risk for amblyopia, and roughly half of those may not be identified before school age.5 In China, it is believed that there are tens of millions of children affected by the visual disorder.6

Amblyopia has become a serious social and economic challenge.6

Amblyopia impairs visual functions, which include but are not limited to visual letter acuity, grating acuity, contrast sensitivity,7–8 contour integration,9–11 and global motion perception.12,13 Of these functions, it is letter acuity, measured rapidly with a chart and easy to understand, that is widely used in screening, early intervention, and treatment evaluation of amblyopia. However, it has been suggested that the contrast sensitivity function (CSF), which describes how visual sensitivity varies as a function of grating spatial frequency, characterizes amblyopia’s spatial vision deficits better than does letter acuity.14–25 For example, contrast sensitivity at high spatial frequencies is still abnormal in amblyopes who are deemed “treated” based on the criterion of remediated visual acuity.26

CSF testing could improve screening, diagnosis, and evaluation of treatment outcomes of amblyopia.

Current clinical CSF tests typically use preprinted letter/grating charts such as the Vistech contrast sensitivity chart (Vistech, Hartford, CT), the Functional Acuity Contrast Test (FACT), the Pelli-Robson chart,27–28 and the CVS-1000 series chart (VectorVision, Greenville, OH). Although chart tests are convenient in clinical use, their limited number of contrast levels greatly limits the range and resolution of the test grating stimuli.29–33 In addition, charts using broadband letters as stimuli (e.g., the Pelli-Robson chart), are typically insensitive to frequency-specific deficits.34 In contrast, the CSF can be measured with higher precision and accuracy in laboratory psychophysical tests. However, even with adaptive methods such as the QUEST35 or ψ method,36 a reasonable threshold estimate at a single spatial frequency usually takes 50 to 100 trials. Consequently, a CSF measurement obtained with conventional adaptive methods, sampling at seven different spatial frequencies, requires up to 700 trials, can take 1 hour, and is therefore too time-consuming, making it unsuitable for clinical use.

Lesmes et al.37 recently developed a novel adaptive psychophysical procedure, the qCSF method, which estimates the full CSF rapidly (~5–10 minutes) with reasonable precision (2–3 dB). (In this paper, 1 dB = 0.05 log units, 12%; but in the clinical definition of decibels [1 dB = 0.1 log units], the precision and bias were half of the corresponding estimates.)

The method applies a Bayesian adaptive test strategy, first developed to estimate the threshold and slope of psychometric functions60–68 to directly estimate CSF parameters. The CSF

From the 1Vision Research Laboratory, School of Life Science, University of Science and Technology of China, Hefei, Anhui, China; the 2Laboratory of Brain Processes, Department of Psychology, University of Southern California, Los Angeles, California; the 3Vision Center Laboratory, Salk Institute for Biological Studies, La Jolla, California; and the 4Department of Ophthalmology, the First Affiliated Hospital of Anhui Medical University, Hefei, Anhui, China.

Supported by Grant 2009CB941303 from the National Basic Research Program of China, Grant 306630027 from the Natural Science Foundation of China, and Grant EY017941 from the National Eye Institute.

Submitted for publication March 3, 2010; accepted April 15, 2010.

Disclosure: F. Hou, None; C.-B. Huang, None; L. Lesmes, P; L.-X. Feng, None; L. Tao, None; Y.-F. Zhou, None; Z.-L. Lu, P

*Each of the following is a corresponding author: Yi-Feng Zhou, Vision Research Laboratory, School of Life Science, University of Science and Technology of China, Hefei, Anhui 230027, China; zhouy@ustc.edu.cn.
Zhong-Lin Lu, Laboratory of Brain Processes, Department of Psychology, SGM 501, University of Southern California, Los Angeles, CA 90089; zhl@usc.edu.
form assumed by the qCSF is the truncated log-parabola, which is specified by four parameters: (1) the peak gain (sensitivity), \( \gamma_{\text{max}} \); (2) the peak spatial frequency \( f_{\text{max}} \); (3) the bandwidth \( \beta \), which describes the function’s full-width at half-maximum (in octaves); and (4) \( \delta \), the low-frequency truncation level (Fig. 1). Lesmes et al. validated the qCSF method with a psychophysical experiment, using an orientation identification task in three normal subjects. Combined with simulation results, they demonstrated that 50 to 100 qCSF trials provided accurate CSF estimates.

The testing time savings provided by the qCSF comes from (1) assuming a CSF model and (2) using an efficient Bayesian test strategy to estimate the model’s free parameters. Briefly, the qCSF model assumes that the observer’s CSF can be described by using a functional form, the truncated log-parabola, with four free parameters. The model also assumes that other parameters underlying psychophysical performance, such as the slope of the psychometric function or the lapse rate, are fixed. After defining a probability distribution over the four CSF parameters, a one-step-ahead search of the stimulus space (over dimensions of spatial frequency and contrast) is used to choose stimuli that maximize the information gained over the CSF parameter space. After the Bayesian update of the probability distribution following each trial, a CSF estimate can be calculated from the expected values of the four parameters.

Despite the great advantages of the qCSF, a model-dependent test strategy poses risks. Model mismatch, between the fixed parameters of the assumed qCSF model and the empirical model presented by the observer, can introduce systematic bias into estimates of the parameters. This risk is especially salient for clinical applications, in which visual disease may introduce model mismatch, by affecting all CSF model parameters, fixed or free. In this study, we evaluated the qCSF’s potential as a clinical CSF assessment procedure by applying it to the measurement of contrast sensitivity functions in normal and amblyopic populations. For independent CSF assessment in amblyopic and normal subjects, we interleaved qCSF assessment with an independent adaptive procedure, the \( \Psi \) method. Such an independent assessment was critical for evaluating the relative accuracy of qCSF estimates, in addition to the accuracy of several model assumptions: (1) Individual CSFs can be well described with a specific functional form; and (2) the slopes of psychometric functions at different spatial frequencies are the same. These assumptions, which allow CSF task performance to be modeled as a relatively simple psychometric surface (Fig. 1), have support in the empiric literature on CSFs in the normal population. However, they must be re-evaluated in clinical populations.

We report that the qCSF method required only 100 trials (5–10 minutes) to provide CSF estimates with a precision (1 SD) of 2.6 dB for normal subjects and 3.2 dB for amblyopes. Moreover, CSF features estimated by 50 qCSF trials distinguished normal subjects and amblyopes with an accuracy of approximately 80%. Furthermore, the assumptions used in the qCSF were valid in 8 of 10 normal participants and in all amblyopic participants. Taken together, the findings in our study indicate that the qCSF method provides efficient CSF measurements in both normal and amblyopic populations and competent performance for early screening and treatment evaluation in amblyopia clinical practice.

**Methods**

**Subjects**

Ten normal subjects (N1–N10) aged from 21 to 24 years (mean, 21.6 ± 1.6 years) and eight amblyopes (A1–A8) aged from 20 to 31 years (mean, 23.3 ± 5.7 years) participated in the experiments. All subjects were naïve to the purpose of the experiment. All normal subjects had normal or corrected-to-normal vision and no prior experience in psychophysical experiments. The amblyopes were diagnosed by two ophthalmologists (LF and LT). Six of them had participated in other psychophysical experiments. The detailed characteristics of the amblyopia group are listed in Table 1. The research adhered to the tenets of...
the Declaration of Helsinki. Written informed consent was obtained from the subjects after explanation of the nature and possible consequences of the study. The experiments were conducted according to the experimental protocol for human subjects approved by the ethics committee (IRB) of the School of Life Science, University of Science and Technology of China.

**Apparatus**

The stimuli were displayed on a computer monitor, driven by a video card (model G220; monitor; Sony, Tokyo, Japan, and 9250 ATI Radeon video card; AMD, Sunnyvale, CA). The monitor had a display area of 32.8 × 24.4 cm, a resolution of 1024 × 768, and a refresh rate of 100 Hz. The mean luminance of the display was 34 cd/m². A special circuit was used to achieve 14-bit grayscale resolution. Participants viewed the stimuli monocularly at a distance of 2.4 m. A chin-forehead-rest was used to minimize head movements during the experiment.

**Stimuli**

The stimuli were static vertical sinusoidal gratings with a size of 3.0° × 3.0°. A spatial envelope (0.5° half-Gaussian) was used to blend the gratings’ edge into the background, resulting in a 2.0° × 2.0° patch with full contrast.

**qCSF Implementation**

As mentioned earlier, in the qCSF procedure, the CSF is described by a truncated log-parabola with four parameters: (1) the peak gain (sensitivity), \( \gamma_{\text{max}} \); (2) the peak spatial frequency \( f_{\text{max}} \); (3) the bandwidth \( \beta \), which describes the function’s full-width at half-maximum (in octaves), and (4) \( \delta \), the low-frequency truncation level. The possible CSFs considered for Bayesian inference were represented by a probability distribution over a gridded parameter space, \( T_{\delta} \), where \( \tilde{c} = (\gamma_{\text{max}}, f_{\text{max}}, \beta, \delta) \). The stimulus space, \( V_{\text{stim}} \), where \( \tilde{x} = (c, f) \), represents all possible grating stimuli with contrast \( c \) and spatial frequency \( f \). The goal of the qCSF method is to efficiently search the stimulus space to gain information in the parameter space. See Appendix A for a more detailed description of the qCSF algorithm.

**Parameter Settings**

For the qCSF method, the stimulus space \( \tilde{x} = (c, f) \) comprises gratings that vary in contrast (from 0.1% to 99%) in 1.5-dB steps and spatial frequency (from 0.5 to 16 cycles per degree [cpd] in 3-dB steps). One exception was observer A1, who was tested at spatial frequencies that varied from 0.5 to 12 cpd in 3-dB steps, based on practice results. For the \( \Psi \) method, stimulus contrast \( c \) was sampled from 0.1% to 99% in 1-dB steps; contrast thresholds were measured independently at 0.69, 1.29, 2.42, 4.54, 8.52, and 16.00 cpd for all subjects except A1. For subject A1, thresholds were measured at 0.67, 1.19, 2.12, 3.78, 6.73, and 12.00 cpd.

A log-Weibull psychometric function was adopted for both the qCSF and \( \Psi \) procedures

\[
\Psi(c) = \gamma + (1 - \gamma - \lambda/2)[1 - \exp(-10^{\log_{10}(c) - \log_{10}(\gamma)})] \tag{1}
\]

where \( \tau \) is the to-be-measured contrast threshold, and the other parameters were fixed: Chance performance rate \( \gamma \) was set to 0.5, the lapse rate \( \lambda \) was set to 0.04, and the psychometric slope \( \eta \) was set to 3.5. (The value of 3.5 was based on its use in the QUEST method.) The Discussion section presents simulation results that demonstrate that matching the assumed slope and the observer’s slope is not critical for threshold estimation.) For the qCSF, the threshold variable, \( \tau \), was defined by the CSF model. For the \( \Psi \) method, \( \tau \) was estimated independently at each spatial frequency.

**Procedure**

A two-interval, forced-choice (2IFC) grating detection task was used. In each trial, a grating was randomly presented in one of two successive intervals. The intervals were separated by 500 ms, with a brief tone signaling each interval’s onset. In both intervals, a 120-ms presentation of the grating stimulus or a blank screen was preceded by a cross-hair frame (250 ms) that aided stimulus localization and fixation. Participants were asked to indicate which interval contained the grating. A new trial started 500 ms after the subject’s response. No feedback was provided.

**Design**

One hundred practice trials were used to familiarize each participant with the task. Both the qCSF and \( \Psi \) methods were used to estimate CSFs from the dominant eye of normal participants and the amblyopic eye of each amblyopic participant (or the worse eye in subjects with binocular amblyopia). For CSF measures obtained with the \( \Psi \) method, contrast thresholds were estimated independently at six spatial frequencies with 50 trials for each spatial frequency and a total of 300 trials. A qCSF run encompassed 300 trials. To compare the results from the qCSF and \( \Psi \) methods, contrast thresholds obtained with the qCSF method were calculated at the six spatial frequencies as follows.

---

**Table 1. Characteristics of Each Amblyopic Participant**

<table>
<thead>
<tr>
<th>No.</th>
<th>Sex</th>
<th>Age (y)</th>
<th>Eye</th>
<th>Type of Amblyopia</th>
<th>Eye Alignment</th>
<th>Correction</th>
<th>Acuity</th>
<th>Test-Retest Interval (d)</th>
<th>Experienced</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>20</td>
<td>AE</td>
<td>Anisometropia</td>
<td>None</td>
<td>1.50DC×90</td>
<td>2.5</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>25</td>
<td>AE</td>
<td>Strabismus/</td>
<td>RXT 10Δ</td>
<td>0.75DS</td>
<td>1.0</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>anisometropia</td>
<td></td>
<td>-1.50DS/0.75DC×180°</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>20</td>
<td>AE</td>
<td>Anisometropia</td>
<td>None</td>
<td>-5.00DS</td>
<td>10.0</td>
<td>12</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>22</td>
<td>AE</td>
<td>Anisometropia</td>
<td>None</td>
<td>+3.00DS;+1.00DC×15°</td>
<td>2.0</td>
<td>15</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>31</td>
<td>AE/LE</td>
<td>Measured</td>
<td>None</td>
<td>+2.50DS;+0.50DC×160°</td>
<td>1.5</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>22</td>
<td>AE/RE</td>
<td>Measured</td>
<td>None</td>
<td>+4.00DS</td>
<td>2.0</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>21</td>
<td>AE</td>
<td>Strabismus/</td>
<td>LET 10Δ</td>
<td>+6.00DS</td>
<td>4.0</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>anisometropia</td>
<td></td>
<td>+3.00DS</td>
<td>2.0</td>
<td>15</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>25</td>
<td>AE</td>
<td>Strabismus/</td>
<td>LET 18Δ</td>
<td>+1.25DS;+0.50DC×180°</td>
<td>1.3</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>anisometropia</td>
<td></td>
<td>+1.00DS</td>
<td>5.0</td>
<td>12</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.50DS;+1.00DC×80°</td>
<td>5.0</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.50DS</td>
<td>1.3</td>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

Acuity is expressed in minimum angle of resolution (MAR). AE: amblyopic eye; FE: fellow eye; LE: left eye, RE: right eye.
All participants completed two sessions of 600 trials, with 300 qCSF trials randomly intermixed with 300 H9023 trials. Because of scheduling conflicts, the interval between the two sessions ranged from 1 to 15 days across subjects (see Table 1 for details).

**RESULTS**

Figure 2 shows plots of CSFs estimated with both the H9023 and qCSF methods. In the following sections, we evaluate the precision and accuracy of qCSF estimates, examine the test–retest reliability of both CSF testing methods, and check qCSF model assumptions. We also evaluate the potential of qCSF in amblyopia screening, by testing how well CSF features estimated with the qCSF can predict membership in the two groups.

**Precision**

The precision of CSF estimates obtained with the qCSF and H9023 methods was estimated through a bootstrap procedure based on 500 simulated repetitions of full experimental runs of 300 qCSF and 300 H9023 trials. For each participant, the slope of the psychometric functions was assumed to be the slope of the best-fitting single-slope Weibull function (see Assumptions Check). For each observer, the precision in a qCSF or H9023 run was characterized by the standard deviation of the distribution of resampled threshold estimates:

\[
SD = \sqrt{\frac{\sum_{k}^{K} \sum_{i}^{N} (20 \log_{10}(\tau_k^i) - 20 \log_{10}(\overline{\tau}))^2}{N \cdot K - 1}}
\]

where \(k\) indexes spatial frequency, \(i\) indexes repetition, and \(\overline{\tau}\) is the mean threshold of \(N\) repetitions in the \(k\)th spatial frequency condition.

The precision of the qCSF method for each participant in each session was calculated trial by trial. There was no significant difference between two sessions for both groups (for 78.3% trials of the normal group and all trials of amblyop-
Bias

To evaluate the relative bias of qCSF estimates, we derived qCSF threshold estimates for the six spatial frequencies used in the \( \Psi \) method and compared them with those directly estimated by the qCSF method. The bias was defined as the difference between the thresholds obtained by the qCSF method and \( \Psi \) methods:

\[
\text{Bias}_i = \frac{\sum_{k=1}^{K}[20\log_{10}(\tau_{qCSF}^k) - \log_{10}(\tau_\Psi^k)]]}{K - 1},
\]

where \( k \) indexes spatial frequency, and \( \log(\tau_{qCSF}^k) \) is the average threshold from the two sessions.

The average bias of all participants in each group was computed as

\[
\text{Bias} = \frac{\sum_i \text{Bias}_i}{N},
\]

where \( i \) is the index of each participant. The SD of the bias was computed as

\[
\text{SD}_{\text{bias}} = \sqrt{\frac{\sum_i (\text{Bias}_i - \text{Bias})^2}{N - 1}},
\]

where \( N \) is the number of participants in each group.

The trial-by-trial bias was calculated for each subject in each session. There was no significant difference between the two sessions in both groups (99.7% for the normal group and 99.3% for the amblyopia group, \( P > 0.10 \); paired \( t \)-test). The average bias over the two sessions plotted as a function of the number of trials for the normal and amblyopia groups in Figure 4. Averaged over sessions, the mean bias of the qCSF method after 50, 100, and 300 trials was 0.32, -0.03, and -0.16 dB in the normal group, and -1.38, -0.78, and -0.24 dB in the amblyopia group. For both groups, these biases were not significantly different from 0 (\( P > 0.10 \); \( t \)-test).

Test–Retest Reliability

For both normal and amblyopic subjects, the test-retest reliability of CSF measures obtained with the qCSF and \( \Psi \) methods was evaluated. For both methods, we calculated the Pearson correlation between log thresholds measured in the first and second sessions, at six different frequencies. The test-retest correlations for the qCSF method were 0.944 and 0.960 (both \( P < 0.001 \)) for the normal and amblyopia groups, respectively (Fig. 5). Linear regression was also applied to analyze thresholds obtained from both groups. The slope for the \( \Psi \) method was 0.857 (confidence interval [CI]: 0.761–0.941), with \( r^2 = 0.892 \) in the normal group, and 0.977 (CI: 0.893–1.06), with \( r^2 = 0.922 \) in the amblyopia group.

Similar correlation and linear regression analyses were performed on the thresholds obtained by the qCSF method after 50, 100, and 300 trials (Fig. 6). For the normal group, the correlations for the qCSF estimates obtained with 50, 100, and 300 trials were 0.903, 0.926, and 0.951 (all \( P < 0.001 \)), respectively. For the amblyopia group, the correlations were 0.924, 0.951, and 0.945 (all \( P < 0.001 \)). For thresholds measured in 50 qCSF trials, the slope of the linear regression between the first and second sessions was 0.717 (CI: 0.627–0.806, \( r^2 = 0.815 \)) in the normal group, and 1.12 (CI: 0.983–1.26, \( r^2 = 0.855 \)) in the amblyopia group. After 100 qCSF trials, the slopes in the normal and amblyopia groups were 0.765 (CI: 0.683–0.847, \( r^2 = 0.858 \)) and 1.078 (CI: 0.975–1.18, \( r^2 = 0.905 \)). After 300 qCSF trials, the slopes in the normal and amblyopia groups were 0.820 (CI: 0.750–0.891, \( r^2 = 0.903 \)) and 1.095 (CI: 0.983–1.21, \( r^2 = 0.893 \)), respectively.
These results showed that, for both groups and both methods, threshold estimates in the two sessions were well matched with as few as 50 qCSF trials.

**Assumptions Check**

The qCSF method makes several assumptions: (1) that CSF can be described by a truncated log-parabolic function, and (2) that the slope of the psychometric function is invariant across spatial frequencies. Violations of these assumptions could introduce bias and imprecision into the qCSF estimates. We used the CSF data independently obtained with the $\Psi$ method to evaluate the assumptions in turn.

To verify how well the truncated log-parabolic model used by the qCSF can account for the empiric CSF data, we fit the model

![Graphs showing the threshold measurements for normal and amblyopic groups with linear regression for different trial counts.](image)

**FIGURE 5.** The thresholds measured by the $\Psi$ method in the first session were plotted against those measured in the second session. Linear regression was performed to evaluate the test-retest reliability.

![Graphs showing the threshold measurements for normal and amblyopic groups with linear regression for different trial counts.](image)

**FIGURE 6.** The thresholds measured by the qCSF method in the first session were plotted against those measured in the second session. Linear regression was performed to evaluate the test-retest reliability.
to the contrast sensitivities estimated at the six different spatial frequencies used in the \( \Phi \) method. Thirty-six sets of CSFs (18 participants in two sessions) measured by the \( \Phi \) method were fitted. In the normal group, the mean \( r^2 \) was 0.985 \( \pm \) 0.014 in the first session and 0.968 \( \pm \) 0.034 in the second session. In the amblyopia group, the mean \( r^2 \) was 0.955 \( \pm \) 0.018 and 0.965 \( \pm \) 0.021 in the first and second sessions, respectively. Consistent with previous reports, the truncated log-parabolic model provided an excellent description of the CSFs in the normal group. Our results also suggest that the same functional form can be used to describe CSFs in amblyopic vision.

In addition, we calculated the slope of the psychometric function at each of the six spatial frequencies, using the data obtained by the \( \Phi \) method. For each subject, data from the two sessions were pooled. The 100 trials in each spatial frequency condition were binned by dividing the log contrast range into five equal parts. The percentage of correct responses in the five bins allowed us to construct a raw psychometric function for each spatial frequency. The raw psychometric functions were then aligned by transferring the abscissa into a log relative contrast unit, \( \log_{10}(\text{contrast}) \).  

We then fit the relative psychometric functions with the Weibull function (equation 1), by using a maximum likelihood procedure.

\[
\text{Likelihood} = \prod_i N_i! K_i! (N_i - K_i)! P_i^{K_i} (1 - P_i)^{N_i - K_i},
\]

where \( i \) indexes the contrast condition, \( N_i \) and \( K_i \) are the number of total and correct trials, and \( P_i \) is the percentage correct predicted by the Weibull function. A full model with

\[
\log_{10} [\text{CSF}(sf)] = \begin{cases} 
\log_{10}(\gamma_{\text{max}}) - \delta, & \log_{10}(sf) < \log_{10}(f_{\text{max}}) - \beta/2 \sqrt{-\delta/\log_{10}(0.5)} \\
\log_{10}(\gamma_{\text{max}}) + \log_{10}(0.5) \left[ \frac{\log_{10}(sf) - \log_{10}(f_{\text{max}})}{\beta/2} \right]^2, & \text{otherwise}
\end{cases}
\]
six independent slopes at six different spatial frequencies and a reduced model with a single slope across six different spatial frequencies were compared by using the $\chi^2$ statistic:

$$\chi(df) = 2 \log \left( \frac{\text{max likelihood}_{\text{full}}}{\text{max likelihood}_{\text{reduced}}} \right),$$

where $df = k_{\text{full}} - k_{\text{reduced}}$ and $k$ is the number of parameters in each model.

In Figure 7, the raw psychometric functions along with the best-fitting reduced model of each participant were plotted. (Given that the $\Psi$ method places most trials near and above threshold, the number of trials in each bin is unevenly distributed. The lowest bins tend to contain only a few trials [1–5]; this results in the lower-than-chance performance [e.g., 0% correct] evident at the lowest relative contrast levels in Fig. 7.) The result showed that, for 8 of 10 normal and all 8 amblyopic participants, the full and the reduced models provided statistically equivalent fits to the data ($P > 0.1$). For these subjects, the slope invariance assumption is correct. Averaged over subjects, the best-fitting slope of the reduced model is $1.70 \pm 0.53$ in the normal group and $1.36 \pm 0.32$ in the amblyopia group, with no significant difference ($t_{15} = 1.69, P > 0.10$). The average slope over the two groups is $1.55 \pm 0.47$.

### CSF Differences between the Normal and Amblyopia Groups

As a summary metric of the CSF function, the area under the CSF (AULCSF) is a broad measure of spatial vision, whereas the spatial frequency cutoff (cutSF) characterizes the high-frequency resolution of the visual system. To characterize the contrast sensitivity differences between the normal and amblyopic participants, we calculated AULCSF and cutSF for each subject. The AULCSF was calculated as the integration from 0.5 cpd to the root of the log-parabola in the high-spatial-frequency range. The cutSF was defined as the spatial frequency at which contrast sensitivity is 2.0 (threshold = 0.5).

The average AULCSF and cutSF for both groups are shown as a function of trial number in Figure 8, along with running $P$-values from the $t$-test. The AULCSF differences between the groups became significant after approximately 40 trials ($2.29 \pm 0.36$ vs. $2.53 \pm 0.46$, $t_{34} = -1.70, P < 0.05$). After 39 trials, the cutoff frequency of the amblyopia group also became significantly lower than that of the normal group (19.4 ± 12.1 vs. 37.4 ± 26.5, $t_{34} = -2.51, P < 0.01$). After 50 trials, the significance level of both differences remained high ($P < 0.01$).

To explore the relationship between CSF features and visual acuity, we examined the correlations between AULCSF and cutSF with visual acuity. In the analysis, the averages of the final AULCSF and cutSF estimates from the two sessions were used, and the visual acuities of normal subjects were set at 1 minute. As showed in Figure 9, there was a negative correlation between the AULCSF and visual acuity ($r = -0.426, P < 0.01$). The correlation between the final cutSF and visual acuity was also significant ($r = -0.374, P < 0.05$). These analyses confirm a relationship between CSF metrics provided by the qCSF and visual acuity (the predominant clinical vision measure). The correlations are not significant if we restrict the analysis to the amblyopia group, perhaps due to the relatively small sample size. Although the CSF correlates with visual acuity, the two parameters may provide different measures of spatial vision.

### Screening Amblyopia According to qCSF Features

The significant difference of the AULCSF and cutSF between the normal and amblyopia groups prompted us to evaluate them as potential measures for amblyopia screening. A logistic
Clinical Application of qCSF  5373

DISCUSSION

In this study, we compared the CSFs of 10 normal subjects and 8 amblyopes measured by the qCSF\(^3\) and \(\Psi\)\(^3\) methods. The precision, accuracy, and test-retest reliability of the qCSF method were evaluated. Our analysis validated the assumptions of the qCSF method in observers with normal and amblyopic vision. The method's CSF model (truncated log-parabola) can well describe individual CSF data, and the psychometric slope is constant across different spatial frequencies for most subjects. Moreover, additional analyses showed that approximately 50 qCSF trials could capture the differences in CSF features (AULCSF and cutSF) between the normal and amblyopic groups. By adopting a logistic regression, we demonstrated the potential of qCSF for screening amblyopia with an accuracy of 77.8% in 50 trials. Taken together, our results demonstrate the potential of the qCSF method, an accurate and precise method of measuring CSFs in both normal people and amblyopes, as a tool for clinical practice.

To simplify the CSF model used in the qCSF, we fixed the psychometric slope at 3.5 across spatial frequencies based on the QUEST method's slope assumptions.\(^3\) However, the mean estimated value of slopes (1.55 \(\pm\) 0.47) for all participants was much less than 3.5. To examine the effects of this parameter mismatch on our results, we simulated CSF measurements on an observer whose underlying psychometric functions have a slope of 1.55, with both the qCSF and \(\Psi\) methods that assume four different predefined slopes (1, 1.55, 2, and 3.5). The simulation had 500 iterations, each of which consisted of 300 qCSF trials and 50 \(\times\) 6 \(\Psi\) trials.

The precision of both methods is plotted as a function of trial number in the top panel of Figure 11. The precision of the CSFs obtained with the \(\Psi\) method was 2.45, 2.21, 2.24, and 2.34 dB for the predefined slopes 1, 1.55, 2.5, and 3.5, respectively. The four precision curves generated by the qCSF method with the four different predefined slopes overlapped almost completely. This result suggests that the predefined slope value did not affect the precision of measurement of either method. In fact, the observed precision of the qCSF method for normal observers after 100 trials was 2.63 dB, with a predefined slope of 3.5. The precision is well matched to the result for normal subjects (2.79 dB), with a predefined slope of 2 reported by Lesmes et al.\(^3\)

The bias of the threshold estimates of both methods is plotted as a function of trial number in the bottom panel of Figure 11. The mean bias for the CSFs estimated using the \(\Psi\) method was 2.45, 2.21, 2.24, and 2.34 dB for the predefined slopes 1, 1.55, 2, and 3.5, respectively. However, the mean bias for the qCSF method was 2.45, 2.21, 2.24, and 2.34 dB for the predefined slopes 1, 1.55, 2, and 3.5, respectively. The four bias curves generated by the qCSF method with the four different predefined slopes overlapped almost completely. This result suggests that the predefined slope value did not affect the precision of measurement of either method. In fact, the observed precision of the qCSF method for normal observers after 100 trials was 2.63 dB, with a predefined slope of 3.5. The precision is well matched to the result for normal subjects (2.79 dB), with a predefined slope of 2 reported by Lesmes et al.\(^3\)

The bias of the threshold estimates of both methods is plotted as a function of trial number in the bottom panel of Figure 11. The mean bias for the CSFs estimated using the \(\Psi\) method was 2.45, 2.21, 2.24, and 2.34 dB for the predefined slopes 1, 1.55, 2, and 3.5, respectively. However, the mean bias for the qCSF method was 2.45, 2.21, 2.24, and 2.34 dB for the predefined slopes 1, 1.55, 2, and 3.5, respectively. The four bias curves generated by the qCSF method with the four different predefined slopes overlapped almost completely. This result suggests that the predefined slope value did not affect the precision of measurement of either method. In fact, the observed precision of the qCSF method for normal observers after 100 trials was 2.63 dB, with a predefined slope of 3.5. The precision is well matched to the result for normal subjects (2.79 dB), with a predefined slope of 2 reported by Lesmes et al.\(^3\)
method was 0.63, 0.09, -0.12 and -0.06 dB for predefined slopes 1, 1.55, 2.5, and 3.5, respectively. For the qCSF method, a predefined slope of 1, smaller than the actual value, led to threshold overestimates and a predefined slope larger than actual value caused threshold underestimates. The final bias of the qCSF method for a predefined slopes of 1, 1.55, 2.5, and 3.5 was 0.80, -0.21, -0.69, and -0.75 dB, respectively. The difference in bias between predefined slopes of 1.55 and 3.5 was -0.54 dB. Given the almost 2.5-fold difference in slope values between the underlying psychometric functions, this difference is very small (~6%). The simulation results suggest minimal effects of slope parameter mismatch on the conclusions of the current experiment, and that the qCSF and Ψ methods may work well without the very precise knowledge of the true slope. These results suggest that an assumed slope of 2.0 may be ideal for future qCSF applications.

To further investigate the effects of parameter mismatch, we obtained an additional data set from another six normal subjects using both the Ψ and qCSF methods with a predefined slope of 1.55. The average bias of these six subjects from the original measurements with a predefined slope of 3.5 and the new measurements are plotted in Figure 12. There was no significant difference between biases under the two slope settings for 98.3% of the trials (P > 0.1; t-test). For 94% of the trials, the bias with slope 1.55 did not differ from 0 (for these trials, P > 0.1; t-test). Taken together with the results of the previous analysis, we conclude that the steep slope value predefined in the qCSF program did not appreciably affect our results.

To further explore the flexibility of the qCSF method, we simulated CSF measurements on observers whose underlying psychometric functions have different slopes, 1, 1.55, 2.5, and 3.5, with both the qCSF and Ψ methods that assumed predefined slopes, 1, 1.55, 2.5, and 3.5 (a 4 × 4 design). The simulation had 500 iterations, each of which consisted of 300 qCSF trials and 50 × 6 Ψ trials. The precision and bias of the qCSF and Ψ methods under these 4 × 4 conditions are listed in Appendix B (see Tables A1 and A2). The results measured with predefined slope 3.5 are plotted in Figure 13. The final (300 trials) precision of the qCSF method, which used a model slope of 3.5, was 2.10, 1.45, 0.84, and 0.26 dB for observer slopes 1, 1.55, 2.5, and 3.5, respectively. The final bias was −1.02, −0.79, −0.29, and 0.01 dB for observer slopes 1, 1.55, 2.5, and 3.5, respectively. The precision and bias were proportional to the deviation between the slope of the observer’s psychometric function and the predefined slope. Even in the worst situation, in which the predefined slope 3.5 deviated well away from the slope of the observer’s psychometric function (1.0), the precision and bias were only 2.10 and −1.02 dB, respectively.

In test–retest analysis, the linear regression slope in the amblyopia group was closer to 1 than that in the normal group. It seemed that both methods were more reliable in the amblyopia group than in the normal group. However, it is worth noting that the normal participants had no previous experience in psychophysical experiments, and so, the slight loss of reliability in the normal group may be due to the learning effect across measurement sessions. To some extent, this finding is similar to that obtained in two other studies reported by Levi and Polat.60,61 In these experiments, novices improved substantially after training, but participants who were highly practiced did not. This phenomenon raises another important question: how to minimize effects of practice during measurement. There is plenty of evidence that even a single testing session can introduce learning effects,62 and overexposure to certain settings for 98.3% of the trials (P > 0.1; t-test). For 94% of the trials, the bias with slope 1.55 did not differ from 0 (for these trials, P > 0.1; t-test). Taken together with the results of the previous analysis, we conclude that the steep slope value predefined in the qCSF program did not appreciably affect our results.

To further explore the flexibility of the qCSF method, we simulated CSF measurements on observers whose underlying psychometric functions have different slopes, 1, 1.55, 2.5, and 3.5, with both the qCSF and Ψ methods that assumed predefined slopes, 1, 1.55, 2.5, and 3.5 (a 4 × 4 design). The simulation had 500 iterations, each of which consisted of 300 qCSF trials and 50 × 6 Ψ trials. The precision and bias of the qCSF and Ψ methods under these 4 × 4 conditions are listed in Appendix B (see Tables A1 and A2). The results measured with predefined slope 3.5 are plotted in Figure 13. The final (300 trials) precision of the qCSF method, which used a model slope of 3.5, was 2.10, 1.45, 0.84, and 0.26 dB for observer slopes 1, 1.55, 2.5, and 3.5, respectively. The final bias was −1.02, −0.79, −0.29, and 0.01 dB for observer slopes 1, 1.55, 2.5, and 3.5, respectively. The precision and bias were proportional to the deviation between the slope of the observer’s psychometric function and the predefined slope. Even in the worst situation, in which the predefined slope 3.5 deviated well away from the slope of the observer’s psychometric function (1.0), the precision and bias were only 2.10 and −1.02 dB, respectively.

In test–retest analysis, the linear regression slope in the amblyopia group was closer to 1 than that in the normal group. It seemed that both methods were more reliable in the amblyopia group than in the normal group. However, it is worth noting that the normal participants had no previous experience in psychophysical experiments, and so, the slight loss of reliability in the normal group may be due to the learning effect across measurement sessions. To some extent, this finding is similar to that obtained in two other studies reported by Levi and Polat.60,61 In these experiments, novices improved substantially after training, but participants who were highly practiced did not. This phenomenon raises another important question: how to minimize effects of practice during measurement. There is plenty of evidence that even a single testing session can introduce learning effects,62 and overexposure to certain settings for 98.3% of the trials (P > 0.1; t-test). For 94% of the trials, the bias with slope 1.55 did not differ from 0 (for these trials, P > 0.1; t-test). Taken together with the results of the previous analysis, we conclude that the steep slope value predefined in the qCSF program did not appreciably affect our results.

To further explore the flexibility of the qCSF method, we simulated CSF measurements on observers whose underlying psychometric functions have different slopes, 1, 1.55, 2.5, and 3.5, with both the qCSF and Ψ methods that assumed predefined slopes, 1, 1.55, 2.5, and 3.5 (a 4 × 4 design). The simulation had 500 iterations, each of which consisted of 300 qCSF trials and 50 × 6 Ψ trials. The precision and bias of the qCSF and Ψ methods under these 4 × 4 conditions are listed in Appendix B (see Tables A1 and A2). The results measured with predefined slope 3.5 are plotted in Figure 13. The final (300 trials) precision of the qCSF method, which used a model slope of 3.5, was 2.10, 1.45, 0.84, and 0.26 dB for observer slopes 1, 1.55, 2.5, and 3.5, respectively. The final bias was −1.02, −0.79, −0.29, and 0.01 dB for observer slopes 1, 1.55, 2.5, and 3.5, respectively. The precision and bias were proportional to the deviation between the slope of the observer’s psychometric function and the predefined slope. Even in the worst situation, in which the predefined slope 3.5 deviated well away from the slope of the observer’s psychometric function (1.0), the precision and bias were only 2.10 and −1.02 dB, respectively.

In test–retest analysis, the linear regression slope in the amblyopia group was closer to 1 than that in the normal group. It seemed that both methods were more reliable in the amblyopia group than in the normal group. However, it is worth noting that the normal participants had no previous experience in psychophysical experiments, and so, the slight loss of reliability in the normal group may be due to the learning effect across measurement sessions. To some extent, this finding is similar to that obtained in two other studies reported by Levi and Polat.60,61 In these experiments, novices improved substantially after training, but participants who were highly practiced did not. This phenomenon raises another important question: how to minimize effects of practice during measurement. There is plenty of evidence that even a single testing session can introduce learning effects,62 and overexposure to certain
tasks could lead to performance decrements.\textsuperscript{63,64} With these considerations, the fewer trials a measurement takes, the less effect it may have on the results. On the other hand, a sufficient number of trials are necessary to achieve the desirable precision. The qCSF method may be one of the solutions to this dilemma in CSF measurement.

In the present study, we used a logistic model to classify amblyopia. The result showed that the CSFs obtained in 50 qCSF trials classified amblyopia with a reasonably high accuracy. The qCSF method may be one of the solutions to this problem.\textsuperscript{65} On the other hand, a sufficient number of trials are necessary to achieve the desirable precision and classified with logistic regression. We believe that with improvements in computing power and more powerful statistical tools, the qCSF method will be appropriate for use in screening amblyopia and other visual impairments.

### APPENDIX A

Before measurement started, a prior probability distribution $p_0(\tilde{v})$, representing our knowledge about the tested CSF, was constructed with hyperbolic secants.\textsuperscript{75} A conditional probability lookup table $p(\text{correct} | \tilde{x}, \tilde{v})$ was initialized by calculating all the probabilities of correct response for all possible stimulus conditions $T_\text{s}$ and all possible parameters $T_\text{p}$. As prescribed by Kostevich and Tyler,\textsuperscript{26} before the $t+1$th trial began, the qCSF program calculated:

1. The probability of a correct response in a given stimulus condition $\tilde{x}$

   \[ p_{t+1}(\text{correct} | \tilde{x}) = \sum_\tilde{v} p(\text{correct} | \tilde{x}, \tilde{v}) p_0(\tilde{v}) \]  

2. The posterior probability distribution $p_{t+1}(\tilde{v})$ after a correct and an incorrect response to each possible stimulus $\tilde{x}$ in trial $t+1$

   \[ p_{t+1}(\tilde{v} | \tilde{x}, \text{correct}) = \frac{p(\tilde{v}) p(\text{correct} | \tilde{x}, \tilde{v})}{\sum_\tilde{v} p(\tilde{v}) p(\text{correct} | \tilde{x}, \tilde{v})} \]  

### Table A1. The Final Precision and Bias of the $\Psi$ Method

<table>
<thead>
<tr>
<th></th>
<th>OS 1</th>
<th>OS 1.55</th>
<th>OS 2.5</th>
<th>OS 3.5</th>
<th>Bias (dB)</th>
<th>OS 1</th>
<th>OS 1.55</th>
<th>OS 2.5</th>
<th>OS 3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS 1</td>
<td>3.19</td>
<td>2.45</td>
<td>2.04</td>
<td>1.90</td>
<td>0.04</td>
<td>0.63</td>
<td>1.23</td>
<td>1.44</td>
<td></td>
</tr>
<tr>
<td>MS 1.55</td>
<td>3.10</td>
<td>2.21</td>
<td>1.77</td>
<td>1.61</td>
<td>-0.35</td>
<td>0.09</td>
<td>0.44</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>MS 2.5</td>
<td>3.27</td>
<td>2.24</td>
<td>1.65</td>
<td>1.48</td>
<td>-0.14</td>
<td>-0.12</td>
<td>0.02</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>MS 3.5</td>
<td>3.56</td>
<td>2.34</td>
<td>1.67</td>
<td>1.37</td>
<td>0.26</td>
<td>-0.06</td>
<td>-0.01</td>
<td>0.04</td>
<td></td>
</tr>
</tbody>
</table>

OS and MS are short for observer slope and model slope separately.

### Table A2. The Precision and Bias of the qCSF Method

<table>
<thead>
<tr>
<th></th>
<th>OS 1</th>
<th>OS 1.55</th>
<th>OS 2.5</th>
<th>OS 3.5</th>
<th>Bias (dB)</th>
<th>OS 1</th>
<th>OS 1.55</th>
<th>OS 2.5</th>
<th>OS 3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 Trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS 1</td>
<td>4.10</td>
<td>3.62</td>
<td>3.02</td>
<td>3.15</td>
<td>-0.28</td>
<td>-0.03</td>
<td>0.26</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>MS 1.55</td>
<td>4.20</td>
<td>3.70</td>
<td>3.42</td>
<td>3.85</td>
<td>-0.24</td>
<td>-0.36</td>
<td>-0.37</td>
<td>-0.41</td>
<td></td>
</tr>
<tr>
<td>MS 2.5</td>
<td>4.61</td>
<td>4.63</td>
<td>3.03</td>
<td>2.89</td>
<td>-0.08</td>
<td>-0.51</td>
<td>-0.45</td>
<td>-0.29</td>
<td></td>
</tr>
<tr>
<td>MS 3.5</td>
<td>4.85</td>
<td>4.01</td>
<td>3.15</td>
<td>2.62</td>
<td>0.13</td>
<td>0.16</td>
<td>0.00</td>
<td>-0.20</td>
<td></td>
</tr>
<tr>
<td>100 Trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS 1</td>
<td>3.06</td>
<td>2.57</td>
<td>2.16</td>
<td>2.05</td>
<td>-0.78</td>
<td>-0.01</td>
<td>0.58</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>MS 1.55</td>
<td>3.16</td>
<td>2.54</td>
<td>2.07</td>
<td>1.79</td>
<td>-0.70</td>
<td>-0.53</td>
<td>-0.10</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>MS 2.5</td>
<td>3.35</td>
<td>2.57</td>
<td>1.86</td>
<td>1.52</td>
<td>-0.50</td>
<td>-0.37</td>
<td>-0.30</td>
<td>-0.09</td>
<td></td>
</tr>
<tr>
<td>MS 3.5</td>
<td>3.44</td>
<td>2.79</td>
<td>1.89</td>
<td>1.46</td>
<td>-0.44</td>
<td>-0.29</td>
<td>-0.29</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>300 Trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS 1</td>
<td>1.83</td>
<td>1.32</td>
<td>1.01</td>
<td>0.96</td>
<td>-0.41</td>
<td>0.80</td>
<td>1.67</td>
<td>2.06</td>
<td></td>
</tr>
<tr>
<td>MS 1.55</td>
<td>1.73</td>
<td>1.14</td>
<td>0.91</td>
<td>0.75</td>
<td>-1.07</td>
<td>-0.21</td>
<td>0.62</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>MS 2.5</td>
<td>1.99</td>
<td>1.24</td>
<td>0.56</td>
<td>0.30</td>
<td>-1.14</td>
<td>-0.69</td>
<td>-0.08</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>MS 3.5</td>
<td>2.10</td>
<td>1.41</td>
<td>0.84</td>
<td>0.26</td>
<td>-1.02</td>
<td>-0.75</td>
<td>-0.29</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

OS, observer slope; MS, model slope.
3. The entropies of the estimated posterior probability $p_{t+1}(\mathbf{v}|\mathbf{x}, \text{correct})$ after a correct and an incorrect response to each possible stimulus $\mathbf{x}$

$$H_{t+1}(\mathbf{x}, \text{correct}) = -\sum_{\mathbf{v}} p_{t+1}(\mathbf{v}|\mathbf{x}, \text{correct}) \times \log[p_{t+1}(\mathbf{v}|\mathbf{x}, \text{correct})]$$

$$H_{t+1}(\mathbf{x}, \text{incorrect}) = -\sum_{\mathbf{v}} p_{t+1}(\mathbf{v}|\mathbf{x}, \text{incorrect}) \times \log[p_{t+1}(\mathbf{v}|\mathbf{x}, \text{incorrect})].$$

4. The expected entropy after a trial with stimulus $\mathbf{x}$

$$\mathbb{E}[H_{t+1}(\mathbf{x})] = H_{t+1}(\mathbf{x}, \text{correct}) p_{t+1}(\text{correct}|\mathbf{x}) + H_{t+1}(\mathbf{x}, \text{incorrect}) p_{t+1}(\text{incorrect}|\mathbf{x}).$$

5. The stimulus providing the lowest expected entropy

$$\tilde{x}_{t+1} = \arg \min_{\mathbf{x}} \mathbb{E}[H_{t+1}(\mathbf{x})].$$

After the participants finished trial $t+1$ with stimulus condition $\tilde{x}_{t+1}$, the qCSF program calculated:

6. The posterior probability distribution $p_{t+1}(\mathbf{v}|\mathbf{x}, \text{correct})$ or $p_{t+1}(\mathbf{v}|\mathbf{x}, \text{incorrect})$ was updated as $p_{t+1}(\mathbf{v}|\mathbf{x}, \text{correct}) = p_{t+1}(\mathbf{v}|\mathbf{x}, \text{correct})$ or $p_{t+1}(\mathbf{v}|\mathbf{x}, \text{incorrect})$.

7. The expected value of $\mathbf{v}$ was calculated

$$\tilde{\mathbf{v}} = \sum_{\mathbf{v}} \tilde{p}_{t+1}(\mathbf{v}).$$

After 300 trials were run, the program was terminated.

**APPENDIX B**

**Parameter Mismatch**

We simulated qCSF measurements on observers whose underlying psychometric functions have slopes of 1, 1.55, 2.5, and 3.5, with both the qCSF and $\Psi$ methods that assumed predefined slopes of 1, 1.55, 2.5, and 3.5. The simulation had 500 iterations, each of which consisted of 300 qCSF trials and 50 × 6 $\Psi$ trials. The precision and bias were calculated. The precision and bias estimated in 50, 100, and 300 trials are shown in Tables A1 and A2 (unit: dB).

**References**


30. Pesudovs K, Hazel CA, Doran RM, Elliott DB. The usefulness of Vistech and FACT contrast sensitivity charts for cataract and


