Using 10AFC to further improve the efficiency of the quick CSF method

Fang Hou Laboratory of Brain Processes, The Ohio State University, Columbus, OH, USA

Luis Lesmes Adaptive Sensory Technology, Boston, MA, USA

Peter Bex Department of Psychology, Northeastern University, Boston, MA, USA

Michael Dorr Institute for Human–Machine Communication, Technische Universität München, Munich, Germany

Zhong-Lin Lu Laboratory of Brain Processes, The Ohio State University, Columbus, OH, USA

The contrast sensitivity function (CSF) provides a fundamental characterization of spatial vision, important for basic and clinical applications, but its long testing times have prevented easy, widespread assessment. The original quick CSF method was developed using a two-alternative forced choice (2AFC) grating orientation identification task (Lesmes, Lu, Baek, & Albright, 2010), and obtained precise CSF assessments while reducing the testing burden to only 50 trials. In this study, we attempt to further improve the efficiency of the quick CSF method by exploiting the properties of psychometric functions in multiple-alternative forced choice (*m*-AFC) tasks. A simulation study evaluated the effect of the number of alternatives *m* on the efficiency of the sensitivity measurement by the quick CSF method, and a psychophysical study validated the quick CS method in a 10AFC task. We found that increasing the number of alternatives of the forced-choice task greatly improved the efficiency of CSF assessment in both simulation and psychophysical studies. The quick CSF method based on a 10-letter identification task can assess the CSF with an averaged standard deviation of 0.10 decimal log unit in less than 2 minutes.

Introduction

The contrast sensitivity function (CSF), which describes visual sensitivity (1/contrast threshold) to narrow-band stimuli of different spatial frequencies, provides a comprehensive measure of the visual system over a wide range of spatial frequencies in both normal and abnormal vision (Ginsburg, 1981, 2003; Hess, 1981). The CSF is closely related to daily visual functions (Ginsburg, 2003), because our visual environment consists of visual stimuli with a wide range of contrasts and a broad spatial frequency spectrum. The CSF has proved important in characterizing functional deficits in visual disorders such as amblyopia (Hess & Howell, 1977; Jindra & Zemon, 1989; Onal, Yenice, Cakir, & Temel, 2008), multiple sclerosis (Ginsburg, 1981; Shandiz et al., 2010), and glaucoma (Arden & Jacobson, 1978; Richman et al., 2010). It has been suggested that the CSF characterizes spatial vision deficits better than letter acuity (Hess & Howell, 1977; Jindra & Zemon, 1989; Marmor, 1981, 1986; Marmor & Gawande, 1988; Montes-Mico & Ferrer-Blasco, 2001; Onal et al., 2008; Yenice et al., 2007). It has also been reported that even when acuity appears normal, patients may have evident CSF deficits (Huang, Tao, Zhou, & Lu, 2007; Jindra & Zemon, 1989; Woods & Wood, 1995). Although many visual disorders are related to general reductions in contrast sensitivity, selective contrast sensitivity reduction in different ranges of spatial frequencies has been documented (Regan, 1991). The variability in CSF deficits between and within visual pathologies suggests that measuring

Citation: Hou, F., Lesmes, L., Bex, P., Dorr, M., & Lu, Z.-L. (2015). Using 10AFC to further improve the efficiency of the quick CSF method. *Journal of Vision, 15*(9):2, 1–18, doi:10.1167/15.9.2.

doi: 10.1167/15.9.2 Received November 25, 2014; published July 10, 2015 ISSN 1534-7362

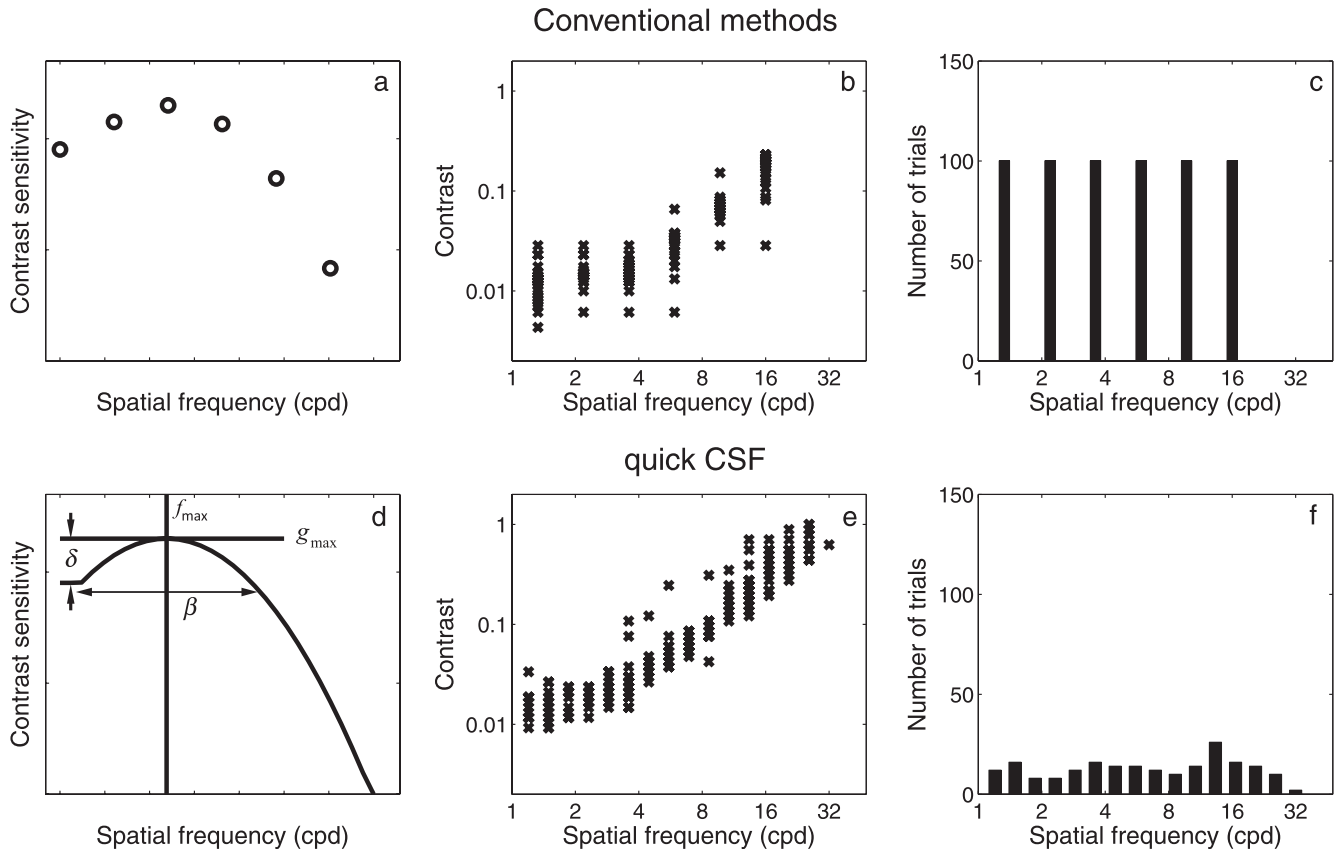

Figure 1. Comparison between the quick CSF and convention methods. Conventional methods measure contrast sensitivity a few predefined spatial frequencies (a), select stimuli in contrast space (b) at each spatial frequency, and distribute trials evenly over spatial frequencies (c). In contradistinction, the quick CSF method adopts a four-parameter log parabolic functional form (d), selects stimuli in both contrast and frequency spaces (e), and allocates trials more efficiently over two-dimensional space (f). To achieve the same precision, the quick CSF method requires many fewer trials.

the full CSF over a wide range of frequencies is clinically important.

Despite the importance of assessing the full CSF, the testing time needed for precise assessment has prevented its clinical application. Conventionally, a typical CSF assessment estimates a set of contrast thresholds at 5–10 predetermined spatial frequencies (Figure1a). Each threshold is measured by an adaptive method (Treutwein, 1995) with about 100 trials. In total, a full CSF assessment requires 500–1,000 trials that are evenly distributed over tested spatial frequencies (Figure 1b and c) and takes 30–60 minutes (Kelly & Savoie, 1973). Such testing times might be acceptable for measuring a single CSF in the laboratory, but are prohibitive in situations requiring assessment of multiple CSFs (e.g., both eyes) and/or in clinical settings.

Recently, Lesmes et al. (2010) developed the quick CSF method, a novel Bayesian adaptive psychophysical method, which accurately estimates the CSF in less than 50 trials. In this method, the CSF is characterized by a *truncated log parabola* (Lesmes et al., 2010; Watson & Ahumada, 2005; Figure 1d) with four parameters: peak gain g_{max}, peak spatial frequency f_{max}, bandwidth at half-height β (in octaves), and low-frequency truncation level δ. Using a Bayesian adaptive algorithm (Cobo-Lewis, 1996; Kim, Pitt, Lu, Steyvers, & Myung, 2014; King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994; Kontsevich & Tyler, 1999; Kujala & Lukka, 2006; Lesmes, Jeon, Lu, & Dosher, 2006; Watson & Pelli, 1983) to select the optimal test stimulus and update the posterior probabilities of CSF parameters following each trial, the quick CS method directly estimates the entire CSF curve instead of sensitivities at some predetermined spatial frequencies (See Appendix A for more details). Unlike the conventional methods that select stimuli adaptively in only contrast space, the quick CSF method searches stimuli in both contrast and frequency spaces (Figure 1e and f), making it more efficient. For a 2AFC grating orientation identification task, only 5–10 minutes are needed to obtain a CSF with a 0.10–0.20 decimal log unit standard deviation, comparable to conventional methods using much longer testing times.

Since its debut, the quick CSF method has been further tested and applied in several studies. Hou et al. (2010) validated the quick CSF method in patients with

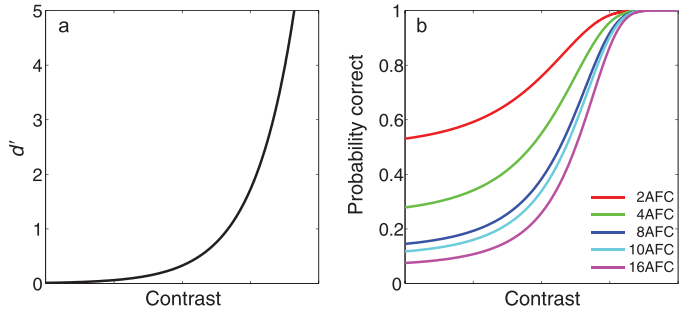

Figure 2. The probability correct psychometric functions for 2, 4, 8, 10, and 16 AFC tasks (b) with the same underlying d' psychometric function (a). Different colors indicated different numbers of alternatives.

amblyopia and demonstrated that the procedure can achieve a 0.24 log unit standard deviation with 50 trials. Dorr, Lesmes, Lu, & Bex (2013) implemented the procedure on a tablet device and demonstrated that CSFs obtained on a tablet device are comparable to those obtained with specialized laboratory equipment. The method has been applied to investigate the dynamic effects of visual adaptation (Gepshtein, Lesmes, & Albright, 2013) and emotional arousal (Lee, Baek, Lu, & Mather, 2014) on the contrast sensitivity function. It has also been validated in applications studying peripheral vision (Rosén, Lundström, Venkataraman, Winter, & Unsbo, 2014) and second-order perception (Reynaud, Tang, Zhou, & Hess, 2014), with a 0.07 log unit standard deviation with approximate 10 min of testing time reported by Rosen et al. (2014). With the procedure implemented on a tablet device, Kalia et al. (2014) demonstrated surprising visual development in a unique sample of patients who experienced extended early-onset blindness before removal of bilateral congenital cataracts. The quick CSF method has also been validated in populations with macular degeneration (Lesmes, Wallis, Jackson, & Bex, 2013; Lesmes, Wallis, Lu, Jackson, & Bex, 2012; Rosen et al., 2015) and glaucoma (Ramulu, Dave, & Friedman, 2015). Several quick CSF studies on clinical populations, including age-related macular degeneration (AMD), amblyopia, and glaucoma, showed that a similar or slightly (<25%) higher number of trials was required to achieve the same precision in clinical populations as in normal subjects, and the test precision did not depend on the patients' overall level of visual deficits (Hou et al., 2010; Lesmes et al., 2013; Lesmes et al., 2012; Rosen et al., 2015).

The original quick CSF method was based on a 2AFC task. In this study, we attempt to further improve the efficiency of the quick CSF method by exploiting the properties of psychometric functions in multiple-alternative forced choice (m-AFC) tasks. Increasing the number of alternatives in m-AFC tasks has two effects on psychometric functions. First, it reduces the guessing rate and therefore makes each trial more informative. Second, it increases the slope of the psychometric function (Figure 2b).

Several studies have suggested that increasing the slope of the psychometric function improves the efficiency of threshold estimates. Taylor (1971) found that the ideal sweat factor (Taylor & Creelman, 1967) of an adaptive method is related to the slope of the psychometric function. Patterson, Foster, and Heron (1980) found that increasing the slope of the psychometric function improved the precision of the estimated threshold. Alcalá-Quintana and García-Pérez (2004) demonstrated that the variance of estimated parameters in a Bayesian adaptive method was related to the steepness of psychometric function. In addition, many studies have shown that psychophysical efficiency increases with the number of alternatives in m-AFC tasks. For example, Hall (1983) and Shelton and Scarrow (1984) found that thresholds obtained with a 3AFC auditory task were less variable than those obtained with a 2AFC task. Bi, Lee, and O'Mahony (2010) concluded that the 4AFC task was statistically more powerful than the 2AFC task in food flavor discrimination. Using a contrast sensitivity function assessment test, Jäkel and Wichmann (2006) also reported that a 4AFC task was 3.5 times more efficient than a 2IFC task in contrast detection. In a simulation study, Leek, Hanna, and Marshall (1992) used the up-down staircase procedure with 2AFC, 3AFC, and 4AFC tasks to estimate both the threshold and slope of the psychometric function. They found that the test efficiency for slope estimate, defined as the sweat factor, was highest for the 4AFC task and declined when the number of alternatives decreased.

In the quick CSF method, the subjects' behavior on each trial is modeled by a psychometric function. We hypothesize that performance of the method depends on the shape of the psychometric functions, and that increasing the number of alternatives (m) in m-AFC tasks would improve the efficiency of the quick CSF method. Indeed, an earlier simulation study found that the average standard deviation of CSFs obtained from the quick CSF method decreased from 0.13 to 0.07 log unit when the slope of the log-Weibull psychometric function increased from 1.55 to 3.5 (Hou et al., 2010).

In this paper, we first describe a systematic simulation study of the effect of the number of alternatives (m) in m-AFC tasks on the precision of the quick CSF method. We then report a psychophysical validation experiment of the quick CSF method in a 10AFC task. The CSFs of five normal observers were measured with both the quick CSF and conventional methods. We also compared the efficiency of the quick CSF method based on the 10AFC task with that based on a 2IFC task in a published study (Hou et al., 2010).

Simulation

Psychometric functions in *m*-AFC

In the simulation study, we used the quick CSF method to measure the contrast sensitivity function corresponding to thresholds at a fixed $d' = 1.5$[1] for a simulated observer performing an *m*-AFC task, with $m = 2$, 4, 8, 10, and 16. The d' psychometric function of the simulated observer is described by the following equation (Foley & Legge, 1981; Legge, Kersten, & Burgess, 1987; Figure 2a):

$$d'(c,f) = 1.5\left(\frac{c}{\tau(f)}\right)^{\zeta} \quad (1)$$

where c is the contrast of the stimulus, $\tau(f)$ is the contrast threshold at $d' = 1.5$ in the f spatial frequency condition, and ζ is the log-log slope of the d' psychometric function. We set $\zeta = 2.35$ in the simulation based on typical values in the literature (Foley & Legge, 1981; Legge et al., 1987; Lu & Dosher, 1999). In this formulation, ζ is independent of the threshold level $\tau(f)$; that is, the d' psychometric functions in different spatial frequency conditions have different thresholds but exactly the same shape and are only shifted horizontally on the low contrast axis.

In an *m*-AFC task, the probability correct psychometric function of the simulated observer can be derived from the d' psychometric function (Hacker & Ratcliff, 1979):

$$P(c,f,m) = \int_{-\infty}^{+\infty} \phi\Big(x - d'(c,f)\Big)\Phi^{m-1}(x)dx \quad (2)$$

where $\phi()$ and $\Phi()$ are the probability density and cumulative probability density functions of a standard normal distribution, $d'(c,f)$ is the d' value associated with a stimulus with contrast c and spatial frequency f, and m is the number of alternatives. Although the slope of the d' psychometric function ζ is invariant to the number of alternatives in the *m*-AFC task, the slope of the probability correct psychometric function depends on the number of alternatives, m (Figure 2b).

Human observers inevitably make occasional finger errors in experiments. The behavior is modeled by considering a lapse rate λ that is independent of stimulus level[2] (Klein, 2001; Wichmann & Hill, 2001):

$$P'(c,f,m) = (1-\lambda)P(c,f,m) + \gamma\lambda \quad (3)$$

where $P(c,f,m)$ is the psychometric function without lapse (Equation 2). For the simulated observer and in the quick CSF method, λ was set to 0.04 (Lesmes et al., 2010; Wichmann & Hill, 2001).

Based on signal detection theory, Equation 2 provides a concise way to describe psychometric functions in *m*-AFC tasks. However, the computational load for integration in Equation 2 is very heavy; Weibull functions are used in our simulation and the quick CSF procedure to approximate these functions and greatly reduce the computational load. In Appendix B, we provide details on Weibull approximations.

After specifying the parameters of the underlying psychometric functions, the response probabilities of the simulated observer in the *m*-AFC tasks in all possible stimulus conditions can be computed. We then used these probabilities to simulate the response of the observer in the quick CSF procedure, which is used to infer the underlying contrast threshold function, $\tau(f)$, of the observer.

Simulation

In the simulation, we assumed that the underlying CSF parameters of the simulated observer were $\theta^{\text{true}} = (g_{\max}^{\text{true}}, f_{\max}^{\text{true}}, \beta^{\text{true}}, \delta^{\text{true}}) = (80, 1.07, 3.6, 0.3)$. The values were based on the average CSF parameters from the psychophysical experiment (see Psychophysical study). These values were used to calculate the contrast thresholds $\tau^{\text{true}}(f)$, which in turn were used to generate the simulated observer's response in each trial by Equation 3.

The quick CSF method was used to estimate the CSF underlying the response of the simulated observer. In the procedure, we assume that the shape of the underlying psychometric function, defined by ζ, γ, and λ (Equation 3), is known. Hou et al. (2010) demonstrated that the assumption of the log-invariant psychometric function was largely correct. Only $\tau(f)$ was being estimated by the quick CSF method. Again, the detailed quick CSF algorithm is described in Appendix A.

Evaluation procedure

The quick CSF method implemented with 2, 4, 8, 10, and 16 AFC tasks was used in the simulation. For each task, the observer "ran" the quick CSF procedure 500 times, with 300 trials in each run. To obtain CSF estimates from each run, 1,000 sets of CSF parameters were sampled from the posterior distribution of CSF parameters, $p_t(\theta)$, and used to construct 1,000 CSF curves. Based on these CSF curve samples, we obtained the empirical distribution of the CSF, $p_t(\tau)$. Each CSF curve was evaluated at 20 spatial frequencies ranging from 0.5 to 32 cpd, evenly distributed in log space. This resampling procedure automatically takes into account the covariance structure in the posterior distribution of

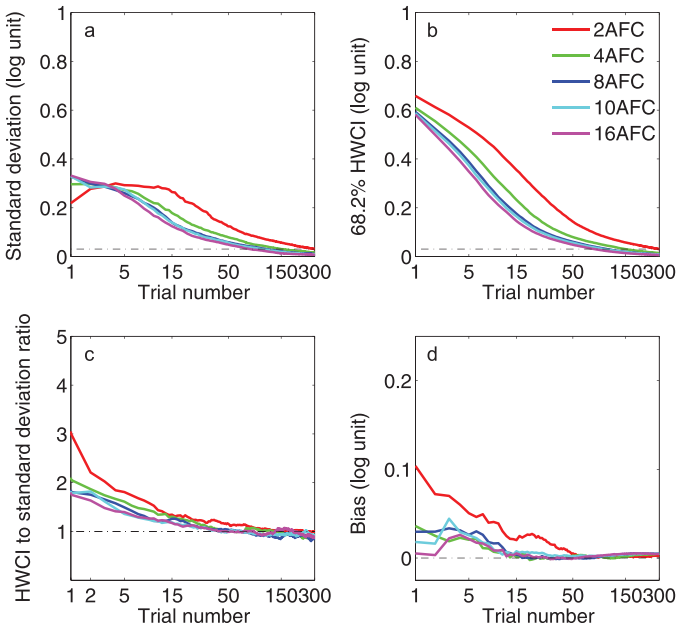

Figure 3. (a) Standard deviation of the CSFs obtained by the quick CSF method with *m*-AFC tasks as functions of trial number. (b) HWCI of the CSF obtained by the quick CSF method with *m*-AFC tasks as functions of trial number. (c) The ratios between HWCI and standard deviation for all *m*-AFC tasks as functions of trial number. (d) Bias of the CSFs obtained by the quick CSF method as functions of trial number. Different *m*-AFC tasks are represented by different colors.

the CSF parameters, and allows us to compute variance of the estimated CSF curve.

The performance of the quick CSF method, gauged by the precision and accuracy, was evaluated after each trial. The precision of the method is defined as the reciprocal of the standard deviation of the estimated sensitivities across different quick CSF runs:

$$Precision = \frac{1}{Standard\ \ deviation} \tag{4}$$

with

$$Standard\ \ deviation = \frac{\sum_{k=1}^{20}\sqrt{\frac{\sum_{j=1}^{500}\left(\left(\log_{10}(\tau_j^k)-\log_{10}(\tau^{k'})\right)^2\right)}{500-1}}}{20} \tag{5}$$

where $\tau_j^k = \frac{\sum_{i=1}^{1000}\tau_{ij}^k}{1000}$ is the estimated sensitivity (in log units) at the kth spatial frequency in the jth run, τ^k is the average of τ_j^k across all runs; $k = 1, 2, \ldots 20$ is the index of spatial frequencies, $i = 1, 2, \ldots 1{,}000$ is the index of CSF samples from the posterior distribution of a single quick CSF run, and $j = 1, 2, \ldots 500$ is the index of the quick CSF runs. The accuracy of the quick CSF method, defined by the bias of the estimated CSF is calculated as the mean differences between the measured and true sensitivities:

$$Bias = \frac{\sum_{k=1}^{20}\sum_{j=1}^{500}\left(\log_{10}(\tau_j^k) - \log_{10}(\tau^{\mathrm{true},k})\right)}{20\times 500} \tag{6}$$

where $\tau^{k'} = \frac{\sum_{j=1}^{500}\tau_j^k}{500}$ and $\tau^{\mathrm{true},k}$ are the estimated and true contrast sensitivity at the kth spatial frequency. The standard deviation and bias of sensitivity estimates were averaged across all spatial frequencies and are both in decimal log sensitivity units.

The standard deviation of the estimated CSF is computed from repeated measures. Alternatively, the variability of the measured sensitivity can be described by the width of the credible interval of the posterior distribution $p_t(\tau)$ from a single quick CSF run. A 68.2% credible interval represents the shortest interval that contains the actual value with 68.2% probability (Clayton & Hills, 1993). Since it is uncommon to repeat the same measurement multiple times in clinical practice, the credible interval of the posterior distribution is a valuable tool to gauge the precision of a test in a single run. Here we report the half width of the 68.2% credible interval (HWCI) for a single quick CSF run. We choose 68.2% credible interval because if the posterior distribution is Gaussian, the 68.2% HWCI is equal to the standard deviation of the distribution. The HWCI is also in unit of decimal log sensitivity.

The efficiency of a procedure is defined as its precision divided by the number of trials. To compare efficiencies of different procedures, we define the relative efficiency as the efficiency ratio between two procedures.

Results

Figure 3a and b presents the standard deviation and 68.2% HWCI of the CSF measured by the quick CSF procedure as functions of trial number for 2, 4, 8, 10, and 16 AFC tasks. Both types of functions demonstrate similar convergence patterns. For both standard deviation and HWCI, the curves from the *m*-AFC tasks with different numbers of alternatives are approximately laminated. The downward shift between curves reflects increasing precision as the number of alternatives increases from 2 to 16.

The standard deviation of CSFs obtained with the 2, 4, 8, 10, and 16 AFC quick CSF procedure after 50 trials is 0.12, 0.08, 0.06, 0.06, and 0.05 log unit, respectively. With 300 trials, the standard deviation is 0.03, 0.02, and 0.01 log unit for 2AFC, 4AFC and 8AFC, respectively, and <0.01 log unit for 8AFC and 16AFC. The HWCI for CSFs obtained with the 2, 4, 8, 10, and 16 AFC quick CSF procedure after 50 trials is 0.14, 0.08, 0.06, 0.06, and 0.05 log unit, respectively.

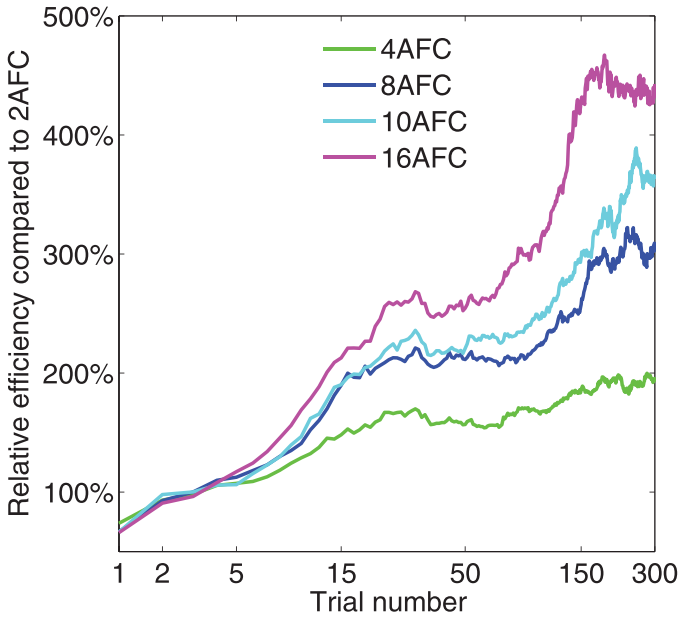

Figure 4. The relative efficiencies of the quick CSF procedures with 4, 8, 10, and 16 AFC tasks as functions of trial number.

With 300 trials, the HWCI is 0.03 and 0.01 log unit for 2 and 4 AFC, and <0.01 log unit for 8, 10, and 16 AFC. In addition, the magnitudes of standard deviation and HWCI are very similar. We plot the ratio of HWCI to standard deviation for all the quick CSF procedures in Figure 3c. Generally, after 15 trials, the ratio is very close to 1 for all the quick CSF procedures. The result indicates that the HWCI captured within a single run is very closely related to the standard deviation from repeated measures. This property can be potentially exploited to estimate the precision of a single quick CSF run in practical applications.

Figure 3d shows the biases of CSF estimates, obtained from the quick CSF method using 2, 4, 8, 10, and 16 AFC tasks, as functions of trial number. No systematic bias was found for all these procedures. In about 15 trials, the bias of threshold estimates obtained in all but the 2AFC tasks is less than 0.01 log unit. To reach 0.01 log unit bias, the quick CSF procedure with the 2AFC task requires at least 44 trials. The bias is 0.003, 0.005, 0.005, 0.005, and 0.005 log unit for the 2, 4, 8, 10, and 16 AFC tasks after 300 trials, respectively.

Figure 4 shows the relative efficiencies of the 4, 8, 10, and 16 AFC quick CSF procedures compared to the 2AFC quick CSF procedure. All relative efficiencies increase with increasing number of alternatives and number of trials. After 50 trials, the relative efficiency of the quick CSF procedure with 4, 8, 10, and 16 AFC is 156%, 211%, 221%, and 255%, respectively. After 300 trials, the relative efficiency of the quick CSF procedure with 4, 8, 10, and 16 AFC is 195%, 305%, 366%, and 443%, respectively. Taken together, the simulation results show that increasing the number of alternatives in an *m*-AFC task can substantially improve the efficiency of the quick CSF method. It also suggests that the benefit of large alternative number is even greater with more trials.

It has been known that one major concern about Bayesian adaptive methods is that these methods may be unstable if there are some lapse trials at the beginning of the experiment (Kontsevich & Tyler, 1999), which usually contains stimuli way above the threshold. To evaluate the effect of lapse, we simulated three observers with a 100% lapse rate (Equation 3) in the first 1, 3, and 5 trials, respectively. The observers are otherwise the same as the simulated observer in our simulation study who had a 0.04 lapse rate. The CSFs of the simulated observers were measured by the 10AFC quick CSF procedure. The standard deviation, HWCI and bias of the estimated CSFs are plotted as functions of trial number in Figure 5, along with those of the 10AFC observer from Figure 2. Lapse in the beginning of the quick CSF procedure significantly impacted the quality of CSF estimated by the quick CSF procedure. To reach a standard deviation of 0.1 log unit, the observer who makes a lapse in the first 1, 3, and 5 trials needs 27, 38,

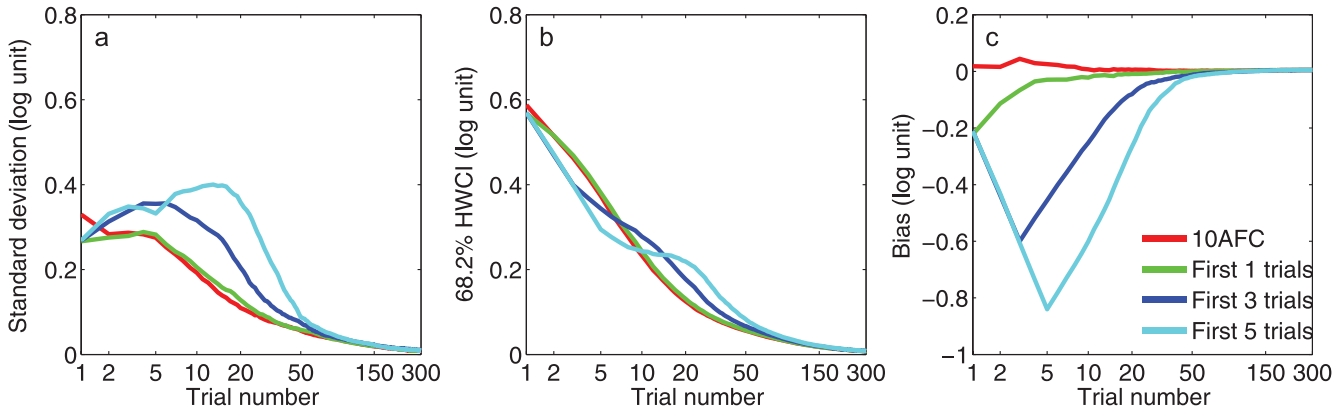

Figure 5. (a) Standard deviation, (b) half width of 68.2% credible interval, and (c) bias of the CSF measured by the 10AFC quick CSF procedure as functions of trial number. Red curves represent the result of an observer with a 0.04 lapse rate; green, blue, and cyan curves represent the results of an observer with 100% lapse in the first 1, 3, and 5 trials, respectively.

and 50 trials, respectively, compared to 24 trials for the observer with a 0.04 lapse rate. To reach a HWCI of 0.1 log unit, the observer who makes a lapse in the first 1, 3, and 5 trials needs 28, 35, and 45 trials, respectively, compared to 27 trials for the observer with a 0.04 lapse rate. To reach a 0.025 log unit bias, the observer who makes a lapse in the first 1, 3, and 5 trials needs 9, 37, and 47 trials, respectively, compared to 7 trials for the observer with a 0.04 lapse rate. These results suggest that the quick CSF procedure can recover from the detrimental impact of multiple lapses in the beginning of the experiment after about 50 trials.

Psychophysical study

The simulation study demonstrated the substantial benefits of increasing the number of alternatives in an *m*-AFC task. In this section, we report a psychophysical validation study of the 10AFC quick CSF procedure with a 10-letter identification task, in which observers were asked to identify a randomly chosen letter presented on the screen in each trial.

Our stimuli were based on the Sloan letters, C, D, K, H, N, O, R, S, V, and Z, which are widely used in optometry clinics to provide nearly identical percent correct for identification across the letter set (NAS-NRC Committee on Vision, 1981; Sloan, Rowland, & Altman, 1952). The original letter images are broadband in spectrum. They were band-pass filtered and resized to generate narrow-band stimuli to assess contrast sensitivity in different central spatial frequencies (Alexander, Xie, & Derlacki, 1994; McAnany & Alexander, 2006). In addition, the empirical Weibull psychometric function for 10AFC identification of filtered Sloan letter task has been extensively studied (Hou, Lu, & Huang, 2014). We set the slope of the Weibull psychometric function to 2.74 based on the literature.

To validate the 10AFC quick CSF procedure, CSFs were obtained with the procedure and compared to the "true" CSFs obtained with a conventional procedure that measured individual contrast thresholds in several spatial frequencies. We chose the Psi method (Kontsevich & Tyler, 1999) and set the slope of psychometric function equal to 2.74 in the procedure to measure individual thresholds. We also compared the relative efficiencies of the quick CSF method based on the data from the 10AFC quick CSF procedure in this study and data from a 2IFC quick CSF procedure in a published paper (Hou et al., 2010).

Method

Observers

The first author (S1) and four other observers (S2–S5), aged 23 to 33 years, participated in the study. All observers had normal or corrected-to-normal vision. All observers except S1 were naive to the purpose of the study. The study was approved by the institutional review board of human subjects research of The Ohio State University. Written informed consent was obtained before the experiment.

Apparatus

All programs used in this study were coded in MATLAB (MathWorks Corp., Natick, MA) with Psychtoolbox extensions (Kleiner, Brainard, & Pelli, 2007) and run on a PC computer. Stimuli were displayed on a gamma-corrected ViewSonic CDE3201LED 32″ monitor with a mean luminance of 120 cd/m^2, a 1920 × 1080 pixel resolution, and a vertical refresh rate of 60 Hz. A special circuit was used to achieve 14-bit grayscale resolution (Li & Lu, 2012; Li, Lu, Xu, Jin, & Zhou, 2003). Participants viewed the stimuli at a distance of 5 m in a dark room.

Stimuli

To generate the stimuli, each 256 × 256 pixel white (RGB value 255) letter was centered in a 512 × 512 pixel black (RGB value 0) background and filtered with a raised cosine filter (Chung, Legge, & Tjan, 2002):

$$filter(f) = \frac{1 + \cos\left(\frac{\log(f) - \log(f_0)}{\log(f_{\text{cutoff}}) - \log(f_0)}\pi\right)}{2} \quad (7)$$

where f denotes radial spatial frequency, $f_0 = 3$ cycles per object (cpo) is the center frequency of the filter, and $f_{\text{cutoff}} = 2f_0$ was chosen such that the full bandwidth at half height is 1 octave. The pixel intensity of each filtered image was normalized by the maximum absolute intensity of the image such that, after normalization, the maximum absolute Weber contrast of the image is 1.0 (Figure 6a). Stimuli with different contrasts were obtained by scaling the intensities of the normalized images with corresponding values. The filtered images were rescaled to 16 different sizes to generate stimuli with 16 evenly spaced (in log space) central spatial frequencies ranging from 1.33 to 32.0 cpd for the quick CSF procedure (Figure 6b). For the conventional method, stimuli at six evenly spaced (in log space) central spatial frequencies ranging from 1.33 to 16.0 cpd were generated.

Procedure

Each trial began with the presentation of a crosshair fixation pattern (367 ms), followed by a blank screen with mean luminance (183 ms), and stimulus presentation (183 ms). A response screen with all 10 letters was shown 500 ms after stimulus presentation to facilitate the

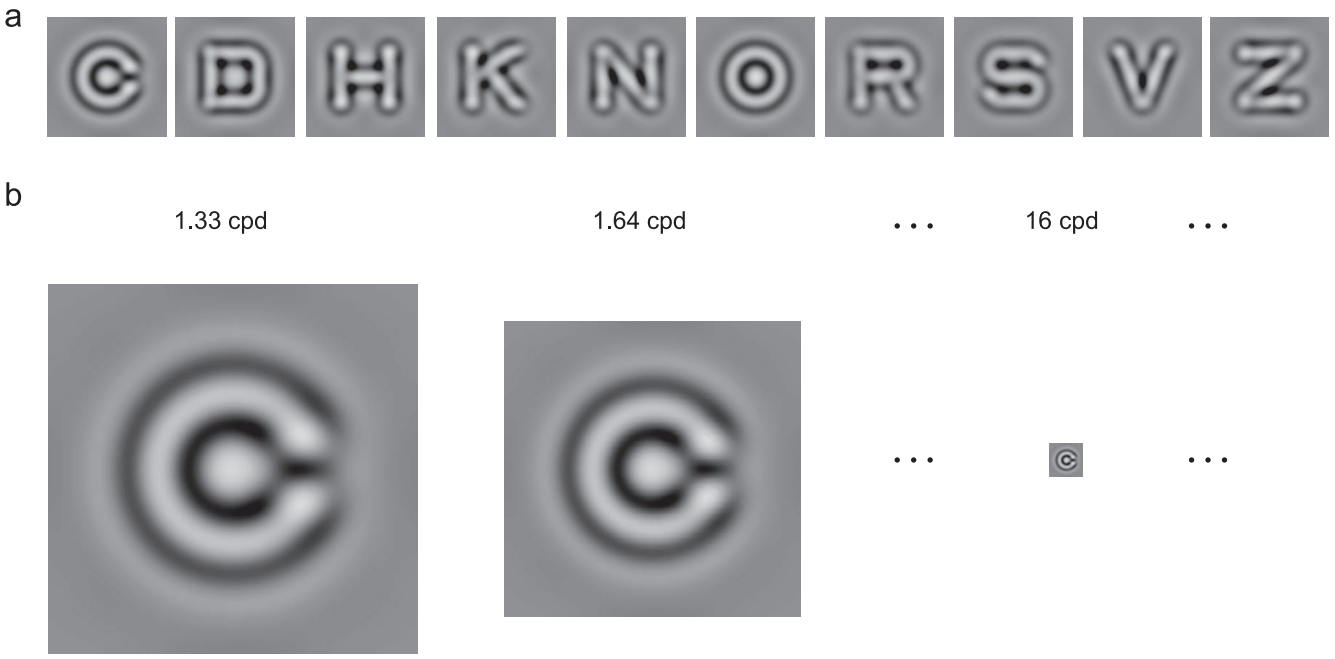

Figure 6. (a) Ten filtered letters. (b) Illustration of filtered letter "C" in different spatial frequency conditions.

response collection. Unfiltered letters were used on the response screen. The size of the letters on the response screen was 0.80 degree × 0.80 degree with a 0.70 degree gap between adjacent letters. Letters were arranged as a 2 × 5 matrix and presented in the center of the display. With a Weber contrast of −0.4, the RMS contrast of the response screen was comparable to the average RMS contrast of the stimuli used throughout the experiment. Observers were instructed to use the keyboard to type or mouse to select the letter they saw. No feedback was provided. A new trial started 500 ms after response.

Each observer finished one experimental session that, with voluntary breaks, lasted approximately 50 min and consisted of 600 trials (300 quick CSF and 300 conventional). The session comprised six consecutive quick CSF runs (300 = 6 runs × 50 trials each) interleaved with the conventional method that was used to estimate contrast sensitivities at individual spatial frequencies (300 = 50 trials × 6 frequency conditions).

Results

Validity of the 10AFC quick CSF procedure

In Figure 7, the CSFs measured in different quick CSF runs for all observers are plotted along with the CSF obtained by the conventional method for each individual. For the CSF measured in the third quick CSF run, the 68.2% HWCI of the estimated CSF is shown as shaded region, and for CSF measured by the conventional method, the 68.2% HWCI is represented as error bars. Inspection of the estimated CSFs from the quick CSF and Psi methods suggests excellent agreement.[3]

In order to show the agreement between the two methods, we compared the contrast sensitivities estimated with 10, 20, and 50 quick CSF trials to those measured by the conventional method with 300 trials (Figure 8). Sensitivities at six spatial frequency conditions (1.3, 2.2, 3.6, 5.9, 9.7, and 16.0 cpd) were used. CSFs from all observers and in all six quick CSF runs were pooled together. The Pearson correlation coefficient between the contrast thresholds obtained in the two methods was 0.952, 0.971, and 0.987 with 10, 20, and 50 quick CSF trials ($p < 0.001$ for all observers). In addition, all the data points are distributed along the unity line: the slope of the linear regression line is 1.0, 0.987, and 0.99 for CSFs obtained with 10, 20, and 50 quick CSF trials, respectively, which is not significantly different from 1.0. The results show excellent agreement between the quick CSF and conventional methods.

Comparison with the 2IFC quick CSF procedure

The average standard deviation and HWCI of the CSFs obtained from the 10AFC quick CSF procedure in this study and those from a 2IFC quick CSF procedure in a published paper (Hou et al., 2010) were computed. The bias of the CSFs from the 10AFC quick CSF procedure was also calculated. The results are shown in Figure 9.[4] In computing bias, CSFs obtained from the conventional method were used as the "true" values.

The average standard deviation of the CSFs obtained with the 10AFC quick CSF procedure was 0.15 ± 0.07

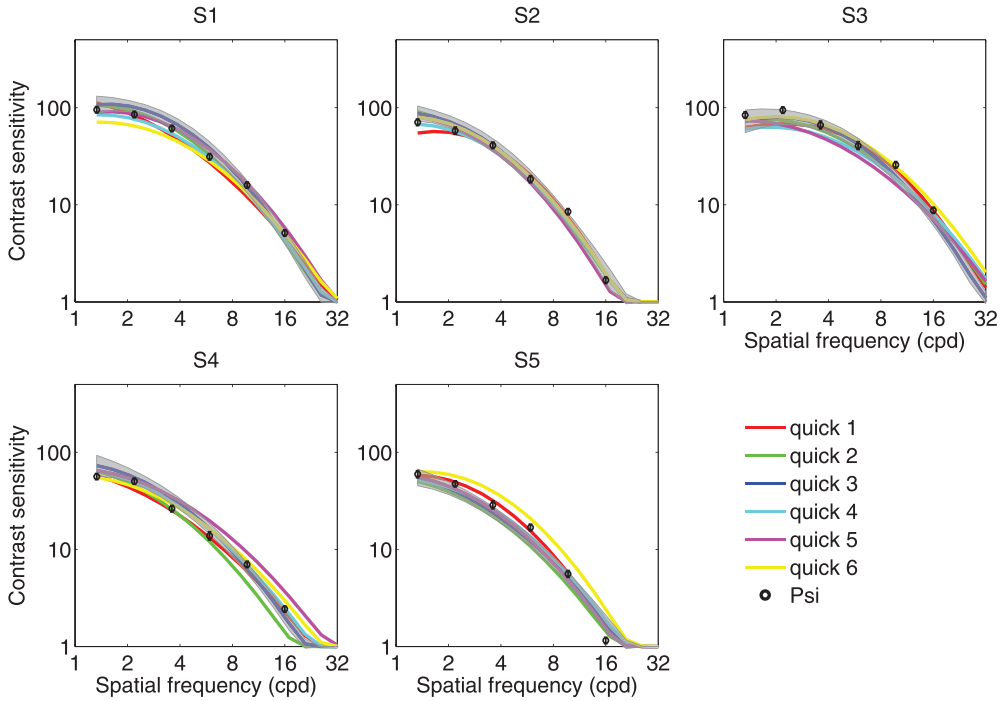

Figure 7. CSFs obtained from different quick CSF runs are plotted along with that measured by the conventional method for all observers. HWCI of 68.2% is indicated by shaded region for the CSF measured in the third quick CSF run, and by error bars for the CSF measured by the conventional method.

(mean ± *SD*), 0.10 ± 0.04, and 0.06 ± 0.03 log unit after 10, 20, and 50 trials, respectively. The average standard deviation of the CSFs obtained with the 10AFC quick CSF procedure is less than that of the 2IFC quick CSF with 50 trials when the trial number is greater than 8 ($p < 0.001$). The relative efficiency of the 10AFC quick CSF procedure was 214%, 274%, and 336% at 10, 20, and 50 trials, respectively, relative to the 2IFC quick CSF procedure. For CSFs obtained with the 10AFC procedure, the average HWCI was 0.20 ± 0.06, 0.12 ± 0.04, and 0.06 ± 0.02 log unit after 10, 20, and 50 trials, respectively. For CSFs obtained with the 2IFC procedure, the average HWCI was 0.22 ± 0.06 log unit after 50 trials. With about 10 trials, the HWCI of the CSFs from the 10AFC procedure became narrower than that of the CSFs from the 2IFC procedure with 50 trials ($p < 0.001$). For comparison, the average HWCI of CSFs from the conventional method with 300 trials is 0.023 ± 0.001 log unit. It took at least 150 trials of the conventional method to reach the same HWCI obtained by the 10AFC quick CSF procedure in 50 trials.

Finally, the standard deviation from repeated measures and HWCI from a single run became closer as trial number increased. After 50 trials, the ratio

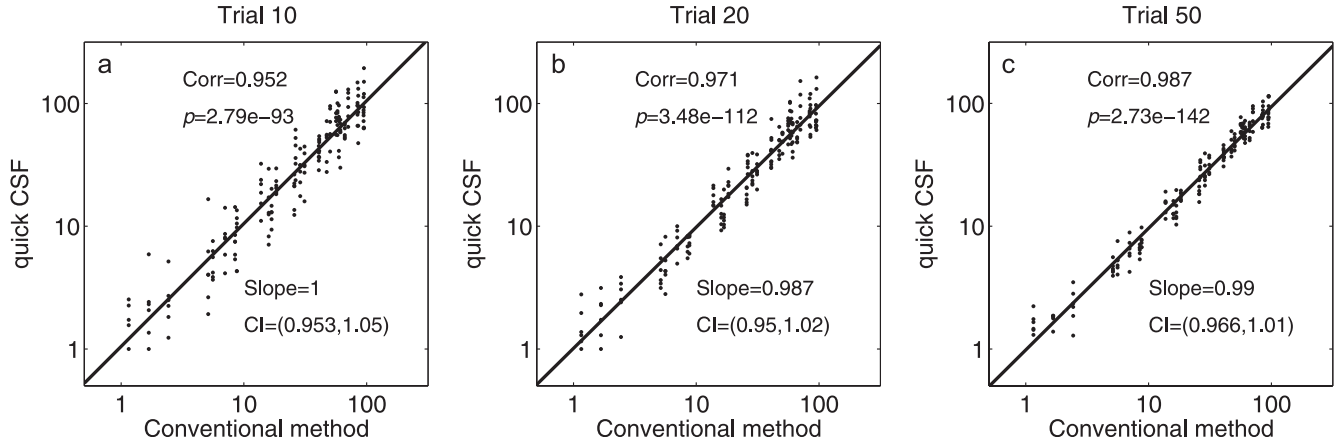

Figure 8. CSFs measured by the quick CSF procedure with 10, 20, and 50 trials were compared against those measured by the conventional method. Pearson correlation coefficients and linear regression fits are shown.

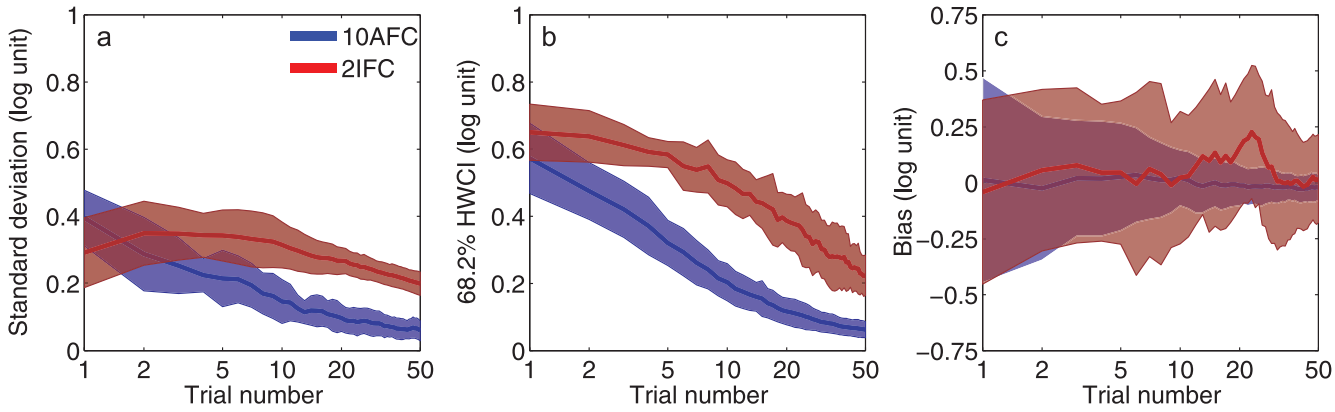

Figure 9. (a) Standard deviation, (b) HWCI, and (c) bias of the CS from the quick CSF procedure with 10AFC and 2IFC tasks as functions of trial number. Blue and red curves represent the results from the 10AFC and 2IFC quick CSF procedures, respectively. Shaded regions represent ±1 *SD*.

between the standard deviation and HWCI was 1.05 ± 0.05 for the 10AFC quick CSF procedure and 1.12 ± 0.05 for the 2IFC quick CSF procedure.

The average bias of the CSFs obtained from the 10AFC quick CSF procedure was 0.02 ± 0.12, −0.01 ± 0.09, and −0.02 ± 0.06 log unit after 10, 20, and 50 trials, respectively. The bias was not significantly different from 0 after only a few quick CSF trials ($p >$ 0.05 for all trials except trial 38 to 44). No consistent bias was found. The variance of the bias of the CSFs obtained from the 10 AFC quick CSF procedure decreased much faster than that from the 2IFC quick CSF procedure.

To quantify the test–retest reliability of the quick CSF method, we performed a Bland-Altman analysis (Bland & Altman, 1999) and calculated the coefficient of repeatability (COR), which describes the 95% confidence limits (2.77 × *SD*) for repeated measures. The average COR of CSFs obtained from the 2IFC quick CSF procedure were 1.38 ± 0.19, 1.07 ± 0.22, and 0.62 ± 0.17 log10 unit after 10, 20, and 50 trials, respectively. The average COR of CSFs obtained with the 10AFC quick CSF procedure were 0.41 ± 0.18, 0.27 ± 0.11, and 0.16 ± 0.08 log10 unit after 10, 20, and 50 trials, respectively. The CSFs obtained with the 10AFC quick CSF procedure exhibited much lower COR than those obtained with the 2IFC quick CSF procedure, indicating better test–retest reliability.

Summary and discussion

The goal of the current study was to further improve the efficiency of the quick CSF method (Lesmes et al., 2010). Inspired by studies in the literature (Alcalá-Quintana & García-Pérez, 2004; Hou et al., 2010; Jäkel & Wichmann, 2006; Leek et al., 1992; Taylor, 1971), we hypothesized that increasing the number of alternatives *m* in *m*-AFC tasks would improve the efficiency of the quick CSF method. The hypothesis was tested and confirmed in both computer simulations and a human psychophysics experiment.

Results from the simulation study showed that increasing *m* in an *m*-AFC task greatly improved the efficiency of the quick CSF procedure. With 50 trials, the relative efficiency of the quick CSF procedure with 4, 8, 10, and 16 AFC was 156%, 211%, 221%, and 255%, respectively, compared to that with a 2AFC task.

We further tested the hypothesis in a human psychophysics experiment. First we validated the newly designed 10AFC quick CSF procedure by comparing estimated CSFs directly to those obtained in a conventional method based on the Psi method (Klein, 2001; Kontsevich & Tyler, 1999). We showed that CSFs obtained from the 10AFC quick CSF method were in excellent agreement with those obtained from the conventional method. In addition, the quick CSF procedure exhibited high precision and excellent test-retest reliability. We compared the average standard deviation, credible interval, and bias of CSF obtained from the 10AFC quick CSF procedure with that from a 2IFC quick CSF procedure in a published study (Hou et al., 2010). The average standard deviation of the CSFs obtained with the10AFC quick CSF procedure in 10 trials was 0.15 ± 0.07 log unit and was less than that from the 2IFC procedure in 50 trials. The relative efficiency of the 10 AFC quick CSF procedure with 50 trials is 336% compared to that of a 2IFC quick CSF procedure. In addition, we showed that CSF variability estimated in two different ways—standard deviation from repeated measures and credible interval from a single run—generated closely matched estimates.

So far, we focused on the performance of the quick CSF method in measuring sensitivities. However, other important metrics related to different aspects of visual

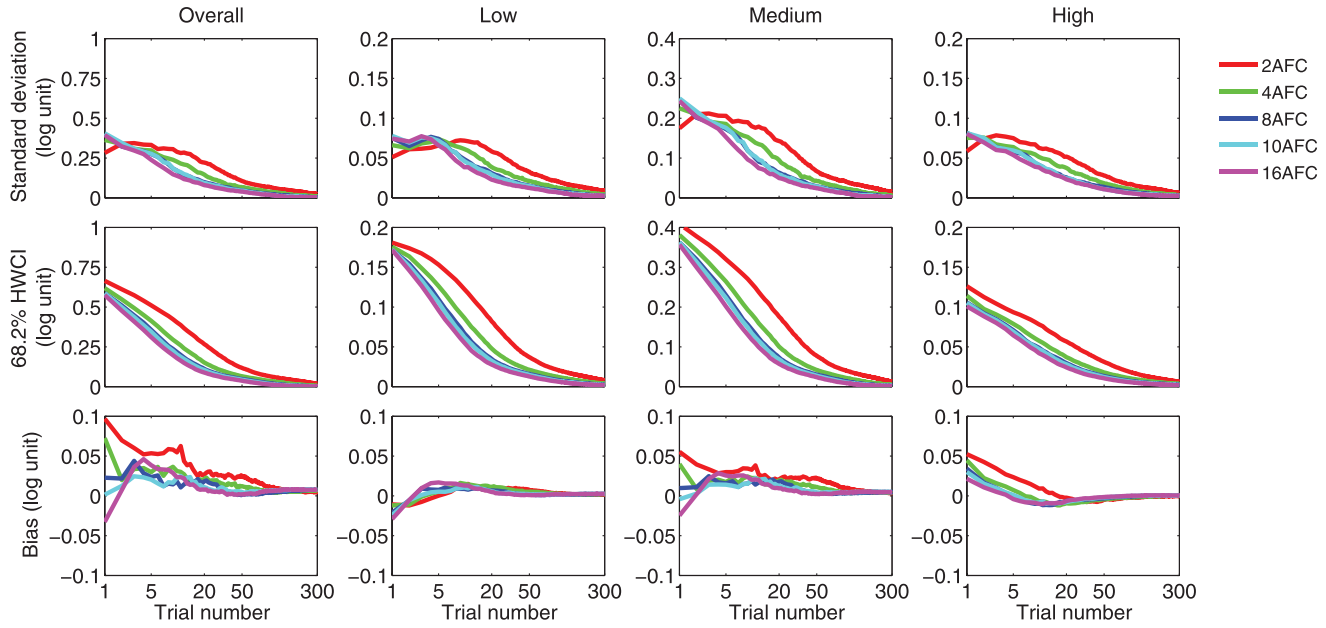

Figure 10. Standard deviation, HWCI, and bias as function of trial number for the overall, low, medium, and high frequency AULCSF. Different colors represent results from different *m*-AFC quick CSF procedures.

performance or particular characteristics of visual disease can be derived from the CSF. For example, it has been reported that the area under log CSF curve (AULCSF) is correlated with optical aberration of the human eye (Applegate, Howland, Sharp, Cottingham, & Yee, 1998; Oshika, Okamoto, Samejima, Tokunaga, & Miyata, 2006). The AULCSF between 3 to 12 cpd, called the subjective quality factor (SQF), has been used as an image quality indicator (Barten, 1999; Granger & Cupery, 1972). The AULCSF in different spatial frequency ranges may be a powerful metric to represent different aspects of visual performance. Patients may exhibit significant CSF deficits in specific spatial frequency bands due to different morphological/pathological characteristics (Huang et al., 2007; Midena, Degli Angeli, Blarzino, Valenti, & Segato, 1997; Regan, 1991). To explore the performance of the quick CSF method with different *m*-AFC task in estimating AULCSF, we calculated the standard deviation, HWCI and bias of measured AULCSF of our simulated observer in a number of different spatial frequency ranges. Because the most common spatial frequencies used in clinical testing are 1.5, 3, 6, 12, and 18 cpd (American National Standards Institute, 2001; Montes-Mico & Charman, 2001; Pesudovs, Hazel, Doran, & Elliott, 2004), the frequency ranges used in our analysis are low frequencies (1.5–3 cpd), medium frequencies (3–12 cpd), high frequencies (12–18 cpd), and overall frequencies (1.5–18 cpd).

As shown in the first column of Figure 10, the results for the overall AULCSF exhibited the same pattern as that for sensitivity (see Figure 2 for comparison). Increasing *m* in an *m*-AFC task improves the precision of the measurement. Similar results are also obtained for AULCSFs in low, medium, and high frequency ranges (see the second, third, and fourth columns of Figure 10). The precision also increases with trial number. With 50 trials, the standard deviation of the estimated CSF is 0.11, 0.07, 0.05, 0.05, and 0.04 log unit for the 2, 4, 8, 10, and 16 AFC tasks, respectively, and the relative efficiencies of the quick CSF procedure for AULCSF with 4, 8, 10, and 16 AFC tasks are 158%, 228%, 237%, and 275%, respectively. The standard deviation and HWCI are essentially the same after about 30 trials for all *m*-AFC tasks.

The precision of the estimated AULCSF depends on the spatial frequency range. After 50 quick CSF trials, the standard deviation of the overall, low, medium, and high frequency AULCSF for the 10AFC task were 0.05, 0.02, 0.03, and 0.01 log unit, respectively, and the HWCI of the overall, low, medium, and high frequency AULCSF for the 10AFC task were 0.04, 0.02, 0.03, and 0.01 log unit, respectively. The bias of the overall, low, medium, and high frequency AULCSF for the 10AFC task were 0.004, 0.002, 0.004, and −0.002 log unit at 50 trials.

Consistent with previous studies (Alcalá-Quintana & García-Pérez, 2004; Hou et al., 2010; Jäkel & Wichmann, 2006; Leek et al., 1992; Taylor, 1971), our results suggest that the shape of the psychometric function could have a profound impact on the efficiency of adaptive procedures that search optimal stimuli in a two-dimensional stimulus space. In a particular experimental setting, the slope of the d' psychometric function is related to the internal noise distribution and transducer of the observer (Lu & Dosher, 1998, 2008,

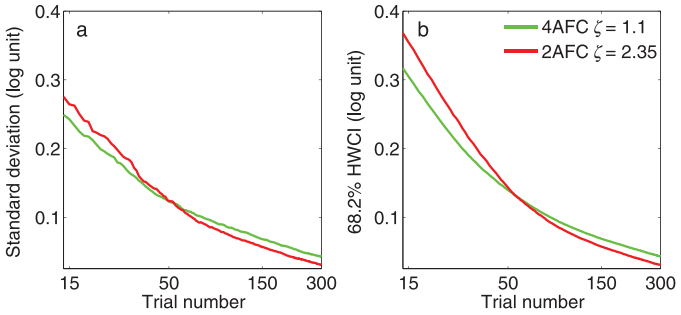

Figure 11. The standard deviation and HWCI the CSF from a 2AFC quick CSF procedure (red) and a 4AFC procedure (green). The slope of the d' psychometric function is $\zeta = 2.35$ in the 2AFC procedure and 1.1 in the 4AFC procedure.

2014; May & Solomon, 2013) and is not easy to manipulate. However, for a single d' psychometric function, it is possible to reduce the guessing rate and increase the slope of the percent correct psychometric function by increasing the number of alternatives in an m-AFC task, and therefore increase the efficiency of the adaptive procedure. The benefit of a larger number of alternatives in m-AFC tasks may not only apply to the quick CSF procedure, but also to other Bayesian adaptive testing procedures such as QUEST, ZEST, Psi, quick TvC, and quick Partial Report, all of which are based on some underlying parametric psychometric functions (Baek, Lesmes, & Lu, 2014; King-Smith et al., 1994; Kontsevich & Tyler, 1999; Kujala & Lukka, 2006; Lesmes et al., 2006; Lesmes et al., 2010; Watson & Pelli, 1983). It would be worthwhile to perform further studies to test the magnitude of improvements for those methods.

The relative efficiency of the 10AFC quick CSF procedure at 50 trials found in empirical experiment was higher than that found in simulation (336% vs. 221%). This is because the psychometric functions in the simulation and psychophysics experiment were different. The shape of the percent correct psychometric function is jointly determined by the underlying d' psychometric function and the task configuration. In a given experimental setting, increasing m in an m-AFC task will lead to reduced guessing rate and increased slope of the percent correct psychometric function. However, across experimental settings, the underlying d' psychometric function may be quite different (Eckstein, Abbey, & Bochud, 2000) and can affect the shape of the probability correct psychometric function (Equation 2). One can't simply look at the m in an m-AFC task to determine if a procedure is more efficient. Figure 11 shows the standard deviation and HWCI curves of CSF obtained by two quick CSF procedures, one is based on a 2AFC task with a slope ζ (the log-log slope of the d' psychometric function, Equation 1) of 2.35, and the other is based on a 4AFC task with a slope $\zeta = 1.1$. The precision of the 4AFC task with a shallower psychometric function is lower than that of the 2AFC task with a steeper psychometric function.

It should be noted that increasing the number of alternatives in forced choice tasks may introduce some complications that, if not dealt with properly, may adversely affect CSF measurements. For example, increasing the number of alternatives may generate higher working memory or attention demand. Jäkel and Wichmann (2006) found that in sinusoidal grating contrast discrimination, an 8AFC task led to worse performance than a 4AFC task. They suggested that observers may have to attend to more spatial locations in the 8AFC task because the eight alternative stimuli were distributed spatially. The 10AFC quick CSF procedure uses an identification task in which a single letter stimulus is presented to the subject at a single spatial location in a single temporal interval in each trial; plus, all 10 letters are well known to the subjects, minimizing attention and working memory demands.

The response screen used to facilitate response collection may introduce side effects. We have carefully chosen the display parameters to minimize potential aftereffects. First, the response screen was shown 500 ms after stimuli presentation. This interval is much longer than the typical temporal integration window of the human visual system (Breitmeyer, 1984; Lu, Jeon, & Dosher, 2004). The response screen was also presented in a Weber contrast of −0.4 (letter was darker than background), which, if converted into RMS contrast, matched the average RMS contrast of letter stimuli in our experiment, so that it would not disrupt the adaption status of our observers in the experiment.

Another potential complication with the use of the letter stimuli is that the letter stimuli in the same stimulus conditions (e.g., same spatial frequency and contrast) may have unequal visibility (Alexander et al., 1994). To reduce this problem, we normalized the contrast of the filtered letters to their individual maximum absolute intensity such that, after normalization, the maximum absolute Weber contrast of the image is 1.0. The RMS contrasts of the 10 letters after normalization were essentially the same (0.115 ± 0.010). In addition, a letter can sometimes be confused with other letters in subject's responses (Mueller & Weidemann, 2012), violating the assumption that all alternatives are orthogonal and equivalent (Hacker & Ratcliff, 1979) in the simple formulation of the psychometric function of m-AFC tasks (Equation 2). By applying the same filter to all the letters, we have restricted the spatial content of all the stimuli to a two-octave range and reduced the difference in spatial content among letters. This could significantly reduce the probability of confusion. In fact, Gervais, Harvey, and Roberts (1984) found that the differences in spatial frequency content of letters provided the best predic-

tion of the confusion matrix in letter recognition. Furthermore, there is evidence that CSFs measured with narrowband letters was very similar to those measured with sinewave gratings or D6 patterns (Alexander et al., 1994; Hou & Lu, 2014; McAnany & Alexander, 2006).

Increasing the number of alternatives in *m*-AFC tasks is not the only way to increase the efficiency of quick CSF. Kim et al. (2014) have recently developed a hierarchical adaptive design optimization (HADO) procedure that achieves greater accuracy and efficiency in adaptive information gain by exploiting two complementary schemes of inference with past and future data. HADO extends the standalone quick CSF method to a framework that models a higher-level structure across the population, which can be used as an informative prior for each new assessment. In turn, the parameter estimates from each individual enable the update of the higher-level structure. The judicious application of informative priors used by HADO improves the efficiency of the quick CSF method by approximately 30%.

In summary, the quick CSF method has been validated and applied in a range of experiments (Dorr et al., 2013; Gepshtein et al., 2013; Hou et al., 2010; Kalia et al., 2014; Lee et al., 2014; Reynaud et al., 2014; Rosén et al., 2014). In the current study, we show that the efficiency of the quick CSF method can be further improved by increasing the number of alternatives in multialternative forced choice tasks used in the procedure. Specifically, the quick CSF method utilizing the 10-letter identification task can be used to estimate a CSF with a 0.1 log unit standard deviation in about 20 trials, or less than 2 min.

Keywords: contrast sensitivity function, psychometric function, 2AFC, efficiency, precision, quick CSF

Acknowledgments

We would like to express our gratitude to Dr. Woojae Kim for the discussion on precision in the simulation study. We also thank two anonymous reviewers and the editor for their constructive comments. This research was supported by the National Eye Institute (EY021553 to Z-LL) and by the Elite Network Bavaria (to MD).

Commercial relationships: LAL, PB, MD, and Z-LL have an intellectual property interest in methods for measuring contrast sensitivity (US 7938538, WO2013170091 A1), and equity interest in a company (Adaptive Sensory Technology, MA) commercializing the technology. LAL holds employment in Adaptive Sensory Technology.

Corresponding author: Zhong-Lin Lu.
Email: lu.535@osu.edu.
Address: Laboratory of Brain Processes, The Ohio State University, Columbus, OH, USA.

Footnotes

[1] We normalized Equation 1 by $d' = 1.5$ because the true values of the CSF parameters $\theta^{\text{true}} = (g_{max}^{true}, f_{max}^{true}, \beta^{\text{true}}, \delta^{\text{true}}) = (80, 1.07, 3.6, 0.3)$ were chosen based on empirical data measured with a d' of about 1.5.

[2] We assume that the observer would make a random guess when she is in lapse. The formulation is slightly different from Klein (2001) and Wichmann and Hill (2001), which did not consider guessing in lapse trials. When λ is low (i.e., 0.04), there is no significant difference between the two definitions.

[3] The low frequency truncation is not apparent in these CSFs because of the range of spatial frequencies covered in this study (1.33–32 cpd) is relative high compared to the 0.5–16 cpd range used in Hou et al. 2010. The low frequency truncation parameter δ is necessary in CSF tests that include lower spatial frequency conditions.

[4] The standard deviation reported in Hou et al., (2010) was from a bootstrap procedure because they only had two repeated measures of each CSF.

References

Alcalá-Quintana, R., & García-Pérez, M. (2004). The role of parametric assumptions in adaptive Bayesian estimation. *Psychological Methods, 9*(2), 250–271.

Alexander, K. R., Xie, W., & Derlacki, D. J. (1994). Spatial-frequency characteristics of letter identification. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 11*, 2375–2382.

American National Standards Institute. (2001). *American National Standard for Ophthalmics–Multifocal Intraocular Lenses*. Fairfax, VA: Optical Laboratories Association.

Applegate, R. A., Howland, H. C., Sharp, R. P., Cottingham, A. J., & Yee, R. W. (1998). Corneal aberrations and visual performance after radial keratotomy. *Journal of Refractive Surgery, 14*, 397–407.

Arden, G. B., & Jacobson, J. J. (1978). A simple grating test for contrast sensitivity: Preliminary results indicate value in screening for glaucoma. *Investi-*

gative Ophthalmology & Visual Science, 17(1), 23–32. [PubMed] [Article]

Baek, J., Lesmes, L., & Lu, Z. L. (2014). *Bayesian adaptive estimation of the sensory memory decay function: the quick Partial Report method.* Paper presented at the 2014 Annual meeting of the Vision Science Society, St. Pete Beach, FL.

Barten, P. G. J. (1999). *Contrast sensitivity of the human eye and its effects on image quality*. Knegsel, The Netherlands: HV Press.

Bi, J., Lee, H.-S., & O'Mahony, M. (2010). d' and variance of d' for four-alternative forced choice (4-AFC). *Journal of Sensory Studies, 25*, 740–750.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8*, 135–160.

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. New York: Oxford University Press.

Chung, S. T., Legge, G. E., & Tjan, B. S. (2002). Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vision Research, 42*, 2137–2152.

Clayton, D., & Hills, M. (1993). *Statistical models in epidemiology*. Oxford, UK: Oxford University Press.

Cobo-Lewis, A. B. (1996). *An adaptive method for estimating multiple parameters of a psychometric function*. Paper presented at the 29th Annual Meeting of the Society for Mathematical Psychology, Chapel Hill, NC.

Dorr, M., Lesmes, L. A., Lu, Z. L., & Bex, P. J. (2013). Rapid and reliable assessment of the contrast sensitivity function on an iPad. *Investigative Ophthalmology & Visual Science, 54*(12), 7266–7273. [PubMed] [Article]

Eckstein, M. P., Abbey, C. K., & Bochud, F. O. (2000). Visual signal detection in structured backgrounds. IV. Figures of merit for model performance in multiple-alternative forced-choice detection tasks with correlated responses. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 17*, 206–217.

Foley, J. M., & Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research, 21*, 1041–1053.

Gepshtein, S., Lesmes, L. A., & Albright, T. D. (2013). Sensory adaptation as optimal resource allocation. *Proceedings of the National Academy of Sciences, USA, 110*, 4368–4373.

Gervais, M. J., Harvey, L. O., Jr., & Roberts, J. O. (1984). Identification confusions among letters of the alphabet. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 655–666.

Ginsburg, A. P. (1981). Spatial filtering and vision: Implications for normal and abnormal vision. In L. Proenz, J. Enoch & A. Jampolsky (Eds.), *Clinical applications of visual psychophysics*. (pp. 70–106). Cambridge, UK: Cambridge University Press.

Ginsburg, A. P. (2003). Contrast sensitivity and functional vision. *International Ophthalmology Clinics, 43*(2), 5–15.

Granger, E. M., & Cupery, K. N. (1972). An optical merit function (SQF), which correlates with subjective image judgments. *Photographic Science and Engineering, 16*, 221–230.

Hacker, M., & Ratcliff, R. (1979). A revised table of d' for M-alternative forced choice. *Perception & Psychophysics, 26*, 168–170.

Hall, J. L. (1983). A procedure for detecting variability of psychophysical thresholds. *The Journal of the Acoustical Society of America, 73*, 663–667.

Hess, R. F. (1981). Application of contrast-sensitivity techniques to the study of functional amblyopia. In L. Proenz, J. Enoch, & A. Jampolsky (Eds.), *Clinical applications of visual psychophysics*. (pp. 11–41). Cambridge, UK: Cambridge University Press.

Hess, R. F., & Howell, E. R. (1977). The threshold contrast sensitivity function in strabismic amblyopia: Evidence for a two type classification. *Vision Research, 17*, 1049–1055.

Hou, F., Huang, C. B., Lesmes, L., Feng, L. X., Tao, L., Zhou, Y. F., & Lu, Z. L. (2010). qCSF in clinical application: Efficient characterization and classification of contrast sensitivity functions in amblyopia. *Investigative Ophthalmology & Visual Science, 51*, 5365–5377. [PubMed] [Article]

Hou, F., & Lu, Z. L. (2014). *Comparisons of contrast sensitivity functions measured by two different qCSF implements. Investigative Ophthalmology and Visual Science, 55*, E-abstract 770.

Hou, F., Lu, Z. L., & Huang, C. B. (2014). The external noise normalized gain profile of spatial vision. *Journal of Vision, 14*(13):9, 1–14, doi:10.1167/14.13.9. [PubMed] [Article]

Huang, C., Tao, L., Zhou, Y., & Lu, Z. L. (2007). Treated amblyopes remain deficient in spatial vision: A contrast sensitivity and external noise study. *Vision Research, 47*, 22–34.

Jäkel, F., & Wichmann, F. A. (2006). Spatial four-alternative forced-choice method is the preferred psychophysical method for naive observers. *Journal of Vision, 6*(11):13, 1307–1322, doi:10.1167/6.11.13. [PubMed] [Article]

Jindra, L. F., & Zemon, V. (1989). Contrast sensitivity testing: a more complete assessment of vision. *Journal of Cataract & Refractive Surgery, 15*, 141–148.

Kalia, A., Lesmes, L. A., Dorr, M., Gandhi, T., Chatterjee, G., Ganesh, S. ... Sinha, P. (2014). Development of pattern vision following early and extended blindness. *Proceedings of the National Academy of Sciences, USA, 111*, 2035–2039.

Kelly, D. H., & Savoie, R. E. (1973). A study of sine-wave contrast sensitivity by two psychophysical methods. *Perception & Psychophysics, 14*, 313–318.

Kim, W., Pitt, M. A., Lu, Z. L., Steyvers, M., & Myung, J. I. (2014). A hierarchical adaptive approach to optimal experimental design. *Neural Computation, 26*, 2465–2492.

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research, 34*, 885–912.

King-Smith, P. E., & Rose, D. (1997). Principles of an adaptive method for measuring the slope of the psychometric function. *Vision Research, 37*, 1595–1604.

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception and Psychophysics, 63*, 1421–1455.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? [Abstract] *Perception, 36,* 14.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research, 39*, 2729–2737.

Kujala, J. V., & Lukka, T. J. (2006). Bayesian adaptive estimation: The next dimension. *Journal of Mathematical Psychology, 50*, 369–389.

Lee, T. H., Baek, J., Lu, Z. L., & Mather, M. (2014). How arousal modulates the visual contrast sensitivity function. *Emotion, 14*, 978–984.

Leek, M. R., Hanna, T. E., & Marshall, L. (1992). Estimation of psychometric functions from adaptive tracking procedures. *Perception and Psychophysics, 51*, 247–256.

Legge, G. E., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 4*, 391–404.

Lesmes, L., Wallis, J., Jackson, M. L., & Bex, P. (2013). The reliability of the quick CSF method for contrast sensitivity assessment in low vision. *Investigative Ophthalmology and Visual Science, 54*, E-abstract 2762.

Lesmes, L., Wallis, J., Lu, Z.-L., Jackson, M. L., & Bex, P. J. (2012). Clinical application of a novel contrast sensitivity test to a low vision population: The quick CSF method. *Investigative Ophthalmology and Visual Science, 53*, E-abstract 4358.

Lesmes, L. A., Jeon, S. T., Lu, Z. L., & Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick TvC method. *Vision Research, 46*, 3160–3176.

Lesmes, L. A., Lu, Z.-L., Baek, J., & Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: The quick CSF method. *Journal of Vision, 10*(3):17, 1–21, doi:10.1167/10.3.17. [PubMed] [Article]

Li, X., & Lu, Z. L. (2012). Enabling high grayscale resolution displays and accurate response time measurements on conventional computers. *Journal of Visualized Experiments, 60*, e3312, doi:10.3791/3312.

Li, X., Lu, Z. L., Xu, P., Jin, J., & Zhou, Y. (2003). Generating high gray-level resolution monochrome displays with conventional computer graphics cards and color monitors. *Journal of Neuroscience Methods, 130*, 9–18.

Lu, Z. L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*, 1183–1198.

Lu, Z. L., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 16*, 764–778.

Lu, Z. L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review, 115*(1), 44–82.

Lu, Z. L., & Dosher, B. A. (2014). *Visual psychophysics: From laboratory to theory*. Cambridge, MA: The MIT Press.

Lu, Z. L., Jeon, S. T., & Dosher, B. A. (2004). Temporal tuning characteristics of the perceptual template and endogenous cuing of spatial attention. *Vision Research, 44*, 1333–1350.

Marmor, M. F. (1981). Contrast sensitivity and retinal disease. *Annals of Ophthalmology, 13*, 1069–1071.

Marmor, M. F. (1986). Contrast sensitivity versus visual acuity in retinal disease. *The British Journal of Ophthalmology, 70*, 553–559.

Marmor, M. F., & Gawande, A. (1988). Effect of visual blur on contrast sensitivity. Clinical implications. *Ophthalmology, 95*, 139–143.

May, K. A., & Solomon, J. A. (2013). Four theorems on the psychometric function. *PLoS One, 8*(10), e74815.

McAnany, J. J., & Alexander, K. R. (2006). Contrast sensitivity for letter optotypes vs. gratings under conditions biased toward parvocellular and mag-

nocellular pathways. *Vision Research, 46*, 1574–1584.

Midena, E., Degli Angeli, C., Blarzino, M. C., Valenti, M., & Segato, T. (1997). Macular function impairment in eyes with early age-related macular degeneration. *Investigative Ophthalmology & Visual Science, 38*(2), 469–477. [PubMed] [Article]

Montes-Mico, R., & Charman, W. N. (2001). Choice of spatial frequency for contrast sensitivity evaluation after corneal refractive surgery. *Journal of Refractive Surgery, 17*, 646–651.

Montes-Mico, R., & Ferrer-Blasco, T. (2001). Contrast sensitivity function in children: Normalized notation for the assessment and diagnosis of diseases. *Documenta Ophthalmologica, 103*(3), 175–186.

Mueller, S. T., & Weidemann, C. T. (2012). Alphabetic letter identification: Effects of perceivability, similarity, and bias. *Acta Psychologica (Amst), 139*, 19–37.

NAS-NRC Committee on Vision, W. G. M. (1981). Recommended standard procedures for the clinical measurement and specification of visual acuity: Report of working group 39. *Advances in Ophthalmology, 41,* 103–148.

Onal, S., Yenice, O., Cakir, S., & Temel, A. (2008). FACT contrast sensitivity as a diagnostic tool in glaucoma: FACT contrast sensitivity in glaucoma. *International Ophthalmology, 28*, 407–412.

Oshika, T., Okamoto, C., Samejima, T., Tokunaga, T., & Miyata, K. (2006). Contrast sensitivity function and ocular higher-order wavefront aberrations in normal human eyes. *Ophthalmology, 113*, 1807–1812.

Patterson, V. H., Foster, D. H., & Heron, J. R. (1980). Variability of visual threshold in multiple sclerosis: Effect of background luminance on frequency of seeing. *Brain, 103*, 139–147.

Pesudovs, K., Hazel, C. A., Doran, R. M., & Elliott, D. B. (2004). The usefulness of Vistech and FACT contrast sensitivity charts for cataract and refractive surgery outcomes research. *British Journal of Ophthalmology, 88*, 11–16.

Ramulu, P., Dave, P., & Friedman, D. (2015). Precision of contrast sensitivity testing in glaucoma. *Investigative Ophthalmology and Visual Science, 56,* E-abstract 2225.

Regan, D. (1991). Spatiotemporal abnormalities of vision in patients with multiple sclerosis. In D. Regan (Ed.), *Spatial vision* (pp. 239–249). Boca Raton, FL: CRC Press.

Reynaud, A., Tang, Y., Zhou, Y., & Hess, R. (2014). A unified framework and normative dataset for second-order sensitivity using the quick Contrast Sensitivity Function (qCSF). *Journal of Vision, 14*(10): 1428, doi:10.1167/14.10.1428. [Abstract]

Richman, J., Lorenzana, L. L., Lankaranian, D., Dugar, J., Mayer, J., Wizov, S. S., & Spaeth, G. L. (2010). Importance of visual acuity and contrast sensitivity in patients with glaucoma. *Archives of Ophthalmology, 128*, 1576–1582.

Rosen, R., Jayaraj, J., Bharadwaj, S. R., Weeber, H. A., Van der Mooren, M., & Piers, P. A. (2015). Contrast sensitivity in patients with macular degeneration. *ARVO Investigative Ophthalmology and Visual Science, 56*, E-abstract 2224.

Rosén, R., Lundström, L., Venkataraman, A. P., Winter, S., & Unsbo, P. (2014). Quick contrast sensitivity measurements in the periphery. *Journal of Vision, 14*(8):3, 1–10, doi:10.1167/14.8.3. [PubMed] [Article]

Shandiz, J. H., Nourian, A., Hossaini, M. B., Moghaddam, H. O., Yekta, A. A., Sharifzadeh, L., Marouzi, P. (2010). Contrast sensitivity versus visual evoked potentials in multiple sclerosis. *Journal of Ophthalmic & Vision Research, 5*, 175–181.

Shelton, B. R., & Scarrow, I. (1984). Two-alternative versus three-alternative procedures for threshold estimation. *Perception & Psychophysics, 35*, 385–392.

Sloan, L., Rowland, W. M., & Altman, A. (1952). Comparison of three types of test target for the measurement of visual acuity. *Quarterly Review of Ophthalmology and Otorhinolaryngology, 8,* 4–17.

Taylor, M. M. (1971). On the efficiency of psychophysical measurement. *The Journal of the Acoustical Society of America, 49*(2), Suppl 2, 505–508.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America, 41*(4A), 782–787.

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research, 35*, 2503–2522.

Watson, A. B., & Ahumada, A. J., Jr. (2005). A standard model for foveal detection of spatial contrast. *Journal of Vision, 5*(9):6, 717–740, doi:10.1167/5.9.6. [PubMed] [Article]

Watson, A. B., & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*, 113–120.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*, 1293–1313.

Woods, R. L., & Wood, J. M. (1995). The role of contrast sensitivity charts and contrast letter charts in clinical practice. *Clinical and Experimental Optometry, 78*, 43–57.

Yenice, O., Onal, S., Incili, B., Temel, A., Afsar, N., & Tanridag, T. (2007). Assessment of spatial-contrast function and short-wavelength sensitivity deficits in patients with migraine. *Eye, 21*, 218–223.

Appendix A: The quick CSF algorithm

The quick CSF method uses the following steps to estimate $\tau(f)$ (Lesmes et al., 2010):

1. Define a CSF functional form. $\tau(f)$ is the reciprocal of contrast sensitivity $S(f)$:

$$\tau(f) = \frac{1}{S(f,\theta)},$$

which is described by the truncated log parabola with four parameters (Lesmes et al., 2010; Watson & Ahumada, 2005):

$$\log_{10}\Big(S(f,\theta)\Big) =$$

$$\begin{cases} \log_{10}(g_{\max}) - \delta, f < f_{\max} \wedge S_0 < \log_{10}(g_{\max}) - \delta \\ \log_{10}(g_{\max}) - \dfrac{4}{\log_{10}(2)}\left(\dfrac{\log_{10}(f) - \log_{10}(f_{max})}{\beta}\right)^2, f > f_{max} \end{cases}, \quad \text{(A1)}$$

where $\theta = (g_{\max}, f_{\max}, \beta, \delta)$ represents the four CSF parameters: peak gain $g_{\max}$, peak spatial frequency $f_{\max}$, bandwidth at half-height β (in octaves), and low-frequency truncation level δ.

2. Define the stimulus and parameter spaces. The application of Bayesian adaptive inference requires two basic components: (a) a prior probability distribution, $p(\theta)$, defined over a four-dimensional space of CSF parameters θ, and (b) a two-dimensional space of possible letter stimuli with contrast c and spatial frequency f.

In our simulation study, the ranges of possible CSF parameters were: 2–2000 for peak gain, 0.2–20 cycles per degree (cpd) for peak frequency, 1–9 octaves for bandwidth, and 0.02–2 for truncation. The ranges for possible grating stimuli were 0.1%–100% for contrast c and 0.5–32 cpd for frequency f. Both parameter and stimuli spaces were sampled evenly in log unit.

3. Priors. Before the beginning of the experiment, an initial prior, $p_{t=0}(\theta)$, which represents the knowledge about the observer's CSF before any data is collected, was defined by a hyperbolic secant function with the best guess of parameters $\theta_{i,\ guess}$ and width of $\theta_{i,\ \text{confidence}}$ for $i = 1$, 2, 3, and 4 (King-Smith & Rose, 1997; Lesmes et al., 2010).

$$p_{t=0}(\theta) =$$

$$\prod_{i=1}^{4} \operatorname{sech}(\theta_{i,\text{confidence}} \times (\log_{10}(\theta_i) - \log_{10}(\theta_{i,\text{guess}}))), \quad \text{(A2)}$$

where $\operatorname{sech}(x) = \frac{2}{e^x + e^{-x}}$, $\theta_i = g_{\max}, f_{\max}, \beta$ and δ for $i = 1$, 2, 3, and 4, respectively, $\theta_{i,\ \text{guess}} = 100$, 2, 3, and 0.5 for $i = 1$, 2, 3, and 4, respectively. $\theta_{i,\ \text{confidence}.} = 2.48$, 3.75, 7.8, and 3.12 for $i = 1$, 2, 3, and 4, respectively.

4. Bayesian adaptive inference. After subject's response is collected in trial t, knowledge about CSF parameters $p(\theta)$ is updated, given the evidence provided by the observer's response $r_x =$ "correct" or "incorrect" to the stimulus $x = (c, f)$ with contrast c and spatial frequency f in the trial. The outcome of trial t is incorporated into a Bayesian inference step that updates the prior knowledge about CSF parameters $p_{t\text{-}1}(\theta)$,

$$p_t(\theta) = p_t(\theta |\ r_x) = \frac{p_{t-1}(\theta)p(r_x|\theta)}{\sum_\theta p_{t-1}(\theta)p(r_x|\theta),} \quad \text{(A3)}$$

where $p_t(\theta \mid r_x)$ is the posterior distribution of parameter vector θ after obtaining a response r_x at trial t; $p(r_x =$ correct $\mid \theta) = \Psi(x,\theta)$ is the percent correct psychometric function given stimulus x, and $p(r_x =$ incorrect $\mid \theta) = 1 - \Psi(x,\theta)$; $p_{t-1}(\theta)$ is our prior about θ before trial t, which is also the posterior in trial $t-1$.

5. Stimulus search. To increase the quality of the evidence obtained on each trial, the quick CSF calculates the expected information gain for all possible stimuli x,

$$I_t(\theta; r_x) = h\left(\int p_t(\theta)\Psi(x,\theta)d\theta\right) - \int p_t(\theta)h\Big(\Psi(x,\theta)\Big)d\theta, \quad \text{(A4)}$$

where $h(p) = -p\log(p) - (1-p)\log(1-p)$ is the information entropy of the distribution p. Before each trial, we find out the candidate stimuli that correspond to the top 10% of the expected information gain over the entire stimulus space. Then we randomly pick one among those candidates as x_t for presentation. In this way, the quick CSF avoids large regions of the stimulus space that are not likely to provide useful information to the current knowledge about θ.

6. Reiteration and stopping rule. The procedure reiterates steps 4 and 5 until 300 trials are run.

7. Analysis. After step 6, we obtain the posterior distribution of CSF parameters $p_t(\theta)$ (see Figure A1 for the marginal prior and posterior distributions for the four CSF parameters). A resampling procedure is used that samples directly from the posterior distributions of the CSF parameters and generates the CSF estimates based on all the CSF samples. The procedure automatically takes into account the covariance structure of the CSF parameters in the posterior distribution

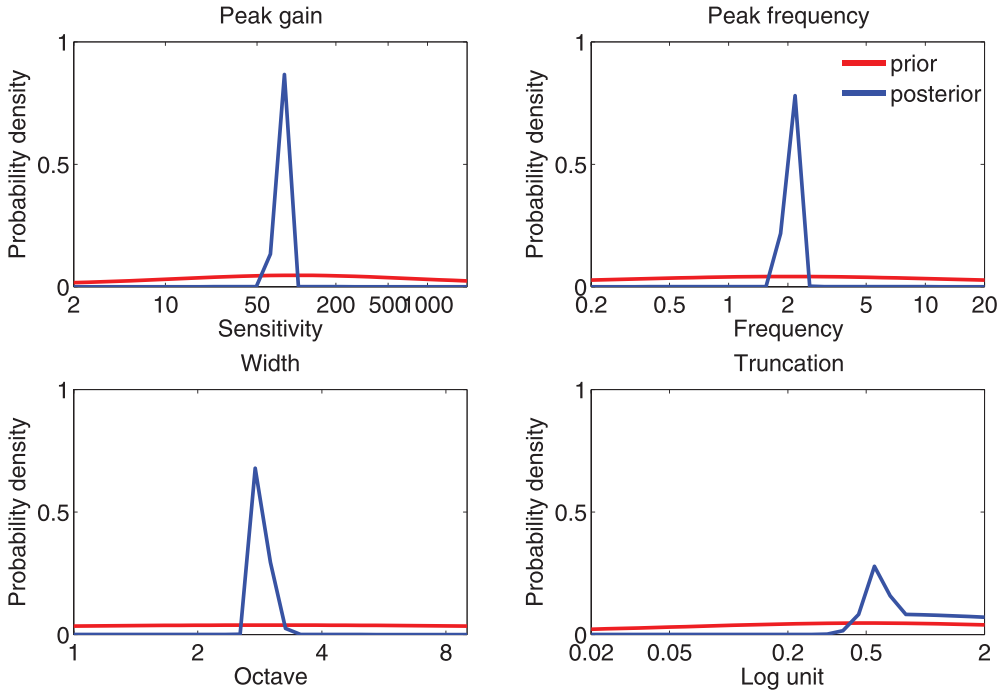

Figure A1. An illustration of the marginal distributions of the four parameters before (prior: red) and after (posterior: blue) measurement. The plot is based on the simulation of a single quick CSF run with 100 trials.

and allows us to compute the credible interval of the estimated CSF functions.

Appendix B: Using the Weibull function to approximate the d' psychometric function

The computational load of integration in Equation 2 is very heavy. In order to facilitate the computation, a Weibull function

$$P_w(c,f,m) = \gamma(m) + \left(1-\gamma(m)\right)\left(1-\exp\left(-\left(\frac{c}{\tau_w(f,m)}\right)^{b(m)}\right)\right) \quad \text{(B1)}$$

was used to approximate the psychometric functions in m-AFC tasks (Equation 2). The approximation made the simulation about 20 times faster. The terms $\gamma(m)$ and $b(m)$ are the guessing rate and slope of the Weibull psychometric function in an m-AFC task, and $\tau_w(f,m)$ is the Weibull contrast threshold in spatial frequency condition f in an m-AFC task, which can be computed from $\tau(f)$.

$$\log_{10}\left(\tau_w(f,m)\right) = \log_{10}\left(\tau(f)\right) - \frac{1}{b(m)}\log_{10}\left(\log\left(\frac{1-\gamma(m)}{1-p_{1.5}(m)}\right)\right), \quad \text{(B2)}$$

where $p_{1.5}(m)$ is the fraction of correct responses corresponding to $d' = 1.5$ in an m-AFC task. $\gamma(m)$ and $p_{1.5}(m)$ are listed in Table B1 for a range of m values used in our simulation study.

Equation B1 was fit to the psychometric functions described in Equation 2 with $m = 2, 4, 8, 10$, and 16 and $\gamma = 1/m$. The Weibull provided an excellent approximation to the psychometric functions in Equation 2 with an average $r^2 = 0.999$. The best fitting $b(m)$s are listed in Table B1. With predetermined $\tau(f)$, b, γ, and λ, the response probabilities of the simulated observer in the m-AFC tasks in all possible stimulus conditions can be computed.

Number of alternatives (m)	γ	b	$p_{1.5}$
2	0.5	3.06	0.856
4	0.25	3.45	0.702
8	0.125	3.90	0.553
10	0.1	4.05	0.509
16	0.0625	4.39	0.421

Table B1. The values of b and γ for different m-AFC tasks.