Vision Research 154 (2019) 21–43

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Efficient assessment of the time course of perceptual sensitivity change

Yukai Zhao[a], Luis Lesmes[b], Zhong-Lin Lu[a,*]

[a] Department of Psychology, The Ohio State University, Columbus, OH, USA
[b] Adaptive Sensory Technology Inc., San Diego, CA, USA

ARTICLE INFO

Keywords:
Adaptive procedure
Dark adaptation
Bayesian inference
Information gain
Time course of perceptual sensitivity

ABSTRACT

Perceptual sensitivity is usually estimated over trials and time intervals, which results in imprecise and biased estimates when it changes rapidly over time. We develop a novel procedure, the quick Change-Detection (qCD) method, to accurately, precisely, and efficiently assess the trial-by-trial time course of perceptual sensitivity change. Based on Bayesian adaptive testing, qCD selects the optimal stimulus, and updates, trial by trial, a joint probability distribution of the parameters that quantify perceptual sensitivity change over time. We demonstrate the utility of the method in measuring the time course of dark adaptation. Simulations showed that the accuracy and precision of the estimated dark adaptation curve after *one* qCD run (root mean squared error (RMSE): 0.002; the half width of the 68.2% credible interval (HWCI): 0.016; standard deviation (SD): 0.020; all in $\log_{10}$ units) was higher than those obtained by *ten* runs of the quick Forced-Choice (qFC) procedure (RMSE: 0.020; HWCI: 0.032; SD: 0.031) and *ten* runs of a weighted up-down staircase procedure (RMSE: 0.026; SD: 0.031). Further, the dark adaptation curve obtained from one qCD run in a psychophysics experiment was highly consistent with the average of four qFC runs (RMSE = 0.076 $\log_{10}$ units). Overall, qCD provides a procedure to characterize the detailed time course of perceptual sensitivity change in both basic research and clinical applications.

1. Introduction

Perceptual sensitivity, defined as 1/threshold, varies as a function of stimulus characteristics such as background luminance (Hou et al., 2016; Sloane, Owsley, & Jackson, 1988), spatial frequency (Campbell, 1983; Campbell & Robson, 1968), temporal frequency (Robson, 1966), and retinal eccentricity (Rovamo, Franssila, & Nasanen 1992; Strasburger, Rentschler, & Jüttner, 2011). It also changes with observer state over a wide range of time scales, from seconds to minutes in dark adaptation (Lamb & Pugh, 2004; Reuter, 2011; Winsor & Clark, 1936) and attention (Cheal & Lyon, 1991; Lu & Dosher, 1998; Nakayama & Mackeben, 1989) to days and beyond in perceptual learning (Dosher & Lu, 2017; Sagi, 2011), visual development and aging (Braddick & Atkinson, 2011; Jackson, Owsley, & McGwin, 1999; Monge & Madden, 2016; Owsley, Sekuler, & Siemsen, 1983; Tahir et al., 2017), and disease progression (Jackson & Edwards, 2008; Jackson, Felix, & Owsley, 2006; Owsley et al., 2016) and intervention (Levi & Li, 2009; Maniglia, Cottereau, Soler, & Trotter, 2016). Precise and efficient measurement of the time course of perceptual sensitivity change is not only essential for understanding basic properties of the visual system (Campbell, 1983; Campbell & Robson, 1968; Cheal & Lyon, 1991; Hou et al., 2016; Lamb & Pugh, 2004; Reuter, 2011; Sagi, 2011; Sloane et al., 1988; Winsor & Clark, 1936), but can also provide critical diagnostic information for a number of diseases, such as retinal diseases (Jackson & Edwards, 2008; Jackson et al., 2006; Owsley et al., 2016), hearing loss (Moore, 1996), reading and language disorders (Demb, Boynton, & Heeger, 1998; Wright, Bowen, & Zecker, 2000), Parkinson's disease (Archibald, Clarke, Mosimann, & Burn, 2009; Weil et al., 2016), schizophrenia (Butler et al., 2007; D'Souza et al., 2005), and autism (Bertone, Mottron, Jelenic, & Faubert, 2005).

Many psychophysical methods have been developed to measure perceptual sensitivity (Cornsweet, 1962; Derman, 1957; Kaernbach, 1991; Kesten, 1958; King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994; Kontsevich & Tyler, 1999; Levitt, 1971; Robbins & Monro, 1951; Watson & Pelli, 1983; see Lu & Dosher, 2013 for a review). Traditionally, the method of constant stimuli (MCS) is used to measure perceptual sensitivity for stimuli that are fixed in all but one feature dimension such as contrast. To measure perceptual sensitivity as a function of

* Corresponding author at: 1835 Neil Ave., Department of Psychology, The Ohio State University, Columbus, OH 43202, USA.
E-mail address: lu.535@osu.edu (Z.-L. Lu).

https://doi.org/10.1016/j.visres.2018.10.009
Received 25 August 2018; Received in revised form 18 October 2018; Accepted 26 October 2018
Available online 12 November 2018
0042-6989/

other stimulus characteristics, such as spatial and/or temporal frequency, the MCS has to be repeated for each value of the characteristics. With the MCS, it usually takes hundreds to thousands of trials, adding up to several hours, to obtain a good estimate of a perceptual sensitivity function.

More efficient measurement of perceptual sensitivity and perceptual sensitivity function can be obtained with adaptive methods (see Leek, 2001; Lu & Dosher, 2013 for reviews). Some non-parametric procedures have been developed to measure a single threshold (Cornsweet, 1962; Derman, 1957; Kaernbach, 1991; Kesten, 1958; Levitt, 1971; Robbins & Monro, 1951). For example, the staircase methods (Cornsweet, 1962; Kaernbach, 1991) can be used to estimate a single threshold more efficiently than the MCS. More recently, methods based on Bayesian adaptive testing have been developed to measure a single threshold (King-Smith et al., 1994; Watson & Pelli, 1983), threshold and slope of a psychometric function (Kontsevich & Tyler, 1999; Prins, 2013), sensitivity threshold and bias in Yes-No and forced-choice tasks (the qYN and qFC methods; Lesmes et al., 2015), visual sensitivity functions (Baek, Lesmes, & Lu, 2016; Hou et al., 2010; Hou, Lesmes, Bex, Dorr, & Lu, 2015; Kujala & Lukka, 2006; Lesmes, Jeon, Lu, & Dosher, 2006; Lesmes, Lu, Baek, & Albright, 2010; Watson, 2017), as well as auditory functions (Shen & Richards, 2013; Shen, Sivakumar, & Richards, 2014; Song et al., 2015; Song, Garnett, & Barbour, 2017).

However, all the methods reviewed above provide an *average* perceptual sensitivity in multiple trials over some time interval, and could lead to biased and imprecise estimates if perceptual sensitivity changes rapidly over time. One potential solution (Fig. 1) is to aggregate data from multiple runs of the same experimental procedure, and use the data obtained at the same delay from the beginning of each run of the procedure to estimate a threshold at each time delay (Cornsweet & Teller, 1965; Mollon & Polden, 1980; Mollon, Stockman, & Polden, 1987; Pianta & Kalloniatis, 2000). The solution requires many repetitions of the same procedure to obtain enough trials to provide an estimated threshold at each delay and is therefore extremely inefficient. In addition, the method cannot be used in procedures that do not allow multiple repetitions, such as perceptual learning (Dosher & Lu, 2007).[1]

In this study, we develop a new method, the quick Change-Detection method (qCD: Zhao, Lesmes, & Lu, 2017), to accurately, precisely and efficiently estimate the time course of perceptual sensitivity change. The method explicitly models the functional form of the time course of perceptual sensitivity change and applies the Bayesian adaptive framework (Baek et al., 2016; Hou et al., 2015; King-Smith et al., 1994; Kontsevich & Tyler, 1999; Kujala & Lukka, 2006; Lesmes et al., 2006, 2010, 2015; Prins, 2013; Watson, 2017; Watson & Pelli, 1983) to update, trial by trial, a joint probability distribution of the parameters that characterize perceptual sensitivity and its change over time. In addition, it selects the optimal stimulus condition for each trial by optimizing the expected information gain in the next trial. Using the posterior distribution of the parameters following each trial, the method provides a trial-by-trial estimate of perceptual sensitivity and therefore a detailed time course of perceptual sensitivity change. Therefore, the qCD adaptively measures human behavior and continuously estimates perceptual sensitivity over time. In the qCD, optimal stimulus selection and trial-by-trial posterior update (thus threshold estimates) are two integral components. As an additional step, we develop a post-hoc (offline) procedure to segregate the trial-by-trial posterior distributions into temporal segments with significantly different posteriors distributions (e.g., different time constants), each of which can be modeled with a single posterior distribution, to further improve the precision and accuracy of the qCD method through aggregation of information in each segment.

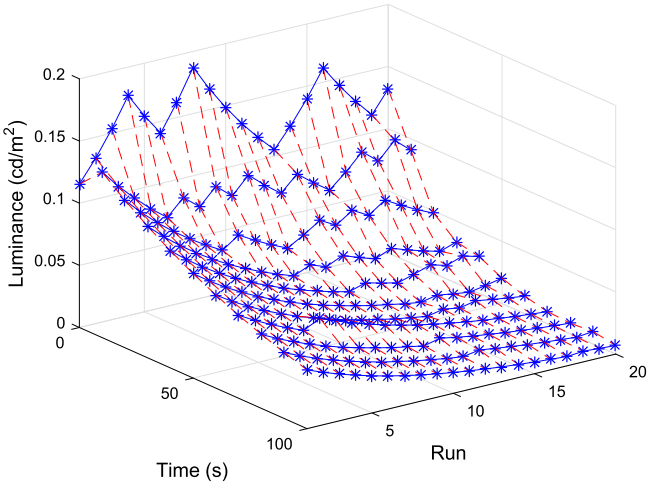


Fig. 1. Illustration of a staircase procedure used to estimate the time course of perceptual sensitivity change. In the first run of the experiment, the conventional staircase procedure is used (illustrated by the asterisks connected by the dashed line closest to the "Time" axis). That is, a pre-determined stimulus intensity is used in the first trial, and the staircase procedure determines the stimulus intensities in all subsequent trials. This sets the "initial" stimulus intensity at each time delay. In all subsequent runs, the stimulus intensity at each time delay is determined by the staircase procedure based on stimulus intensities tested at the same time delay across runs. Asterisks connected by solid lines across runs illustrate the trial sequences in the staircases.

As a first demonstration of the qCD method, we applied it to characterize perceptual sensitivity change in dark adaptation. Before describing the qCD method, we briefly review the relevant literature on dark adaptation.

The human eye can function over 15 orders of magnitude of luminance levels. This remarkable ability is achieved partly through retinal adaptation (Enroth-cugell & Shapley, 1973; Jakiela, Enroth-cugell, & Shapley, 1976; Purpura, Tranchina, Kaplan, & Shapley, 1990; Tranchina, Gordon, & Shapley, 1984) in which the photosensitive cells of the retina adjust their dynamic range to the current light level. Dark adaptation (Aubert, 1865) is a phenomenon in which, following exposure to intense or prolonged light, the perceptual sensitivity of the affected eye gradually recovers to the levels of fully dark-adapted eye. It requires regeneration of rhodopsin (see Lamb & Pugh, 2004; Reuter, 2011 for review), a pigment in the photoreceptors of the retina, that is bleached in response to light (Stuart & Birge, 1996; Wald, Durell, & Stgeorge, 1950). When most pigments are bleached after an intense light exposure, perceptual sensitivity of the affected eye of a normal subject typically decreases by 5 $\log_{10}$ units and recovers in a biphasic time course. The first phase, which happens during the first few minutes, is a rapid cone-mediated recovery to about 3 $\log_{10}$ units below the dark-adapted perceptual sensitivity, following an exponential time course (Gaffney, Binns, & Margrain, 2014; Hecht, Haig, & Chase, 1937; Hecht, Haig, & Wald, 1935; Hollins & Alpern, 1973; Mote & Riopelle, 1951). The second phase is a slow rod-mediated recovery from the cone sensitivity to the fully dark-adapted rod sensitivity, taking another 30 min or more (Hecht et al., 1935; Lamb, 1981; Pugh, 1975; Rushton & Powell, 1972). The rod-recovery phase can be further divided into a fast ***S2*** component and a slow ***S3*** component (Hecht et al., 1937; Lamb, 1981; Pugh, 1975). For bleach levels between 0.5% and 98%, the log sensitivity of the ***S2*** component can be fit by a straight line with a slope of 0.24 $\log_{10}$ units min^{-1} (i.e., the time constant of the exponential

[1] Recently, Kattner, Cochrane, and Green (2017) proposed a post-hoc procedure to analyze data obtained with the method of constant stimuli in perceptual learning experiments in which trial-by-trial data were fit with psychometric functions with time-varying bias and thresholds. However, as the authors noted, the amount of information gained in each trial is limited by the non-adaptive stimulus selection procedure.

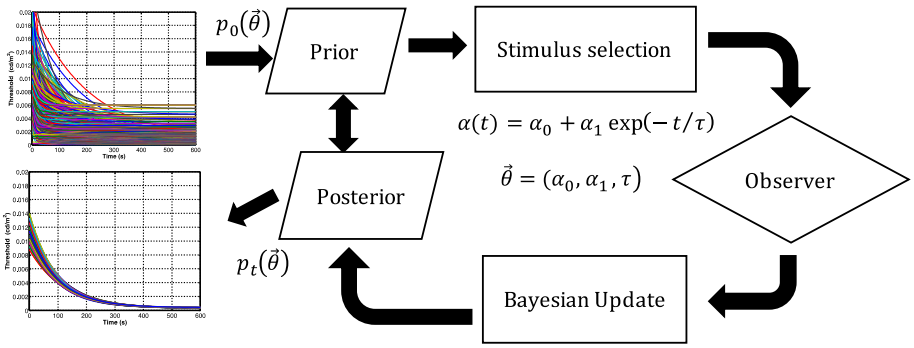


Fig. 2. The qCD method consists of five steps. (1) The time course of the luminance detection threshold is defined as an exponential function with three parameters, $\vec{\theta} = (\alpha_0, \alpha_1, \tau)$,and their joint prior distribution, $p_0(\vec{\theta})$. The top-left graph illustrates exponential functions with parameters randomly sampled from their joint prior distributions. (2) The luminance of the test stimulus in the next trial is selected to optimize the expected information gain on $\vec{\theta}$. (3) The posterior distribution of $\vec{\theta}$ is updated by Bayes' rule based on the observer's response after each trial. (4) Step (2) and (3) are repeated until the stop criterion is met (e.g., pre-determined number of trials). (5) The bottom-left graph illustrates exponential functions with parameters randomly sampled from their joint posterior distributions. Trial-by-trial and post-hoc segment-by-segment thresholds are computed from the posterior distribution.

decay function is about 110 s), and the log sensitivity of the ***S3*** component can be fit by a straight line with a bleach-level dependent slope, between 0 and 0.6 $\log_{10}$ units min^{-1} (i.e., time constant of the exponential decay function is about 400 s; Lamb, 1981; Pugh, 1975).

Impaired dark adaptation has been shown in early stage age-related macular degeneration (AMD). Brown, Adams, Coletta, and Haegerstrom-Portnoy (1986) found decreased sensitivity for both rods and cones, as well as increased time constant for rod-recovery during dark adaptation in early AMD. Recent studies have provided evidence for AMD related increases in the time constant for cone recovery (Gaffney, Binns, & Margrain, 2011, 2013), delayed rod-cone break (Dimitrov, Guymer, Zele, Anderson, & Vingrys, 2008; Gaffney et al., 2011, 2013; Owsley, Jackson, White, Feist, & Edwards, 2001), and delayed rod recovery (Brown et al., 1986; Dimitrov et al., 2008, 2011; Owsley et al., 2001). Impaired dark adaptation has been found in different retinal locations (Gaffney et al., 2011; Steinmetz, Haimovici, Jubb, Fitzke, & Bird, 1993) and at different bleach levels (Dimitrov et al., 2008; Gaffney et al., 2013). Others showed that the ***S2*** (but not ***S3***) component of rod dark adaptation is impaired in AMD patients (Jackson et al., 1999; Owsley et al., 2001). Moreover, Owsley et al. (2016) suggested that the delayed dark adaptation process could serve as a potential biomarker to detect early AMD, although a full visual field dark adaptation test failed to find any significant difference between AMD patients and age-matched controls (Jackson et al., 2006).

In this paper, we first present the qCD method in the context of a dark adaptation experiment. We then present a simulation study that characterized the precision, accuracy, and efficiency of the qCD method, in comparison with two other adaptive methods, the quick Forced-Choice (qFC) method (Hou et al., 2015; Lesmes et al., 2015) and the weighted up-down staircase method (Kaernbach, 1991). The qFC method belongs to a family of novel Bayesian adaptive methods that estimate the sensitivity threshold in Yes-No and Forced Choice tasks. In these procedures, the psychometric function is modeled with parameters related to perceptual sensitivity and decision bias. A Bayesian adaptive procedure is used to select the optimal stimulus in each trial and estimate the joint posterior distribution of the parameters. Like the staircase procedure, the qFC method does not explicitly model perceptual sensitivity change over its measurement block.[2] Next, we present a psychophysical validation study of the qCD method in which the estimated time course of perceptual sensitivity change obtained from the qCD method was compared with that obtained from the qFC method. Finally, we discuss extensions and variations, potential applications as well as limitations of the qCD method.

2. The quick change-detection method (qCD)

While the qCD is developed to be capable of estimating perceptual change over several stimulus dimensions, for this initial demonstration we have implemented it to measure the time course of perceptual sensitivity change in a single stimulus dimension (luminance) in dark adaptation. Based on the literature of dark adaptation, we assume luminance threshold decays as an exponential function during dark adaptation.

To assess luminance threshold change during dark adaptation, the qCD is implemented as a five-step procedure (see Fig. 2): (1) The time course of the luminance detection threshold is defined as an exponential function with three parameters, $\vec{\theta} = (\alpha_0, \alpha_1, \tau)$, and their joint prior distribution, $p_0(\vec{\theta})$, defined in the three-dimension space of $\vec{\theta}$. (2) The luminance of the test stimulus in the next trial is selected to optimize the expected information gain on the parameters of the exponential function. (3) The posterior distribution of $\vec{\theta}$ is updated by Bayes' rule based on the observer's response after each trial. (4) Step (2) and (3) are repeated until the stop criterion is met (e.g., pre-determined number of trials). (5) Trial-by-trial and post-hoc segment-by-segment threshold are computed from the posterior distributions.

2.1. Step 1. Threshold decay function, Psychometric function, and Priors

The time course of the luminance detection threshold during dark adaptation is described by an exponential function:

$$\alpha(\vec{\theta}, t) = \alpha_0 + \alpha_1 \exp(-t/\tau), \quad (1)$$

where $\vec{\theta} = (\theta_1, \theta_2, \theta_3) = (\alpha_0, \alpha_1, \tau)$, t is the time in seconds elapsed since the beginning of the dark adaptation, $\alpha(\vec{\theta}, t)$ is the luminance detection threshold at the $d' = 1.5$ performance level, α_0 is the asymptotic threshold when the eye is fully dark-adapted, $\alpha_0 + \alpha_1$ is the threshold at $t = 0$ second in the beginning of the dark adaptation, and τ is the time constant of the exponential function.

A broad joint prior distribution $p_0(\vec{\theta})$ is defined in the three-di-

[2] With the assumption of equal threshold at all eight spatial locations and no decision bias, the qFC implemented in the current study is equivalent to the Ψ method (Kontsevich & Tyler, 1999), which has been extensively used in psychophysical studies.

mensional parameter space of $\vec{\theta}$, where the subscript 0 is used to denote the prior that represents *a priori* knowledge of the parameters of the dark adaptation curve before the experiment. A one-dimension stimulus luminance space X is setup to cover all possible test luminance levels $x \in X$.

In an 8-Alternative Forced-Choice (8AFC) task, the psychometric function for luminance detection is approximated by a Weibull function (see Hou et al., 2015; Lesmes et al., 2015):

$$p_t(r = 1|\vec{\theta}, x) = (1 - \lambda)(g + (1 - g)\left(1 - \exp\left(-\left(\frac{x}{\alpha_w(\vec{\theta}, t)}\right)^{\gamma}\right)\right)) + \lambda g, \tag{2}$$

where the subscript t is time, r is the response (1 for correct, 0 for incorrect), x is the luminance of the stimulus, $g = 0.125$ is the guessing rate in an 8AFC task, $\gamma = 3.8959$ is the slope of the Weibull function (Hou et al., 2015), $\lambda = 0.04$ leads to a lapse rate of 0.035 (Swanson & Birch, 1992; Treutwein & Strasburger, 1999), and $\alpha_w(\vec{\theta}, t)$ is the Weibull threshold corresponding to a d' of 1.9 at time t given parameters $\vec{\theta}$:

$$\log_{10}(\alpha_w(\vec{\theta}, t)) = \log_{10}(\alpha(\vec{\theta}, t)) - \frac{1}{\gamma}\log_{10}\left(\log\left(\frac{1 - g}{1 - p_{1.5}}\right)\right), \tag{3}$$

where $p_{1.5} = 0.553$ is the probability correct corresponding to $d' = 1.5$ in an 8AFC task. Therefore, $p_t(r = 1|\vec{\theta}, x)$is the probability of a correct response ($r = 1$) at time t since the beginning of dark adaptation, conditioned on parameters $\vec{\theta}$ and stimulus luminance x. The probability of an incorrect response ($r = 0$) is:

$$p_t(r = 0|\vec{\theta}, x) = 1 - p_t(r = 1|\vec{\theta}, x). \tag{4}$$

2.2. Step 2. One step ahead stimulus search

The luminance of the stimulus to be presented in the next trial is selected among all possible stimuli to optimize the expected information gain (Kujala & Lukka, 2006; Lesmes et al., 2006) in the next trial. The expected information gain of a potential test stimulus with luminance x is defined as $I_t(x)$:

$$I_t(x) = h\left(\sum_{\vec{\theta}} p_t(\vec{\theta}) p_t(r = 1|\vec{\theta}, x)\right) - \sum_{\vec{\theta}} p_t(\vec{\theta}) h(p_t(r = 1|\vec{\theta}, x)), \tag{5}$$

$$h(p) = -p\log(p) - (1 - p)\log(1 - p), \tag{6}$$

where $p_t(\vec{\theta})$ is the joint posterior distribution of the parameters at time t.

2.3. Step 3. Bayesian update

In the n^{th} trial, r_n is the observer's response to a stimulus with luminance x_n presented at time t_n, where the subscript n is the trial number. $p_{t_n}(\vec{\theta})$ is the prior distribution of parameters $\vec{\theta}$ at time t_n. The prior distribution $p_{t_n}(\vec{\theta})$ is updated to the posterior distribution $p_{t_n}(\vec{\theta}|r_n, x_n)$ by Bayes' rule:

$$p_{t_n}(\vec{\theta}|r_n, x_n) = \frac{p_{t_n}(r_n|\vec{\theta}, x_n) p_{t_n}(\vec{\theta})}{p_{t_n}(r_n|x_n)}, \tag{7}$$

$$p_{t_n}(r_n|x_n) = \sum_{\vec{\theta}} p_{t_n}(r_n|\vec{\theta}, x_n) p_{t_n}(\vec{\theta}), \tag{8}$$

where $\Sigma_{\vec{\theta}}$ is the summation over the three-dimensional parameter space, that is, $\Sigma_{\vec{\theta}} = \Sigma_{\alpha_0}\Sigma_{\alpha_1}\Sigma_{\tau}$.

The posterior distribution of $\vec{\theta}$ following the n^{th} trial is used as the prior of the $n+1^{\text{th}}$ trial:

$$p_{t_{n+1}}(\vec{\theta}) = p_{t_n}(\vec{\theta}|r_n, x_n). \tag{9}$$

The marginal posterior distribution of each parameter is computed via summation of the joint posterior distribution over the other two parameters in the three-dimensional parameter space:

$$\begin{aligned} p_{t_n}(\alpha_0|r_n, x_n) &= \sum_{\alpha_1}\sum_{\tau} p_{t_n}(\vec{\theta}|r_n, x_n), \\ p_{t_n}(\alpha_1|r_n, x_n) &= \sum_{\alpha_0}\sum_{\tau} p_{t_n}(\vec{\theta}|r_n, x_n), \\ p_{t_n}(\tau|r_n, x_n) &= \sum_{\alpha_0}\sum_{\alpha_1} p_{t_n}(\vec{\theta}|r_n, x_n). \end{aligned} \tag{10}$$

The expected values of the marginal posterior distributions provide estimates of the parameters of the exponential function after the n^{th} trial:

$$\bar{\theta}_{i,n} = \sum_{\theta_i} \theta_i \cdot p_{t_n}(\theta_i|r_n, x_n), \tag{11}$$

where $\theta_i = \alpha_0, \alpha_1, \text{ and } \tau$, for i = 1, 2 and 3, respectively.

2.4. Step 4. Iteration of Step 2 and Step 3

After the stimulus for the next trial is selected by Step 2 and presented to the observer, the joint posterior distribution of the parameters of the exponential decay function is updated by Step 3 based on the response of the observer. The experiment continues until the stop criterion is reached, which could be a pre-defined number of trials or a pre-defined precision level.

2.5. Step 5. Analysis

A trial-by-trial threshold can be estimated from the posterior distribution of the parameters of the exponential function after each trial. The method does not impose any constraint on the posterior distribution over time. In other words, the posterior distribution could vary from trial to trial, and the underlying exponential function could have different parameters from trial to trial.

We have also developed a post-hoc segmentation method to divide the trial-by-trial posterior distributions into segments such that each segment only contains similar posterior distributions. We then compute the segment-by-segment estimate of the dark adaptation curve to improve the precision and accuracy of the qCD method through aggregation of information in each segment. The detailed procedure is described in Appendix C.

3. Simulations

3.1. Methods

To evaluate the qCD method, we stimulated three observers with the same asymptotic threshold level and dynamics range but different time

constants: $\vec{\theta}_{observer,1}$ = (0.000376, 0.0113, 20), $\vec{\theta}_{observer,2}$ = (0.000376, 0.0113, 45) and $\vec{\theta}_{observer,3}$ =(0.000376, 0.0113, 100), to evaluate the performance of the qCD, the qFC (Appendix A), and the weighted up-down staircase (Appendix B) methods in measuring the time course of perceptual sensitivity change. In Appendix D, we present simulations of two additional observers whose luminance thresholds change as a cascade exponential function in dark adaptation.

The simulated observers performed an 8-alternative forced-choice (8AFC) location identification task in a dark adaptation procedure. Observer's response was simulated by drawing a random number y from a uniform distribution over the interval from 0 to 1. The response was scored as correct $(r = 1)$ if $y < p_t(r = 1|\vec{\theta}, x)$, and incorrect otherwise, with the probability of a correct response $p_t(r = 1|\vec{\theta}, x)$ defined by Eq. (2).

Each simulated experiment lasted 600 s. A 2-s inter-trial-interval (ITI) was implemented in all the simulations.[3] Therefore, 300 trials were simulated in each run for each of the three methods. For the qCD method, a threshold was estimated in every trial (2 s). For the qFC and staircase methods, a threshold was estimated in every five trials (10 s).

In the qCD simulations, the parameter space included 50 log-linearly spaced α_0 values between 0.000075 and 0.0060 cd/m^2, 50 log-linearly spaced α_1 values between 0.0023 and 0.030 cd/m^2, and 50 log-linearly spaced τ values between 5 and 200 s. For α_1, 0 was also included to account for no threshold change. The stimulus space consisted of 120 log-linearly spaced x values between 0.000075 and 0.075 cd/m^2. The prior distribution, $p_0(\vec{\theta})$, was defined by the hyperbolic secant function (King-Smith & Rose, 1997):

$$\begin{aligned} p_0(\vec{\theta}) = {} & \operatorname{sech}(\alpha_{0,\text{confidence}}(\log_{10}(\alpha_0) - \log_{10}(\alpha_{0,\text{mode}}))) \\ & \times \operatorname{sech}(\alpha_{1,\text{confidence}}(\log_{10}(\alpha_1) - \log_{10}(\alpha_{1,\text{mode}}))) \\ & \times (\operatorname{sech}(\tau_{\text{confidence},1}(\log_{10}(\tau) - \log_{10}(\tau_{\text{mode},1}))) \\ & + \operatorname{sech}(\tau_{\text{confidence},2}(\log_{10}(\tau) - \log_{10}(\tau_{\text{mode},2})))), \end{aligned} \tag{12}$$

where $\operatorname{sech}(x) = \frac{2}{e^x + e^{-x}}$; $(\alpha_{0,\text{mode}}, \alpha_{1,\text{mode}}, \tau_{\text{mode},1}, \tau_{\text{mode},2}) = (0.00075, 0.0075, 20, 100)$ are the peaks of the respective secant functions; $(\alpha_{0,\text{confidence}}, \alpha_{1,\text{confidence}}, \tau_{\text{confidence},1}, \tau_{\text{confidence},2}) = (4, 4, 6, 6)$ are the spreads of the respective secant functions. Note that we use a distribution with two modes in setting up the prior for τ based on pilot experiment. The prior $p_0(\vec{\theta})$ is the multiplication of two secant functions of α_0 and α_1 and the sum of two secant functions of τ. The joint prior distribution $p_0(\vec{\theta})$ was updated trial-by-trial throughout the simulated experiment. The simulation was performed 1000 times for each simulated observer.

The accuracy of an estimated parameter is quantified by the bias, i.e., the difference between the estimated and true values, after the n^{th} trial defined as:

$$\sum_m (\log_{10}(\bar{\theta}_{Observer,ainm}) - \log_{10}(\theta_{Observer,ai}))/M, \tag{13.1}$$

$$\bar{\theta}_{Observer,ainm} = \sum_{\theta_{Observer,ai}} \theta_{Observer,ai} \cdot p_{t_{nm}}(\theta_{Observer,ai}|r_{nm}, x_{nm}), \tag{13.2}$$

where $\theta_{Observer,ai}$ is the true parameter value of the simulated observer; $\bar{\theta}_{Observer,ainm}$, t_{nm}, r_{nm}, x_{nm} and $p_{t_{nm}}(\theta_{Observer,ai}|r_{nm}, x_{nm})$ are the estimate of the parameter θ_i for Observer a after the n$^{\text{th}}$ trial in the m$^{\text{th}}$ simulation, the time at which the stimulus was presented, the observer's response, the stimulus luminance, and the marginal distribution defined in Eq. (10) after the nth trial in the mth simulation, respectively; M (=1000) is the total number of simulated runs. The smaller the absolute value of the bias is, the more accurate the estimate is. The precision of an estimated parameter, i.e., the spread of the estimated values, is quantified by the average half width of the 68.2% credible interval (68.2% HWCI) of the estimated parameter θ_a, the half of the width of the interval within which the true value lies with 68.2% probability based on the marginal distribution $p_{t_{nm}}(\theta_a|r_{nm}, x_{nm})$. The HWCI is a measure of within-run imprecision. The smaller the HWCI, the more precise the estimate is.

The estimated threshold after the n^{th} trial in the m^{th} simulation is computed via a resampling procedure.[4] The qCD method estimates the posterior distribution of the parameters in each trial, but not the posterior distribution of the threshold. The resampling procedure was used to generate the posterior distributions of the thresholds from the posterior distributions of the parameters. This resampling procedure takes into account the covariance structure in the joint posterior distributions of the parameters and can be used to directly assess the variability of the estimated exponential decay function. 1000 parameter vectors $\vec{\theta}_{znm}$ are independently sampled based the joint posterior distribution $p_{t_{nm}}(\vec{\theta}|r_{nm}, x_{nm})$, where $z \in [1, 1000]$ and $Z = 1000$ is the total number of samples. $\vec{\theta}_{znm}$ are used to compute 1000 sets of estimated values of the exponential functions at time point t_k, where

$$\begin{cases} t_1 = 10 \\ t_k - t_{k-1} = 10, \forall\ k > 1' \end{cases} \tag{14}$$

Therefore, for the 600-s experiment, the total number of evaluation points K is 60. The estimated threshold $\hat{\alpha}_{\{nm\}(t_k)}$ at timet_k in the n^{th} trial in the m^{th} simulation is the average of the 1000 samples:

$$\hat{\alpha}_{\{nm\}(t_k)} = \sum_z \alpha_{\{znm\}(t_k)}/Z = \sum_z \alpha(\vec{\theta}_{znm}, t_k)/Z\ , \tag{15}$$

where $\alpha(\vec{\theta}_{znm}, t_k)$ refers to Eq. (1). Therefore, an estimated threshold can be computed at any non-negative evaluation time t_k regardless of the stimulus presentation time t_{nm}.

In this paper, the accuracy of the estimated thresholds is quantified by the root mean square error (RMSE) between the true and estimated thresholds. The precision of the estimated thresholds is quantified by the standard deviation (SD) as the cross-run variability and by the 68.2% HWCI as the within-run imprecision. The smaller these values are, the more accurate and precise the estimates are, respectively. If subjects are stable across runs, the SD and 68.2% HWCI would approach the same value with sufficient number of trials (Hou et al., 2015).

[3] The ITI used in the simulation was a conservative approximation based on the actual ITIs obtained in the real experiments. The mean ITIs, averaged across runs and subjects, were 1.16 ± 0.08 and 1.10 ± 0.10 s for the qCD and qFC methods in the psychophysical validation study, respectively. In other words, the mean total number of trials, averaged across runs and subjects, were 516.2 ± 32.7 and 547.4 ± 48.6 trials for the qCD and qFC methods in the experiment, respectively. The onset of the stimulus relative to the onset of the dark adaptation period was used to fit the model in both the qCD and qFC methods. Subjects were encouraged to start the test as soon as possible after the onset of dark adaptation and respond as often as possible during dark adaptation. The mean onset time of the first stimulus and the mean number of trials of the first 10-second interval in the qFC procedure, averaged across runs and subjects, were 1.75 ± 0.86 s and 8.4 ± 1.8 trials, respectively. Therefore, there were sufficient trials during the first 10-second interval to estimate the first threshold in the qFC procedure.

[4] The posterior is a 50 × 51 × 50 three-dimensional grid, representing the probability of each of 127,500 sets of parameter combinations. The three-dimensional posterior is converted into a one-dimension probability vector (1 × 127,500). The sum of the 127,500 probabilities equals to one. The probability vector is then converted into a "cumulative probability" vector with the same size, in which the values of one item is the sum of all the previous items of the probability vector. Each time, a random sample from a uniform distribution between 0 and 1 is generated. The last set of parameters with the "cumulative probability" no greater than the random number are selected and the exponential function is calculated based on this set of parameters. The process is repeated 1000 times to generate the 1000 samples from one posterior distribution.

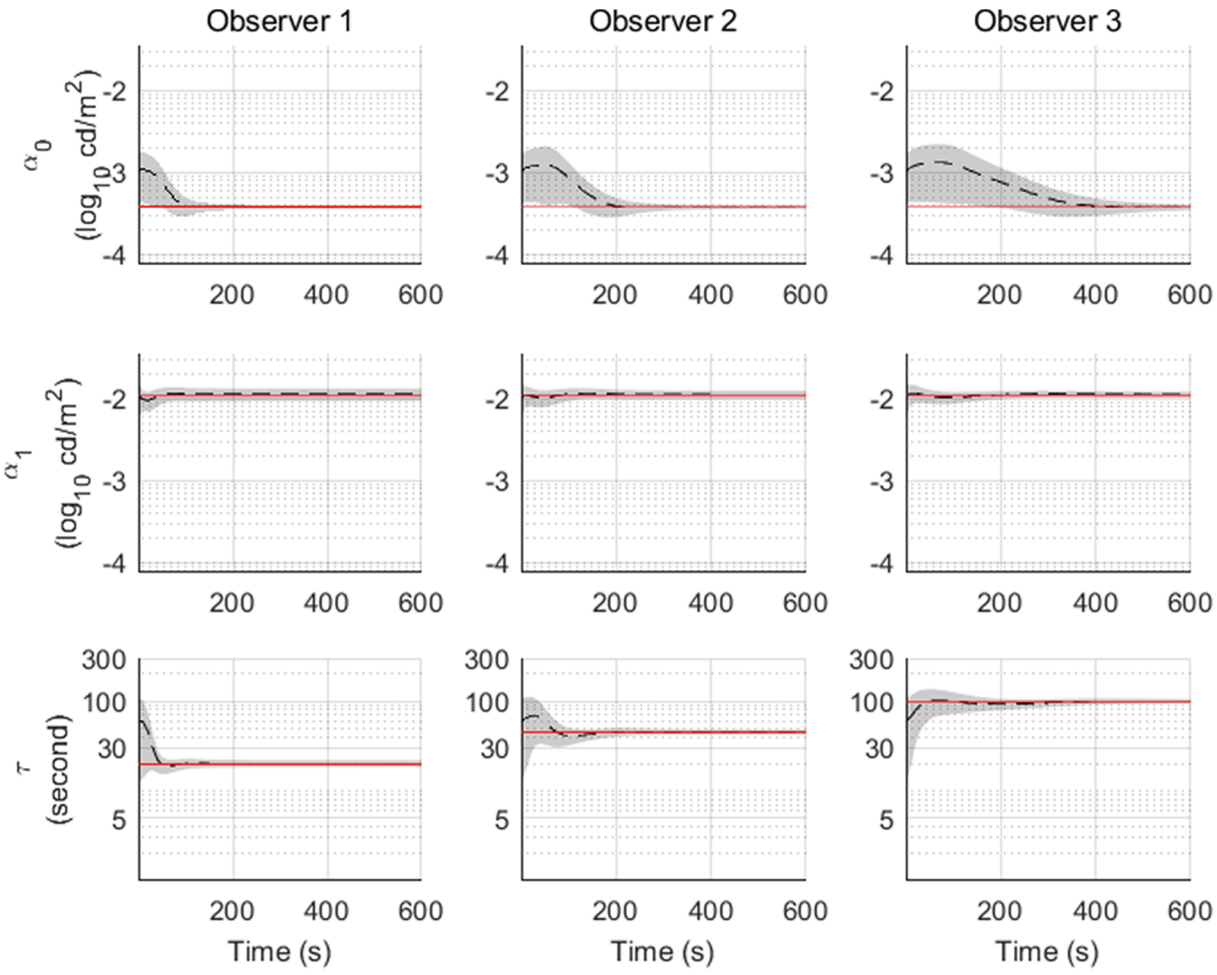


Fig. 3. Trial-by-trial parameter estimates. Results of the three simulated observers are presented in different columns. Parameter estimates for α_0(top row), α_1(middle row), and τ (bottom row) of each observer are plotted as functions of time in seconds since the beginning of the dark adaptation. The red solid horizontal lines represent the true parameter values of the simulated observers, the black dashed lines represent the average estimates, and the grey shaded areas represent the average 68.2% HWCI of the marginalized parameter probability distributions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The accuracy of the estimated thresholds from the qCD method was quantified by the $RMSE(n)$ across independent simulation runs as a function of n:

$$RMSE(n) = \sqrt{\frac{\sum_k \left[\log_{10}\left(\bar{\hat{\alpha}}_{\{n\}(t_k)}\right) - \log_{10}\left(\alpha_{observer(t_k)}\right)\right]^2}{K}}, \tag{16.1}$$

$$\bar{\hat{\alpha}}_{\{n\}(t_k)} = \sum_m \hat{\alpha}_{\{nm\}(t_k)}/M, \tag{16.2}$$

where $\bar{\hat{\alpha}}_{\{n\}(t_k)}$ is the average estimated threshold at evaluation time t_k after the n^{th} trial across M runs; $\hat{\alpha}_{\{nm\}(t_k)}$ is defined in Eq. (15); $\alpha_{observer(t_k)}$ is the true threshold of the simulated observer at time t_k. The precision can be quantified by the SD across independent simulation runs and by the 68.2% HWCI within one run. The standard deviation $SD(n)$ of the estimated threshold after the n^{th} trial across M runs is defined as a function of n:

$$SD(n) = \sqrt{\frac{\sum_k \sum_m \left[\log_{10}\left(\hat{\alpha}_{\{nm\}(t_k)}\right) - \log_{10}\left(\bar{\hat{\alpha}}_{\{n\}(t_k)}\right)\right]^2}{M \times K}}. \tag{17}$$

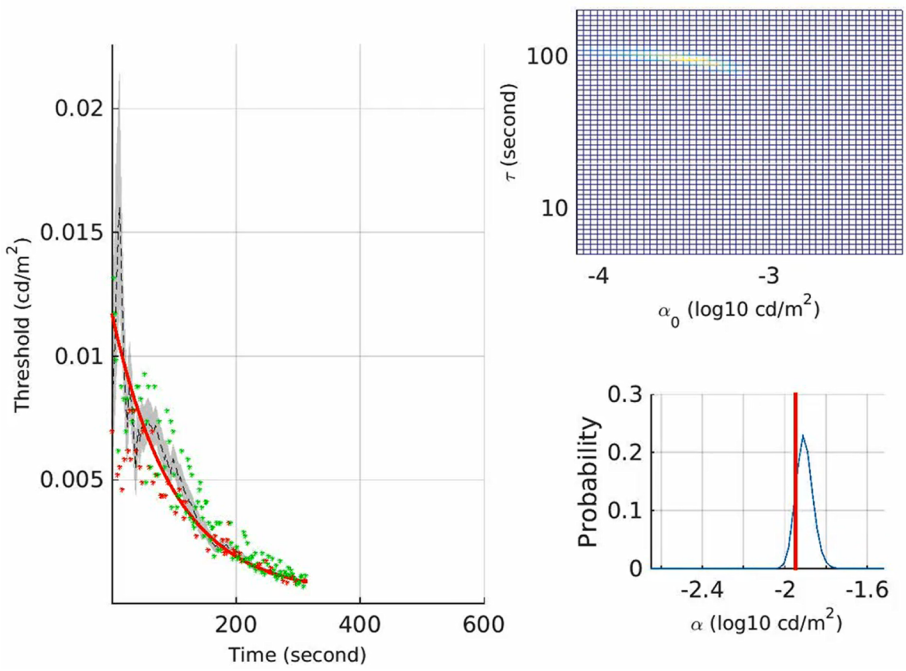
We also computed the average 68.2% HWCI of the threshold estimates, defined as the half width of the interval within which the true value lies with 68.2% probability based on the 1000 samples.

3.2. Results

The qCD started with a broad prior distribution of the three parameters of the dark adaptation curve. As the experiment proceeded, the posterior distribution became narrower and converged towards the true parameter values. A demo of the procedure based on simulated Observer 3 is provided in Movie 1. In Fig. 3, the narrowing of the posterior distribution, i.e., reduction of the variability of the estimates, is quantified by the decrease of the average 68.2% HWCIs (grey shaded area) of the marginal posterior distributions of the parameters. The convergence of the posterior distribution towards the true values (red solid horizontal lines), i.e., improved accuracy of the estimates, is

quantified by the decrease of the bias of the parameter estimates (distance between the black dashed lines and red solid lines). Fig. 4 shows the bias and average 68.2% HWCI of α_0, α_1, and τ for Observers 1–3.

for each simulated observer. Fig. 6 shows the estimated dark adaptation curve in the end of one qCD run for the three simulated observers. For simulated Observer 1, the RMSE, SD, and average 68.2% HWCI was



Movie 1. Trial-by-trial threshold estimates, stimulus placement and maginalized parameter distributions in one example run of the qCD simulation of Observer 3. Left panel: The black dashed line represents trial-by-trial threshold estimates, the grey shaded area represents the corresponding 68.2% HWCI, and the red solid line represents the true thresholds. The green dots represent stimulus lumance in trials with correct responses and the red dots with incorrect responses. Top right panel: marginalized joint probability distribution of τ and α_0. Bottom right panel: marginalized probability distribution of α_1.

Next, we present both the trial-by-trial and post-hoc segment-by-segment estimates of the dark adaptation curve from the qCD method. Fig. 5 shows the trial-by-trial threshold estimates from qCD for the three simulated observers. To compute the trial-by-trial threshold, the posterior distribution after each trial was used. For simulated Observer 1, the RMSE, SD, and average 68.2% HWCI of the estimated thresholds during 600-s dark adaptation was 0.025, 0.032, and 0.031 $\log_{10}$ units, respectively. For simulated Observer 2, the RMSE, SD, and average 68.2% HWCI of the estimated thresholds in the 600-s dark adaptation was 0.023, 0.039, and 0.039 $\log_{10}$ units, respectively. For simulated Observer 3, the RMSE, SD, and average 68.2% HWCI of the estimated thresholds in the 600-s dark adaptation was 0.019, 0.039, and 0.043 $\log_{10}$ units, respectively. These results thus demonstrate the high precision and accuracy of the trial-by-trial qCD estimates.

In the trial-by-trial procedure, the posterior distribution after each trial is used to estimate the threshold in that trial. While the procedure achieved high precision and accuracy, the post-hoc segment-by-segment estimation further improved the accuracy and precision (Appendix C). The criterion MD_0 in Eq. (C2) was set to 6.58. Out of the 1000 simulated runs, 0, 10, and 21 runs resulted in two posterior segments for simulated Observer 1, 2, and 3, respectively while all the remaining runs only indicated one segment for the entire 600-s experiment. Based on this criterion, only one posterior distribution is necessary to describe the dark adaptation curves of the three simulated observers. Therefore, the posterior distributions at 600 s were used to compute the estimated thresholds on the entire dark adaptation curve 0.002, 0.019, and 0.012 $\log_{10}$ units, respectively. For simulated Observer 2, the RMSE, SD, and average 68.2% HWCI was 0.002, 0.020, and 0.017 $\log_{10}$ units, respectively. For simulated Observer 3, the RMSE, SD, and average 68.2% HWCI was 0.002, 0.020, and 0.020 $\log_{10}$ units, respectively.

Next, we present comparisons of the simulation results of the estimated thresholds by the segment-by segment qCD, qFC, and weighted up-and-down staircase methods (Fig. 7) and refer the readers to Appendices A and B for details of the qFC and staircase simulations. Fig. 7a shows the accuracy (RMSE) and precision (SD and HWCI) of the estimated thresholds (of the three simulated observers) as a function of time in seconds in one qCD simulation run. At $t = 0$ s, the RMSE, SD, and average 68.2% HWCI was 0.434, 0.073, and 0.333 $\log_{10}$ units, respectively. These values decreased to 0.002, 0.020, and 0.016 $\log_{10}$ units, respectively, at $t = 600$ s. In comparison, Fig. 7b shows the accuracy (RMSE) and precision (SD and HWCI) as a function of run number for the qCD, qFC and staircase methods. After one single qCD run, ten qFC runs and ten staircase runs, the average RMSE was 0.002, 0.020 and 0.026 $\log_{10}$ units, respectively; the average SD was 0.020, 0.031 and 0.031 $\log_{10}$ units, respectively; the average 68.2% HWCI was 0.016 and 0.032 $\log_{10}$ units. Therefore, estimated thresholds from one single qCD run achieved higher accuracy and precision than those from the 10 qFC runs and the 10 staircase runs.

In the simulations of observers with a single exponential dark adaptation curve, the exponential function obtained with a single run of the post-hoc segment-by-segment qCD showed high accuracy

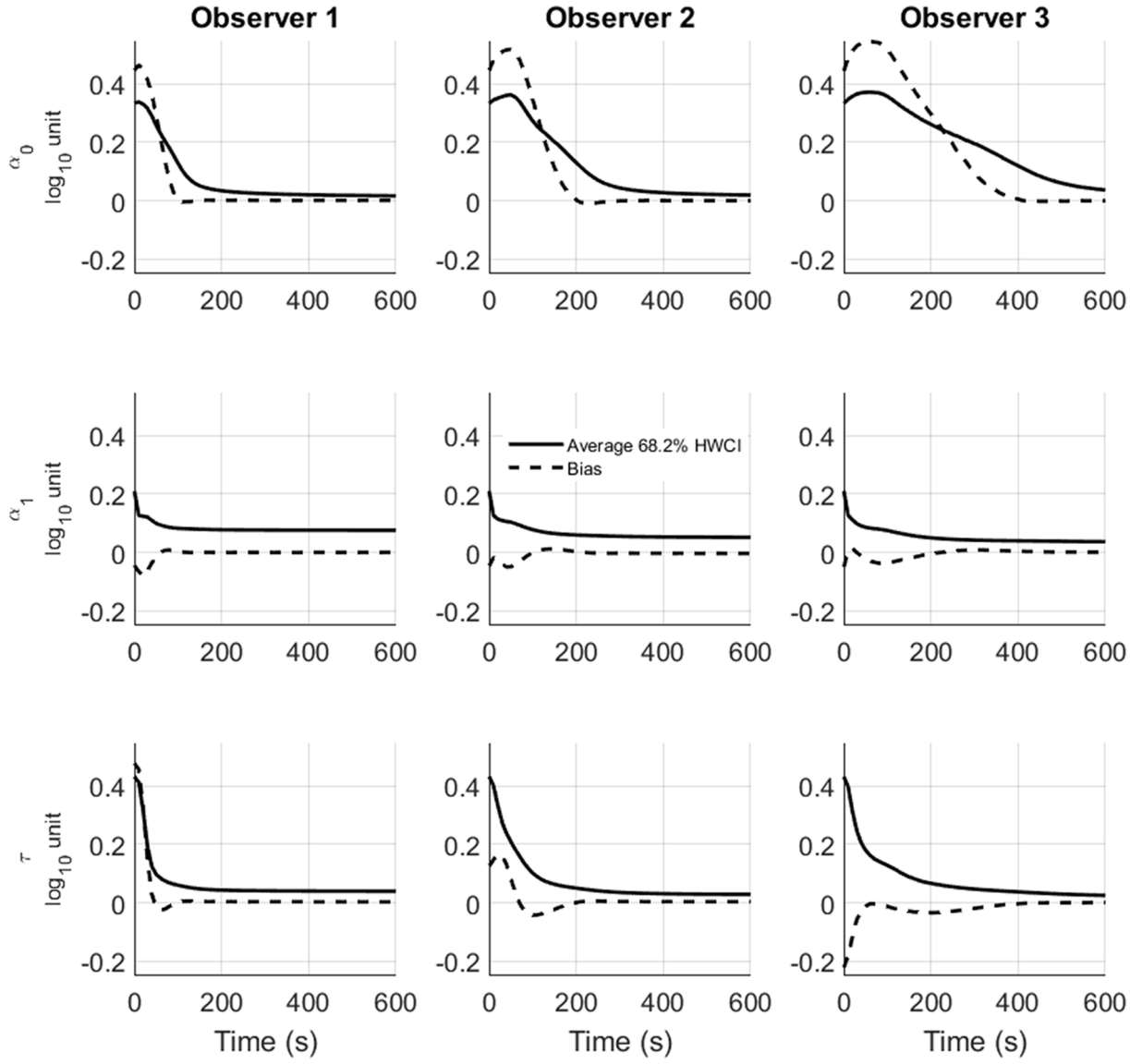


Fig. 4. Accuracy and precision of trial-by-trial parameter estimates, averaged across 1000 simulated qCD runs. Results of the three simulated observers are presented in different columns. Accuracy (bias, dashed lines) and precision (68.2% HWCI, solid lines) of parameter estimates for α_0 (top row), α_1 (middle row), and τ (bottom row) of each observer are plotted as functions of time in seconds since the beginning of the dark adaptation.

(RMSE = 0.002 $\log_{10}$ units) and precision (average 68.2% HWCI = 0.017 $\log_{10}$ units; SD = 0.020 $\log_{10}$ units). However, even after 10 runs, the threshold estimates from both the qFC and staircase methods were less accurate (RMSE = 0.020 and 0.026 $\log_{10}$ units, for qFC and staircase, respectively) and less precise (SD = 0.031 and 0.031 $\log_{10}$ units, for the qFC and staircase methods, respectively; average 68.2% HWCI = 0.032 $\log_{10}$ units for qFC) than those from a single qCD run. Simulation results of two observers with dark adaptation curves consisted of a cascade of exponential functions are presented in Appendix D.

4. Psychophysical validation

A psychophysical experiment was conducted to evaluate and compare the performance of the qCD and qFC methods. Each experimental run consisted of two parts: exposure to high luminance of 75 cd/m^2 for 15 s and 150 cd/m^2 for 120 s with room lights on, which resulted in about 3% bleach of retinal pigments (Hollins & Alpern, 1973; Thomas & Lamb, 1999). This was immediately followed by measurement of luminance detection thresholds during 600-s dark adaptation at 0.000 cd/m^2 mean luminance with room lights off. Only one method, qCD or qFC, was used in each run. Each subject finished four independent qCD runs and four connected qFC runs (see below). The two methods were tested in alternating runs. The method used in the first run was randomly selected for each subject.

In each qFC run, thresholds were estimated every 10 s. The prior distribution $p_0(\theta_{FC} = \alpha_{FC})$ was updated only during the corresponding 10-s interval which resulted in 60 posterior distributions after one run of the 600-s experiment. The 60 posterior distributions obtained in one run were used as the priors in the next run. The qFC procedure was iterated in four runs for a single subject in the psychophysical experiment.

Consistent with the simulations, qCD showed higher precision and efficiency than qFC in all four subjects.

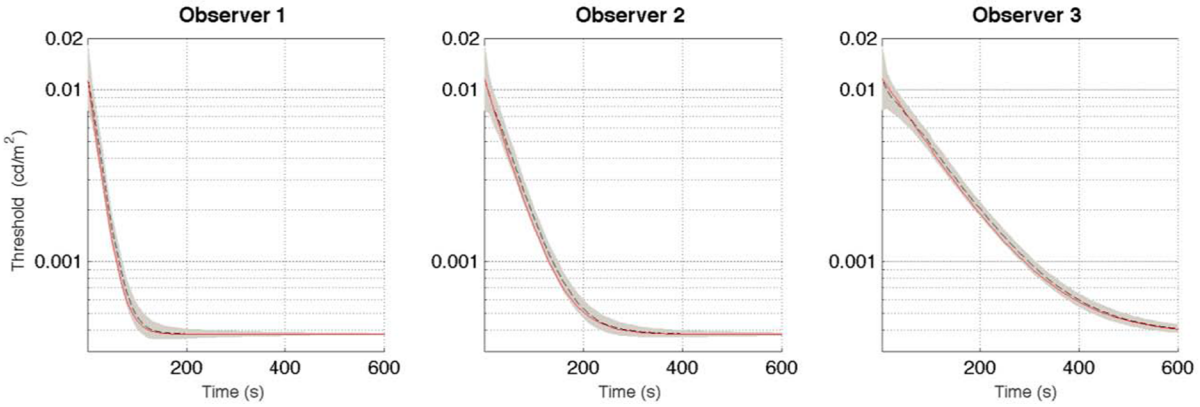


Fig. 5. Trial-by-trial threshold estimates from the qCD simulations. Threshold estimates are plotted as functions of time (in s) elapsed since the beginning of the dark adaptation. Results of the three simulated observers are presented. The solid red lines represent the true values of the simulated observers, the black dashed lines represent the average estimates and the grey shaded areas represent the average 68.2% HWCI. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

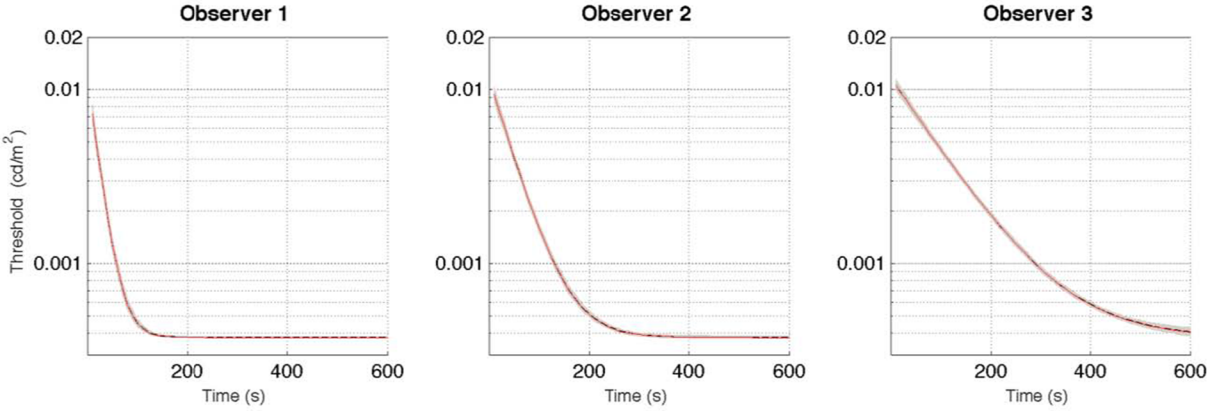


Fig. 6. Post-hoc segment-by-segment analysis. A single exponential function can be used to describe the estimated dark adaptation curve for each simulated observer. Results of the three simulated observers are presented. The solid red lines represent the true exponential decay functions of the simulated observers, the black dashed lines represent the estimated exponential decay functions and the grey shaded areas represent the average 68.2% HWCI. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.1. Method

4.1.1. Apparatus

The experiment was conducted on a PC computer with Pyschtoolbox 3.0.11 extensions in Matlab R2013a. The computer was used to drive a 46-inch NEC MultiSync P463 LCD monitor that was viewed binocularly with natural pupils at a two-meter viewing distance. A chin rest was used to stabilize subject's head. Responses were collected via a wireless keyboard. Luminance was controlled by a pair of goggles with removable neutral density filters with a total of 3.3 ND (one 2.4 ND and one 0.9 ND), reducing light intensity by a factor of $10^{3.3}$ (or 99.95%) in each eye. The goggles were worn throughout the whole experiment but the neutral density filters were flipped on only during the 600-s dark adaptation phase. Room lights were turned off during dark adaptation.

4.1.2. Stimuli

The target in each trial was a 1.7°-diameter luminance disk that could occur at one of eight locations of an imaginary 5°-diameter circle. The luminance of the disk varied trial by trial and was determined by the adaptive procedures. Four 13.4°-long and 0.06°-wide lines crossing the centers of the eight potential locations and the center of the display (with gaps to allow potential stimulus presentation), were presented throughout the experiment to reduce spatial uncertainty of the stimuli. A 0.5°-diameter dot was also presented at the fixation (Fig. 8a) but disappeared during the presentation of the target disk (Fig. 8b). The disappearance of the fixation dot served as a temporal cue. The luminance of the cues was 60, 135, and 0.0376 cd/m^2 during the 15-s exposure, 120-s exposure and 600-s dark adaptation phase, respectively.

4.1.3. Subjects

Three naïve subjects and one of the authors participated in the experiment. All subjects were male between 29 and 49 years old with normal or corrected-to-normal vision. The study was approved by the institutional review board of human subject research at the Ohio State University. Written consent was obtained from each subject in the beginning of the experiment. This research followed the tenets of the declaration of Helsinki.

4.1.4. Procedure

Each experimental run started with a 15-s exposure to a screen with a uniform luminance of 75 cd/m^2. Then the screen luminance doubled to 150 cd/m^2 for 120 s. During each trial of the 120-s exposure phase, one negative contrast single disk with luminance between 135 and 150 cd/m^2 selected by the same adaptive method used in the dark adaptation measurement appeared for 0.2 s randomly at one of the eight possible locations with equal probability. The subject pressed one of eight keys to indicate the location where the disk occurred. No feedback was given. The next trial started 0.2 s after a key was pressed. Subjects were instructed to simply guess if they did not see a disk. The task served as practice before the measurement of dark adaptation.

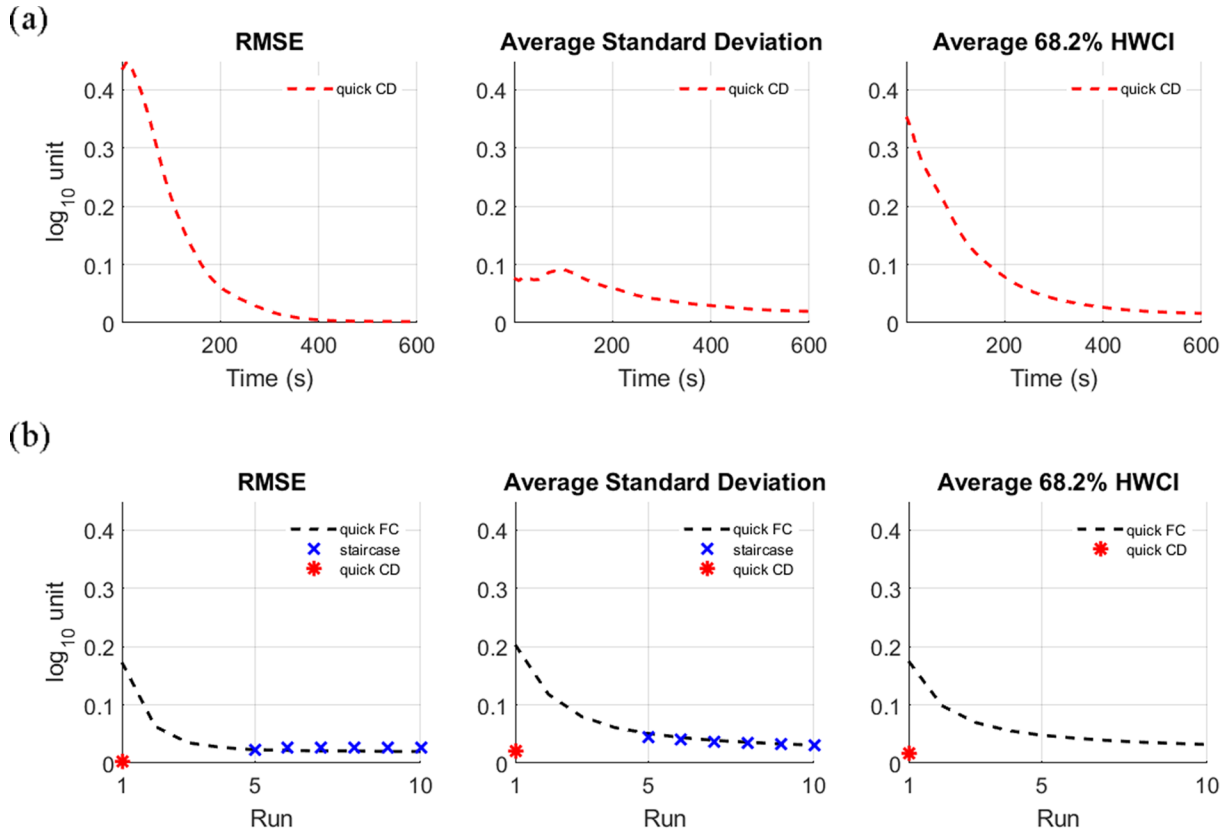


Fig. 7. Comparison of the accuracy (RMSE) and precision (SD and HWCI) of the estimated thresholds from the segment-by-segment qCD, qFC and weighted up-and-down staircase methods. (a) Accuracy and precision as a function of time in seconds in one qCD run. (b) Accuracy and precision as a function of run number. Red asterisks represent the accuracy and precision in the end of a single qCD run. Black dashed lines represent the accuracy and precision of qFC as a function of run number. Blue crosses represent accuracy and precision of the weighted up-and-down staircase method as a function of run number. The staircase did not converge until the 6th run. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The text "Lights off, Filters on" appeared on the screen during the last 3 s of this phase to instruct subject to flip on the neutral density filters as soon as the dark adaptation started. The screen luminance fell to 0.000 cd/m^2 in the end of the 120-s exposure and the room lights were turned off simultaneously. The text "Press 5 when ready" appeared on the screen and instructed the subject to press the "5" key to initiate the first trial in the dark adaptation.

During each trial of the dark adaptation, one single disk appeared for 0.2 s randomly at one of the eight possible locations with equal probability. The subject pressed one of the eight keys to indicate the location where the disk occurred. No feedback was given. The next trial started 0.2 s after a key was pressed. Subjects were instructed to simply guess if they did not see a disk.

4.1.5. Data analysis

The agreement between the qCD and qFC measurements was quantified by the root mean squared error (RMSE) between the estimated thresholds from the two methods:

$$RMSE = \sqrt{\sum_{k} \left[\log_{10}\left(\hat{\alpha}_{\{4\}(t_k)}\right) - \log_{10}(\bar{\hat{\alpha}}_{t_k})\right]^2 / K}\ , \tag{18}$$

$$\bar{\hat{\alpha}}_{t_k} = \sum_{m} \hat{\alpha}_{\{m\}(t_k)} / M_{CD}\ , \tag{19}$$

where $\hat{\alpha}_{\{4\}(t_k)}$ is the estimated threshold of the k^{th} interval in the 4$^{\text{th}}$ qFC run and $\bar{\hat{\alpha}}_{t_k}$ is the average trail-by-trial threshold estimate at time t_k from the qCD across $M_{CD} = 4$ runs.

4.2. Results

The qCD tracked changes of the parameters of the exponential decay function trial-by-trial. Fig. 9 shows that the average 68.2% HWCIs (grey shaded areas) of the marginalized joint probability density were wide in the beginning and usually became narrower as the experiment proceeded. Furthermore, the parameter estimates (dashed lines) changed over time.

Fig. 10 shows the average trial-by-trial threshold estimates from the four independent qCD runs and the threshold estimates from all four qFC runs. Table 1 shows the precision (68.2% HWCI) of the threshold estimates from the two methods. The average 68.2% HWCI of threshold estimates from the four independent qCD runs was 0.037, 0.038, 0.028 and 0.040 $\log_{10}$ units for subjects S1, S2, S3, and S4, respectively. In comparison, the average 68.2% HWCI from the final posteriors of all four qFC runs was 0.040, 0.043, 0.040, and 0.038 $\log_{10}$ units after the 4th qFC run, for subjects S1, S2, S3, and S4, respectively (Table 1).

The average RMSEs between the average trial-by-trial threshold estimates from the qCD and the threshold estimates from all four qFC runs were 0.074, 0.106, 0.057 and 0.069 $\log_{10}$ units for S1, S2, S3, and S4, respectively. The correlation coefficients were 0.989, 0.984, 0.924, and 0.980 for S1, S2, S3, and S4, respectively ($p < 0.001$ for all four subjects). These results indicate that the trial-by-trial threshold estimates from qCD were consistent with those from qFC.

Using a criterion of $MD_0 = 6.58$, the number of segments of the four independent qCD runs are 3, 2, 3, and 3 for S1; 2, 2, 2, and 1 for S2; 1, 1, 1, and 2 for S3; 2, 2, 2, and 2 for S4. Fig. 11a shows the post-hoc segment-by-segment qCD estimates for individual runs of the four subjects. Fig. 11b shows the average post-hoc segment-by-segment

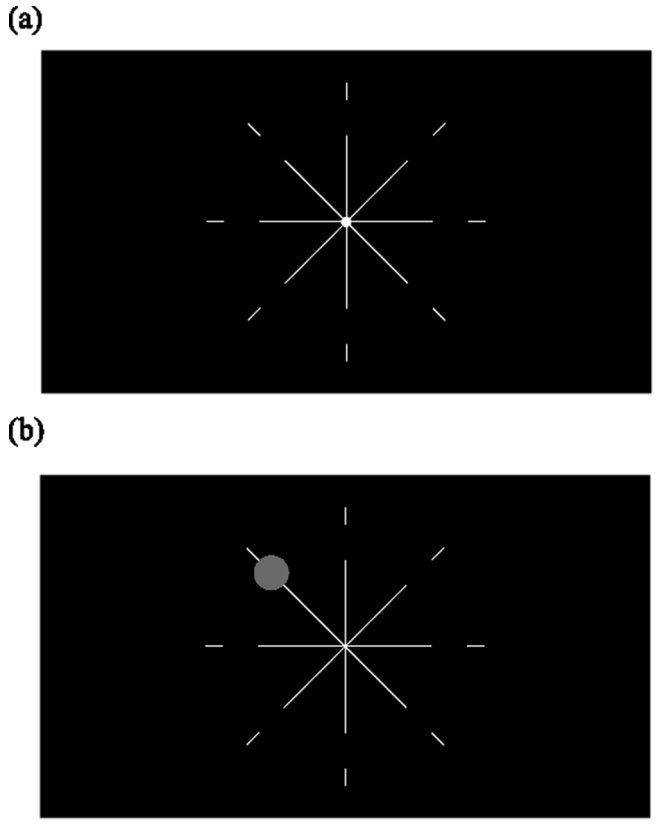


Fig. 8. Illustration of the experimental display. (a) A fixation dot was presented together with the location cues when target was not presented. (b) The fixation dot disappeared when the target was on. The location cues were always on.

threshold estimates of the four independent qCD runs and the threshold estimates from four qFC runs. The average 68.2% HWCI of the post-hoc segment-by-segment threshold estimates from the four qCD runs was 0.036, 0.024, 0.017 and 0.033 $\log_{10}$ units for S1, S2, S3, and S4, respectively. The average RMSEs between the post-hoc segment-by-segment threshold estimates from qCD and the threshold estimates from four qFC runs were 0.074, 0.117, 0.055 and 0.106 $\log_{10}$ units for S1, S2, S3, and S4, respectively. The post-hoc segment-by-segment threshold estimates from the qCD method had higher precision than the threshold estimates from the qFC method (Table 1).

We validated the qCD method using a psychophysical experiment in dark adaptation.[5] It showed good precision - average 68.2% HWCI of 0.036 and 0.028 $\log_{10}$ units for trial-by-trial and post-hoc segment-by-segment threshold estimates, respectively. The estimated time course of perceptual sensitivity change obtained with the qCD also showed good agreement with that obtained with the qFC method. The average RMSE was 0.076 $\log_{10}$ units between trial-by-trial estimates from a single qCD run and estimates from four qFC runs. The average RMSE was 0.126 $\log_{10}$ units between the post-hoc segment-by-segment estimates from a single qCD run and estimates from four qFC runs.

5. Discussion

In this study, we developed the quick Change-Detection (qCD) procedure to accurately, precisely and efficiently measure the trial-by-trial perceptual sensitivity change over time based on the Bayesian adaptive framework (Baek et al., 2016; Hou et al., 2015; King-Smith et al., 1994; Kontsevich & Tyler, 1999; Kujala & Lukka, 2006; Lesmes et al., 2006, 2010, 2015; Prins, 2013; Watson, 2017; Watson & Pelli, 1983). We implemented and tested the qCD method in dark adaptation in which the time course of threshold change is modeled as an exponential decay function. The parameters of the exponential function are characterized by a joint probability distribution. With a prior probability distribution, the qCD selects the stimulus for the next trial that optimizes the expected information gain in the next trial. Then it updates the probability distribution of the parameters by Bayesian inference. The procedure is iterated until a pre-determined stop criterion (e.g., number of trials or precision level) is reached. Based on the joint posterior distribution after each trial, the qCD can provide trial-by-trial estimates of the parameters and the thresholds. Moreover, post-hoc analysis can partition the posterior distributions into segments based on their central tendency. The post-hoc segment-by-segment estimations can further improve the accuracy and precision of the method by utilizing all the information collected during each segment. In summary, the qCD tracks and quantifies the time course of perceptual sensitivity change efficiently with high accuracy and precision by utilizing all available information as well as prior knowledge of the functional form of perceptual sensitivity change over time. Results from both simulations and psychophysical experiment demonstrated the advantage of qCD over two other adaptive procedures (qFC and weighted up-and-down staircase) in measuring the dark adaptation curve. A single run of qCD captures the dark adaptation curve with high accuracy and precision. On the other hand, multiple qFC runs and staircase runs are required to achieve similar accuracy and precision of a single qCD run.

5.1. Extensions and variations of the qCD method

Data collected by other methods such as the staircase method can be rescored with the estimation components of the qCD method to generate trial-by-trial estimates of perceptual sensitivity. However, even after rescoring with the qCD method, the accuracy and precision of the estimated thresholds from the staircase methods would be worse than those obtained directly with the qCD method because the staircase methods do not select the optimal stimulus to maximize the information gain about perceptual sensitivity in each trial.[6]

In this study, weakly informative prior distributions were used for each parameter based on pilot experiments. The efficiency of the qCD can be further improved by informative priors (Gu et al., 2016; Kim, Pitt, Lu, Steyvers, & Myung, 2014), while uniform priors could be used when little prior knowledge is available and/or when the categorization of subpopulations with different time courses is part of the goal (Gu et al., 2016; Kim et al., 2014).

The qCD was tested with an 8AFC task in our study for two reasons. First, high efficiency could be achieved in an 8AFC task since increasing the slope of the psychometric function with increased number of alternative choice resulted in improved efficiency (Hou et al., 2015). Second, the 8AFC task reduces effects of eye movements with the assumption that perceptual sensitivity is approximately the same in the eight locations. However, given the known difference in perceptual sensitivity across the visual field (Strasburger et al., 2011), further work is required to develop procedures that can assess perceptual sensitivity independently at different retinal locations. Moreover, the qCD method

[5] Because the focus of this study is on testing the qCD method and comparing it to qFC, we conducted a binocular dark adaptation experiment with natural pupils. A more careful study focusing on dark adaptation might use monocular tests and/or dilated pupils.

[6] The average accuracy and precision (RMSE: 0.368; HWCI: 0.100; SD: 0.168; all in $\log_{10}$ units) of the threshold estimates from the first staircase run (out of the 10 runs) rescored by the trial-by-trial qCD was much worse than those from the qCD method (RMSE: 0.022; HWCI: 0.038; SD: 0.037; all in $\log_{10}$ units). The average accuracy and precision (RMSE: 0.188; HWCI: 0.039; SD: 0.092; all in $\log_{10}$ units) of the threshold estimates from the first staircase run (out of the 10 runs) rescored by the segment-by-segment qCD was also much worse than those from the qCD method (RMSE: 0.002; HWCI: 0.016; SD: 0.020; all in $\log_{10}$ units).

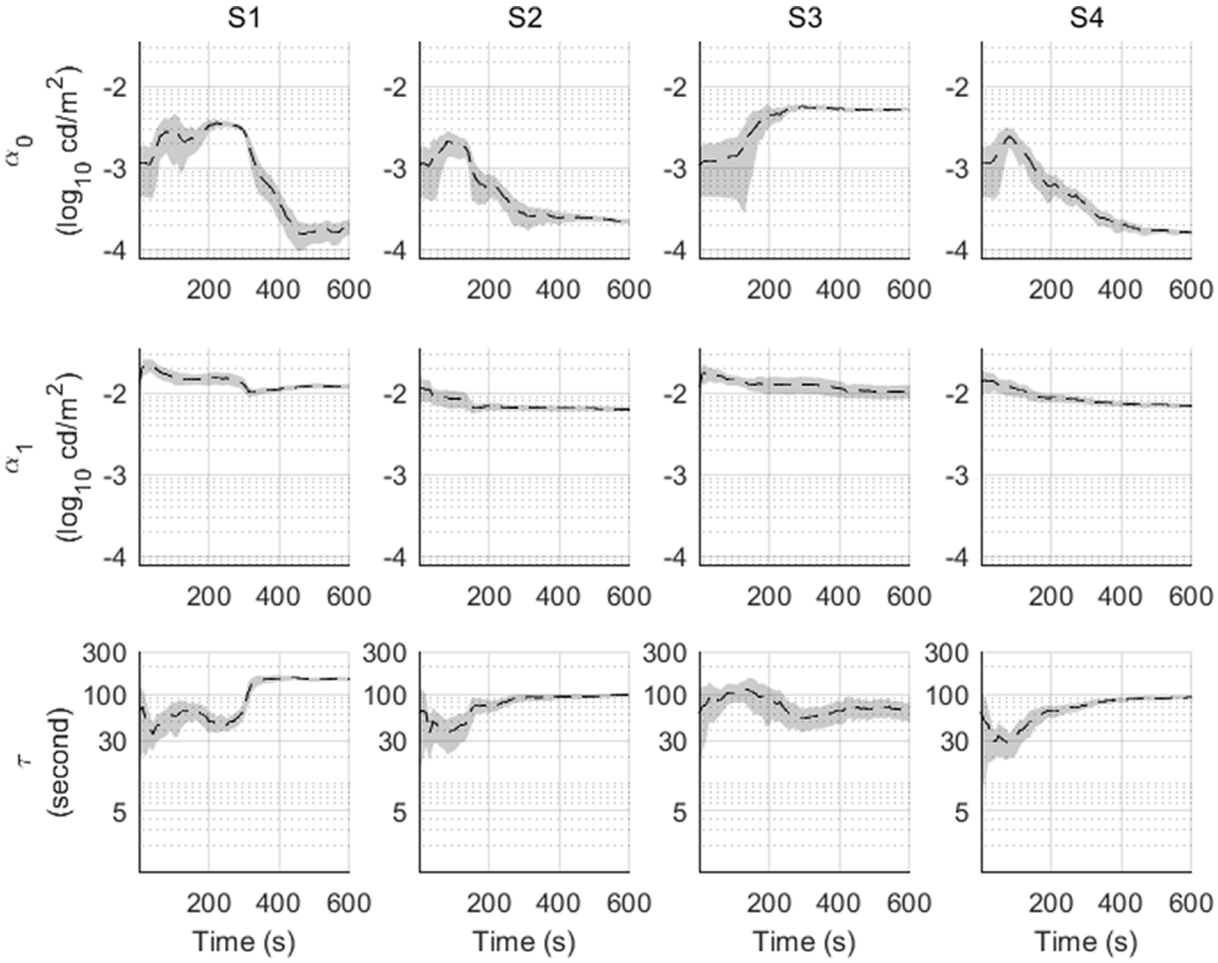


Fig. 9. Trial-by-trial parameter estimates from the qCD (averaged across four runs). Results of the four subjects are presented in different columns. Trial-by-trial parameter estimates of α_0 (top row), α_1 (middle row), and τ (bottom row) of each subject are plotted as functions of time in seconds since the beginning of the dark adaptation. The dashed lines represent the average estimates and the grey shaded areas represent the average 68.2% HWCI.

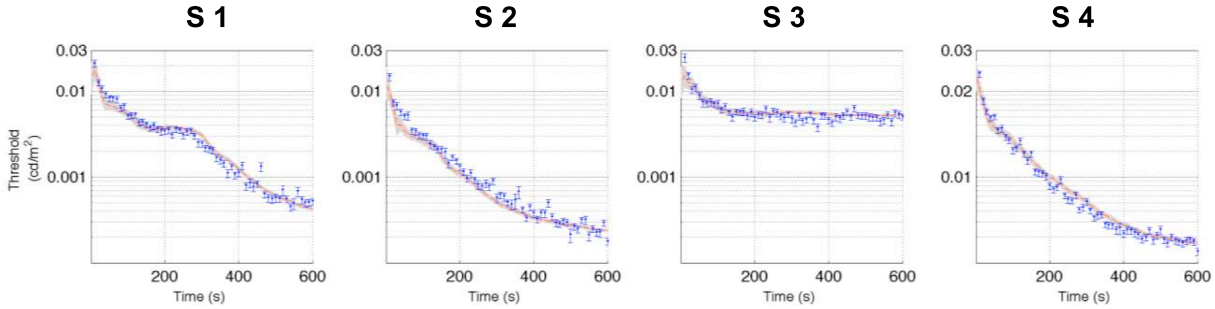


Fig. 10. Trial-by-trial threshold estimates from the qCD. Threshold estimates are plotted as a function of time (in s) elapsed since the beginning of the dark adaptation. The red dashed lines are the estimates from qCD and the grey shaded areas represent the average 68.2% HWCI. The blue dots are the threshold estimates from qFC and the blue error bars are the corresponding average 68.2% HWCI. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

could be used to measure the time course of perceptual sensitivity in tasks other than the 8AFC. For example, in order to measure changes of sensitivity threshold over time at a single location or multiple locations independently in a Yes/No task, the qYN procedure (Lesmes et al., 2015) could be combined with the qCD method. Because the qYN procedure can simultaneously measure both perceptual sensitivity and decision criterion, one added value is the simultaneous measurement of the time courses of perceptual sensitivity and decision criterion during dark adaptation.

Although the exponential function was employed as the functional form of the dark adaptation curve in this study (Lamb & Pugh, 2004; Reuter, 2011), the posterior distributions of the parameters were not constrained across time. In other words, the exponential function can have different parameters at different time points. That is why the qCD was able to track the time course of perceptual sensitivity change that consisted of multiple cascade exponential functions. Simulations (Appendix D) also showed that the qCD can achieve high accuracy (RMSE = 0.032 $\log_{10}$ units) and precision (average 68.2%

Table 1
Precision (68.2% HWCI, $\log_{10}$ units) of the estimated thresholds from the qCD and qFC methods.

		S1	S2	S3	S4
qCD	trial-by-trial	0.037	0.038	0.028	0.040
	segment-by-segment	0.036	0.024	0.017	0.033
qFC	1st run	0.102	0.105	0.100	0.099
	2nd run	0.065	0.074	0.067	0.065
	3rd run	0.048	0.052	0.047	0.046
	4th run	0.040	0.043	0.040	0.038

HWCI = 0.026 $\log_{10}$ units; SD = 0.034 $\log_{10}$ units) in estimating the dark adaptation curves of observers with cascade exponential functions. However, the tracking by the qCD method with the assumption of a single exponential function is non-optimal because the prior in each trial is the posterior of the trial right before it, and if a new exponential function is introduced in a segment of the cascade, the prior in the initial trials of the new segment is not optimal. More research is needed to explore whether explicit modelling of the cascade exponential function as the functional form could further improve the accuracy and precision of the estimates from qCD. In other application domains, different functional forms of the time course, such as the power function, may be used in implementing qCD. When multiple functional forms are possible, hierarchical modeling (Gu et al., 2016; Kim et al., 2014) could be used to test alternative functional forms within the qCD framework. On the other hand, if the functional forms of the time course are unknown, non-parametric methods could be used instead.

The segment-by-segment analysis in the qCD is a post-hoc analysis based on the joint posterior obtained in the end of the experiment to aggregate data from multiple time points. This post-hoc analysis in the qCD might be implemented or improved in other ways, such as the changepoint detection method (for a review, see Jandhyala, Fotopoulos, MacNeill, & Liu, 2013). Because our focus in the current paper is on developing the adaptive measurement procedure, we will leave the optimization of the post-hoc analysis to future research.

5.2. Potential applications in basic and clinical research

Whereas adaptive methods designed to measure thresholds at fixed performance levels over certain test time intervals produces intrinsic bias, the qCD provides estimates of the full time-course with no or very small bias, which is especially critical in assessing processes with rapid changes. This is especially important in estimating the time course of perceptual processes that cannot be measured repeatedly, such as in the case of perceptual learning. For example, conventional methods that only measure the average perceptual sensitivity over some time interval cannot capture fast learning in the early phase of perceptual learning (Kumar & Glaser, 1993). Measuring the average sensitivity over a number of test trials within some test time intervals could result in biased threshold estimates and thus biased conclusions about the learning curve and transfers in perceptual learning.

Although we have only implemented it to assess dark adaptation curve in a single stimulus dimension in this study, qCD can be extended to simultaneously measure the time course of perceptual sensitivity change in multiple stimulus dimensions, such as dark adaptation at multiple spatial locations. Another example is the time course of contrast sensitivity function change. The qCSF (Lesmes et al., 2010) can efficiently, precisely and accurately measure the CSF in a few minutes. Combining the qCD with the qCSF could allow us to measure the detailed time course of CSF changes over time.

Finally, past research has shown that abnormal dark adaptation may be an early sign of AMD (Jackson & Edwards, 2008; Jackson et al., 2006; Owsley et al., 2016). Given its precision, efficiency and accuracy, qCD may find its application in early detection of AMD.

5.3. Limitations

In the current implementation of the qCD method, we used a fixed slope of 3.8959 and a lapse rate of 0.04. To investigate the effects of mismatched slope and lapse rate on the performance of the qCD method, we re-simulated the performance of the qCD method for simulated Observer 1 with mismatched slopes of 2.5 and 5.5, and mismatched lapses rates of 0 and 0.08. The mis-matched slopes and lapse rates only had some very small effects on the performance of the qCD method: The various mismatches increased the RMSE of the estimated thresholds from the segment-by-segment qCD method by a maximum of 0.029 $\log_{10}$ units, and the time to achieve 0.1 $\log_{10}$ units RMSE and 68.2% HWCI by a maximum of 8.4% and 4.6%, respectively. Although we assumed a fixed slope in this paper, we can re-formulate the qCD method to estimate the slope of the psychometric function. It involves adding an extra parameter to the joint probability distribution. If we assume that the slope does not vary over time, estimating it would probably not drastically change the performance of the qCD method.

The applicability of the qCD method depends on its rate of posterior convergence relative to the rate of the to-be-estimated perceptual sensitivity change. In order to obtain accurate and precise estimates of the time course of perceptual sensitivity change, the qCD method should only be used to measure the time course of perceptual sensitivity changes that are slower or at most compatible to the convergence rate of the joint posterior distribution of the parameters of the time course modeled in the qCD procedure. The convergence of the posterior distribution depends on many factors, including the prior distribution, stimulus selection, model complexity, and experimental design. Successful implementation and optimization of the qCD method would require careful consideration of all these factors in each application on a case-by-case basis.

In the qCD method, the prior distribution is constructed based on *a priori* knowledge about the time course of perceptual sensitivity change for the studied population. The more informative the prior is, that is, the more we know about the properties of the population under study, the faster the posterior converges. Kim et al. (2014) developed a hierarchical Bayes extension of adaptive design optimization (HADO) that provided a judicious way to exploit two complementary schemes of inference, with past and future data, to speed up posterior convergence. In the procedure, past data are used to construct and update priors for multiple subject categories, and future data are used to simultaneously categorize new subjects and estimate their individual posterior distributions. Future development of the qCD procedure could use a similar hierarchical Bayesian approach.

Optimizing stimulus selection can increase the information obtained throughout the test and thus increase the convergence rate of the posterior. In this study, the stimulus is selected by a one-step-ahead search algorithm (Kujala & Lukka, 2006; Lesmes et al., 2006). Further improvement of the qCD method can be achieved by multiple-step-ahead search based on dynamic programming (Kim, Pitt, Lu, & Myung, 2017). Optimal stimulus selection is more critical when more than one stimulus dimension is involved, and/or the model is complex.

The complexity of the model of the time course of perceptual sensitivity change, reflected in the number of parameters and the complexity of the parameter space, also affects posterior convergence. Finding regularities in the time course of perceptual sensitivity change could reduce model complexity and improve posterior convergence. For example, we recently found that contrast sensitivity functions measured in different luminance conditions can be described by the same shape (Hou et al., 2017). Such regularity can be exploited to reduce the number of parameters in modeling the time course of contrast sensitivity function change in dark adaptation.

The qCD method, like any other parametric method, is susceptible to model mis-specification. A relatively minor mis-specification could affect the performance of the method in terms of slower posterior convergence. Major mis-specifications could make the method

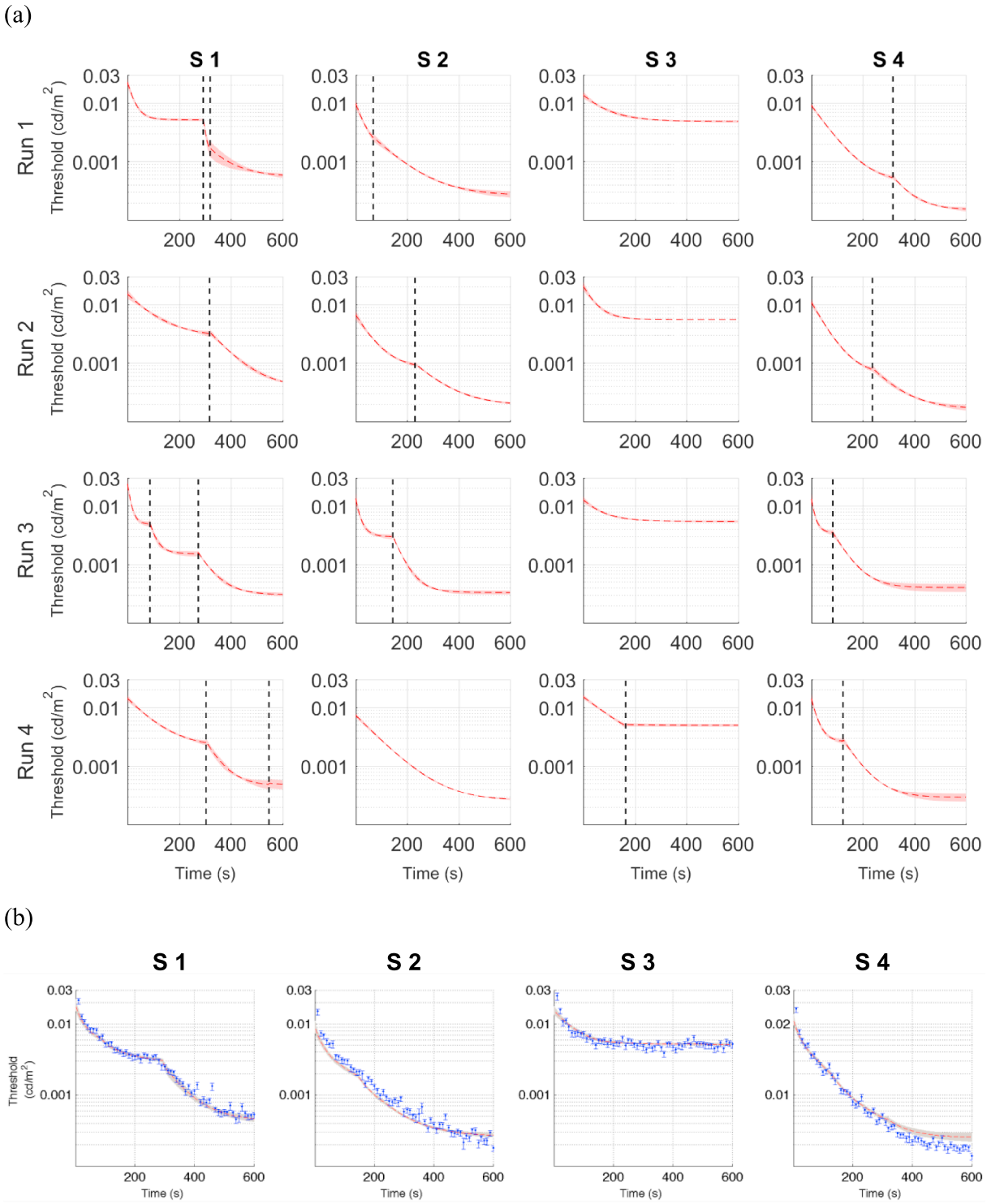


Fig. 11. Estimated post-hoc segment-by-segment dark adaptation curves. Results of the four subjects are presented in different columns. (a) Post-hoc segment-by-segment threshold estimates from individual qCD runs. The red dashed lines represent the estimates from the qCD and the pink shaded areas represent the average 68.2% HWCI. Different segments (if any) are separated by vertical black dashed lines. (b) Average estimated post-hoc segment-by-segment dark adaptation curves. Threshold estimates are plotted as a function of time (in s) elapsed since the beginning of the dark adaptation. The red dashed lines represent the estimates from the qCD and the grey shaded areas represent the average 68.2% HWCI. The blue dots represent the threshold estimates from qFC and the blue error bars represent the average 68.2% HWCI. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

completely inadequate. This is why it is important to understand the general properties of the perceptual phenomena before applying the qCD method. Our choice of an exponential function in dark adaptation is based on the literature (Gaffney et al., 2014; Hecht et al., 1937, 1935; Hollins & Alpern, 1973; Lamb, 1981; Mote & Riopelle, 1951; Pugh, 1975). While the qCD with a single exponential function form showed high efficiency, accuracy and precision in measuring both single exponential and cascade exponential, the delays and deviations shown in simulated Observers 4 and 5 (Figs. D1 and D2) indicated sub-optimal convergence rate. While explicitly modeling the cascade exponential as

the functional form might further improve the accuracy and precision of the estimates, the benefit should be weighed against the cost of the increased model complexity. In applying the qCD to measure other perceptual phenomena, we may use different functional forms.

In addition, a good experimental design could improve the efficiency and convergence rate of the adaptive method. For example, the posterior of the contrast sensitivity function converges much faster in an 8AFC procedure than in a 2AFC procedure (Hou et al., 2015) because each trial in an 8AFC procedure provides more information. A good experimental design could also improve subjects' response rate and quality and therefore increase the rate of information accumulation and posterior convergence.

The aim of this paper is to develop the general algorithm of the qCD method and provide its first demonstration in dark adaptation. To develop an optimized method to measure the time course of perceptual sensitivity change, the specifics of the qCD implementation (factors which affect the convergence rate of the posterior) need to be tailored in each case. We have set up the simulated observers based on the dark adaptation literature and results from pilot experiments in the current study. After we learn more about the population under study, the qCD method can be more thoroughly evaluated by simulating all possible observers. Our experience in developing the qCSF method may be instructive. Since its first publication in 2010 (Lesmes et al., 2010), the qCSF has been improved over the years based on data collected from thousands of observers in many different categories of eye diseases (Dorr et al., 2015; Hou et al., 2010; Lesmes et al., 2016). At the current stage of qCD development, we have only very limited experience with the method. More data on dark adaptation, for example, data from different age groups and eye disease populations, are required to further optimize the method.

6. Conclusions

In conclusion, we have developed the qCD method to efficiently, accurately and precisely assess the trial-by-trial time course of perceptual sensitivity change. In both simulations and a psychophysics experiment, we demonstrated the high efficiency, accuracy, and precision of the method in measuring the dark adaptation curve. The qCD provides a procedure to characterize the detailed time course of perceptual sensitivity change in both basic research and clinical applications.

Acknowledgments

This research was supported by the National Eye Institute (EY021553).

Declaration of interest

YZ, LAL and ZLL have an intellectual property interest in methods for measuring behavioural changes of processes (PCT/US18/21944). In addition, LAL and ZLL have intellectual property interests in methods for measuring and applying contrast sensitivity functions (US 7938538, WO2013170091, PCT/US2015/028657), and equity interest in Adaptive Sensory Technology, Inc. (San Diego, CA). LAL holds employment in AST.

Appendix A:. The quick Forced choice (qFC) method

The algorithm

Based on the Bayesian adaptive framework, the qFC measures the average perceptual sensitivity over multiple trials by updating, trial by trial, a probability distribution of perceptual sensitivity, without explicitly modeling changes in perceptual sensitivity over trials. In this implementation, a single threshold is estimated from qFC in every 10 s (Fig. A1).

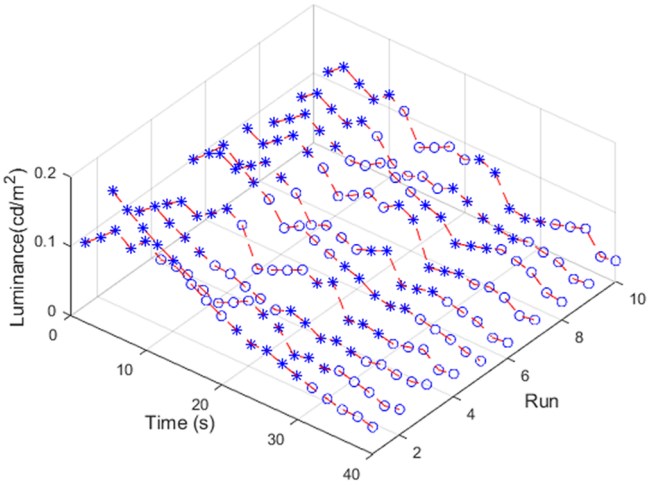


Fig. A1. An example of the simulated trial sequence in the first 40 s of ten repeated staircase runs used to estimate the time course of perceptual sensitivity change. In this procedure, an adaptive procedure (qFC or staircase) is used to measure the average perceptual sensitivity in every 10 s. In the first run of the experiment, the conventional staircase procedure is used (illustrated by the asterisks and open circles connected by the dashed line closest to the "Time" axis). That is, a predetermined stimulus intensity is used in the first trial, and the staircase procedure determines the stimulus intensities in all subsequent trials. This sets the "initial" stimulus intensity at each time delay. In all subsequent runs, the stimulus intensities in each 10-s window are determined by the staircase procedure based on stimulus intensities tested within the 10-s window at the same time delay across runs. Bands of asterisks or open circles across runs illustrate the trial sequences in the staircases.

The qFC method consists of four steps. (1) A broad prior distribution of luminance detection threshold is set. (2) The stimulus luminance for the next trial is selected to optimize the expected information gain on the threshold. (3) The posterior distribution of the threshold is updated by the Bayes' rule based on the observer's response after each trial. (4) Step (2) and (3) are repeated until the stop criterion is met (eg. pre-determined number of trials).

The luminance detection threshold α_{FC} corresponding to $d' = 1.5$ is assumed to be constant for the duration of the measurement. A broad prior distribution $p_0(\theta_{FC})$ is defined in the parameter space $\theta_{FC} = \alpha_{\mathrm{FC}}$ and a one-dimension stimulus luminance space X covers all possible luminance $x \in X$.

A Weibull function is used to approximate the psychometric function in qFC:

$$p_t(r = 1|\theta_{FC}, x) = (1 - \lambda)(g + (1 - g)\left(1 - \exp\left(-\left(\frac{x}{\alpha_w(\theta_{FC})}\right)^{\gamma}\right)\right)) + \lambda g \tag{A1}$$

$$\log_{10}(\alpha_w(\theta_{FC}))=\log_{10}(\alpha_{\mathrm{FC}}) - \frac{1}{\gamma}\log_{10}\left(\log\left(\frac{1-g}{1-p_{1.5}}\right)\right). \tag{A2}$$

The qFC updates the distribution of the sensitivity parameter $p_0(\theta_{FC} = \alpha_{\mathrm{FC}})$. The remaining steps of Bayesian update and one step ahead stimulus search are the same as those in the qCD.

In the qFC simulations, thresholds were estimated every 10s. The prior distribution $p_0(\theta_{FC} = \alpha_{\mathrm{FC}})$ was updated only during the corresponding 10-s interval which resulted in 60 posterior distributions after one run of the 600-s experiment. The 60 posterior distributions obtained in one run were used as the priors in the next run. The qFC was iterated ten times (runs) in each simulation. Therefore, there were 10 iterated runs in one qFC simulation run, while qCD simulation runs were all independent and did not iterate.

Evaluation method

The estimated parameter $\hat{\alpha}_{\{n_{FC}m\}(t_k)}$ is defined as the threshold estimate at time t_k in the end of the k^{th} 10-s measurement interval in the n_{FC}th iterated run in the m^{th} simulation run. In other words, $\hat{\alpha}_{\{n_{FC}m\}(t_k)}$ is the threshold estimate of the k^{th} interval in the ${n_{FC}}^{\mathrm{th}}$ iterated run in the m^{th} simulation run. The $RMSE(n_{FC})$ of the threshold estimates in the ${n_{FC}}^{\mathrm{th}}$ iterated run across M simulation runs is defined as a function of n_{FC}:

$$RMSE(n_{FC}) = \sqrt{\frac{\sum_k \left[\log_{10}(\bar{\hat{\alpha}}_{\{n_{FC}\}(t_k)}) - \log_{10}(\alpha_{observer(t_k)})\right]^2}{K}}\ , \tag{A3}$$

$$\bar{\hat{\alpha}}_{\{n_{FC}\}(t_k)} = \sum_m \hat{\alpha}_{\{n_{FC}m\}(t_k)}/M\ , \tag{A4}$$

where $\bar{\hat{\alpha}}_{\{n_{FC}\}(t_k)}$ is the average of the threshold estimate of the k^{th} interval in the ${n_{FC}}^{\mathrm{th}}$ iterated run across M simulations; $\hat{\alpha}_{\{n_{FC}m\}(t_k)}$ is the estimated threshold of k^{th} interval in the ${n_{SC}}^{\mathrm{th}}$ iterated run in the m^{th} simulation.

The standard deviation $SD(n_{FC})$ is defined as a function of n_{FC}:

$$SD(n_{FC}) = \sqrt{\frac{\sum_k \sum_m \left[\log_{10}(\hat{\alpha}_{\{n_{FC}m\}(t_k)}) - \log_{10}(\bar{\hat{\alpha}}_{\{n_{FC}\}(t_k)})\right]^2}{M \times K}}\ . \tag{A5}$$

Simulation

In the qFC, the prior distribution, $p_0(\theta_{FC})$, was defined by the hyperbolic secant function:

$$p_0(\theta_{FC})=\mathrm{sech}(\alpha_{w,\ \mathrm{confidence}} \times (\log_{10}(\alpha_{w,\ \mathrm{confidence}}) - \log_{10}(\alpha_{w,\ \mathrm{mode}}))\ , \tag{A6}$$

where$\alpha_{w,\ \mathrm{confidence}} = 3.16$ and $\alpha_{w,\mathrm{mode}} = 0.0023$. In qFC, parameter space included 120 log-linearly spaced α_w values between 0.000075 to 0.075 cd/m^2.

The probability of a correct response $p_t(r = 1|\theta_{FC}, x)$ was calculated by Eq. (A1) in qFC. The simulations were performed 1000 times, thus in total 1000 independent simulations of 10 qFC runs were obtained for each observer.

Fig. A2a shows the estimated thresholds after the 1$^{\mathrm{st}}$ and 10$^{\mathrm{th}}$ qFC run for the three simulated observers. Fig. A2b shows the accuracy (RMSE) and the precision (SD and average 68.2% HWCI). For simulated Observer 1, the RMSE (n_{FC} = 1, 2, 3, 10) was 0.204, 0.075, 0.043, and 0.028 $\log_{10}$ units after 1$^{\mathrm{st}}$, 2$^{\mathrm{nd}}$, 3$^{\mathrm{rd}}$ and 10$^{\mathrm{th}}$ run, respectively; the SD (n_{FC} = 1, 2, 3, 10) was 0.223, 0.128, 0.086, and 0.032 $\log_{10}$ units after the 1$^{\mathrm{st}}$, 2$^{\mathrm{nd}}$, 3$^{\mathrm{rd}}$ and 10$^{\mathrm{th}}$ run, respectively; the average 68.2% HWCI was 0.193, 0.105, 0.073, and 0.032 $\log_{10}$ units after the 1$^{\mathrm{st}}$, 2$^{\mathrm{nd}}$, 3$^{\mathrm{rd}}$ and 10$^{\mathrm{th}}$ run, respectively. For simulated Observer 2, the RMSE (n_{FC} = 1, 2, 3, 10) was 0.186, 0.067, 0.035, and 0.018 $\log_{10}$ units after the 1$^{\mathrm{st}}$, 2$^{\mathrm{nd}}$, 3$^{\mathrm{rd}}$ and 10$^{\mathrm{th}}$ run, respectively; the SD (n_{FC} = 1, 2, 3, 10) was 0.210, 0.121, 0.081, and 0.031 $\log_{10}$ units after the 1$^{\mathrm{st}}$, 2$^{\mathrm{nd}}$, 3$^{\mathrm{rd}}$ and 10$^{\mathrm{th}}$ run, respectively; the average 68.2% HWCI was 0.180, 0.101, 0.071, and 0.032 $\log_{10}$ units after the 1$^{\mathrm{st}}$, 2$^{\mathrm{nd}}$, 3$^{\mathrm{rd}}$ and 10$^{\mathrm{th}}$ run, respectively. For simulated Observer 3, the RMSE (n_{FC} = 1, 2, 3, 10) was 0.129, 0.048, 0.027, and 0.014 $\log_{10}$ units after the 1$^{\mathrm{st}}$, 2$^{\mathrm{nd}}$, 3$^{\mathrm{rd}}$ and 10$^{\mathrm{th}}$ run, respectively; the SD (n_{FC} = 1, 2, 3, 10) was 0.177, 0.104, 0.073, and 0.031 $\log_{10}$ units after the 1st, 2nd, 3rd and 10th run, respectively; the average 68.2% HWCI was 0.153, 0.091, 0.066, and 0.031 $\log_{10}$ units after the 1$^{\mathrm{st}}$, 2$^{\mathrm{nd}}$, 3$^{\mathrm{rd}}$ and 10$^{\mathrm{th}}$ run, respectively.

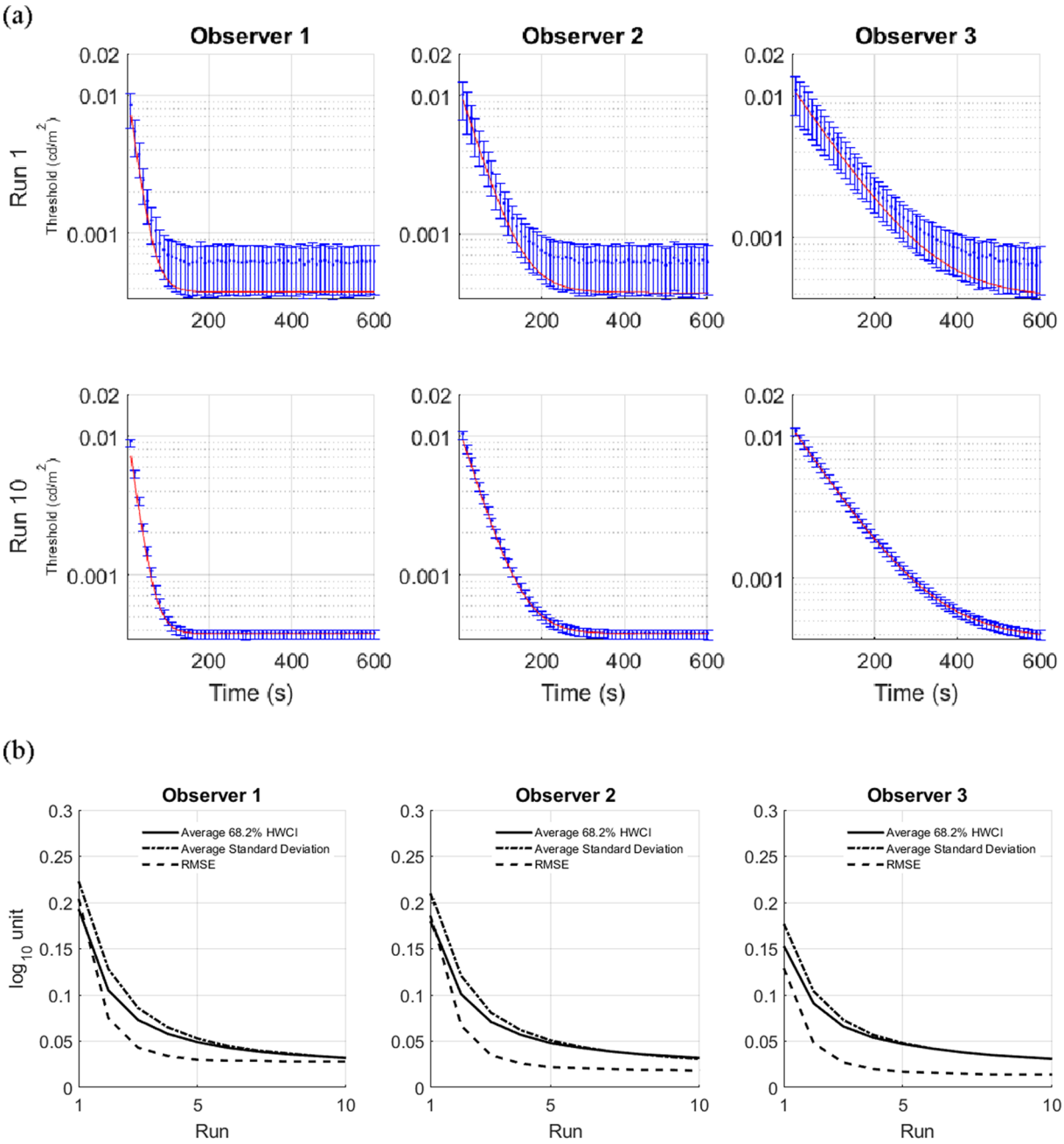


Fig. A2. (a) Threshold estimates in the qFC simulations. Threshold estimates are plotted as a function of time. Results of the three simulated observers are presented in different columns. Estimates after the 1st and 10th runs are plotted in different rows. The solid red lines are the true thresholds of the simulated observers, the blue dots are the threshold estimates from qFC and the blue error bars are the average 68.2% HWCI. (b) Accuracy and precision of the threshold estimates in the qFC simulations. Results of the three simulated observers are presented in different columns. Solid lines represent average 68.2% HWCI; dotted lines represent SD; dashed lines represent RMSE.

Appendix B:. The weighted up-down staircase method

The algorithm

The staircase procedure is designed to converge to a certain accuracy level after a sufficient number of trials. In our implementation, a single threshold is estimated by the staircase in every 10 s (Fig. A1). Same as in the qFC, there were 10 iterated runs in one staircase simulation run. In order to compare the performance by different methods, perceptual sensitivity at the same performance level (d′) needs to be measured. The weighted up-down method (Kaernbach, 1991) was employed to measure the thresholds at the accuracy level $p_{1.5} = 0.553$, the probability correct when $d' = 1.5$ in the 8AFC task. Each correct response results in a decrease, S_{down}, in the luminance level and each incorrect response results in an increase, S_{up}. In order to measure the threshold at $p_{1.5}$, S_{down} and S_{up} are constrained by the equation:

$$S_{down} p_{1.5} = S_{up}(1 - p_{1.5}) \quad \text{(B1)}$$

A Weibull function was used to approximate the psychometric function in the staircase:

$$p_t(r = 1|\theta_{SC}, x) = (1 - \lambda)(g + (1 - g)(1 - \exp(-\left(\frac{x}{\alpha_w}\right)^{\gamma}))) + \lambda g\ , \tag{B2}$$

$$\log_{10}(\alpha_w) = \log_{10}(\alpha_{SC}) - \frac{1}{\gamma}\log_{10}\left(\log\left(\frac{1-g}{1-p_{1.5}}\right)\right), \tag{B3}$$

where $\theta_{SC} = \alpha_{SC}$ and x is the luminance tested at the current step by the staircase.

Evaluation method

For each staircase, the average of the reversals was used as the estimated threshold at t_k in the end of each 10-s interval. The first two reversals were discarded. The last reversal point was also discarded if the number of the reversal points was odd.

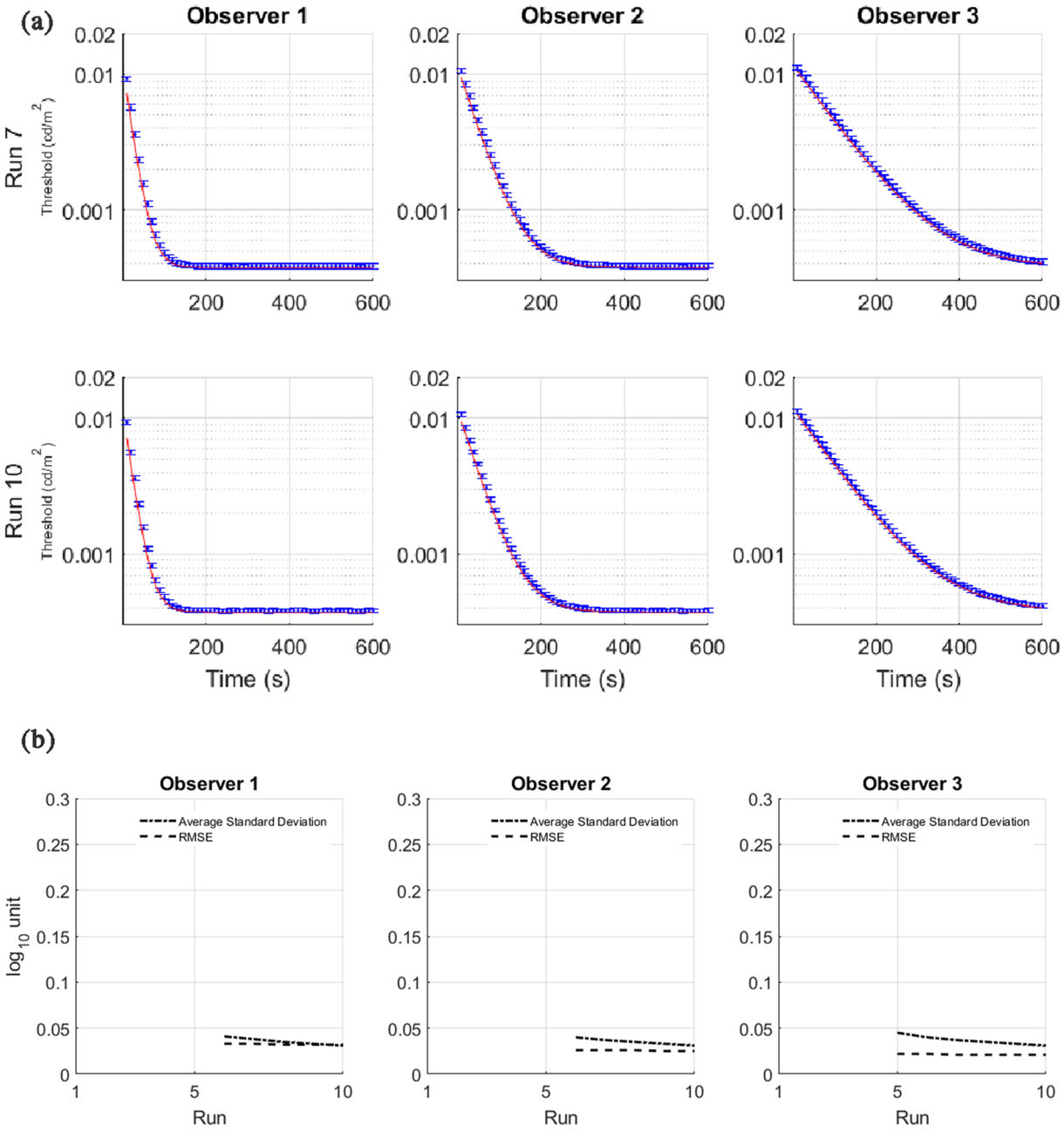


Fig. B1. (a) Threshold estimates in the staircase simulations. Threshold estimates are plotted as a function of time. Results of the three simulated observers are presented in different columns. Estimates after 7th and 10th run are plotted in different rows. The solid red lines are the true values of the simulated observers, the blue dots are the threshold estimates from staircase and the blue error bars are the SD. (b) Accuracy and precision of the threshold estimates in the staircase simulations. Results of the three simulated observers are presented in different columns. The dotted lines represent SD and the dashed lines represent RMSE.

The $RMSE(n_{SC})$ of threshold estimates in the $n_{SC}{}^{\text{th}}$ iterated run is defined as a function of n_{SC}:

$$RMSE(n_{SC}) = \sqrt{\frac{\sum_k [\log_{10}(\bar{\hat{\alpha}}_{\{n_{SC}\}(t_k)}) - \log_{10}(\alpha_{observer(t_k)})]^2}{K}}\ , \tag{B4}$$

$$\bar{\hat{\alpha}}_{\{n_{SC}\}(t_k)} = \sum_m \hat{\alpha}_{\{n_{SC}m\}(t_k)}/M\ , \tag{B5}$$

where $\bar{\hat{\alpha}}_{\{n_{SC}\}(t_k)}$ is the average of the threshold estimate of k^{th} interval in the $n_{SC}{}^{\text{th}}$ iterated run across M simulations; $\hat{\alpha}_{\{n_{SC}m\}(t_k)}$ is the estimated threshold of k^{th} interval in the $n_{SC}{}^{\text{th}}$ iterated run in the m^{th} simulation.$SD(n_{SC}) = \sqrt{\frac{\Sigma_k \Sigma_m [\log_{10}(\hat{\alpha}_{\{n_{SC}m\}(t_k)}) - \log_{10}(\bar{\hat{\alpha}}_{\{n_{SC}\}(t_k)})]^2}{M \times K}}$.

The standard deviation $SD(n_{SC})$ is defined as a function of n_{SC}:

$$SD(n_{SC}) = \sqrt{\frac{\sum_k \sum_m [\log_{10}(\hat{\alpha}_{\{n_{SC}m\}(t_k)}) - \log_{10}(\bar{\hat{\alpha}}_{\{n_{SC}\}(t_k)})]^2}{M \times K}}\ . \tag{B6}$$

Simulation

In the staircase, the range of possible luminance was from 0.000075 to 0.075 cd/m^2. The starting luminance was 0.0075 cd/m^2. The step sizes were 0.05 and 0.062 $\log_{10}$ units for S_{down} and S_{up}, respectively.

In the staircase method, sixty staircases were run, one for each 10-s interval. In the first run of 600-s experiment, the luminance tested was updated continuously. Starting from the second run, each staircase was updated independently only during the corresponding 10-s interval. The staircase method was iterated ten times (runs) in each simulation.

The probability of a correct response $p_t(r = 1|\theta_{SC}, x)$ was calculated by Eq. (B2) in the staircase. The simulations were performed 1000 times, thus in total 1000 independent simulations of 10 staircase runs were obtained for each observer.

The staircases did not converge (with at least 4 reversals) until the 6$^{\text{th}}$, 6$^{\text{th}}$, and 5$^{\text{th}}$ run for simulated Observer 1, 2, and 3, respectively. Fig. B1a shows the estimated thresholds after the 7$^{\text{th}}$ and 10$^{\text{th}}$ staircase run for the three observers. Fig. B1b shows the accuracy (RMSE) and the precision (SD). For simulated Observer 1, the RMSE ($n_{SC} = 10$) and SD ($n_{SC} = 10$) was 0.032 and 0.031 $\log_{10}$ units after the 10$^{\text{th}}$ run. For simulated Observer 2, the RMSE ($n_{SC} = 10$) and SD ($n_{SC} = 10$) was 0.025 and 0.031 $\log_{10}$ units after the 10$^{\text{th}}$ run. For simulated Observer 3, the RMSE ($n_{SC} = 10$) and SD ($n_{SC} = 10$) was 0.021 and 0.031 $\log_{10}$ units after the 10$^{\text{th}}$ run.

Appendix C:. Post-hoc segment-by-segment analysis in qCD

We can partition the trial-by-trial posterior distributions based on their central tendency across time. The distance between the central tendency of the posterior distributions, $p_{t_{n_1}}(\vec{\theta}|r_{n_1}, x_{n_1})$ and $p_{t_{n_2}}(\vec{\theta}|r_{n_2}, x_{n_2})$, at two time points, t_{n_1} and t_{n_2} is quantified by a modified Mahalanobis distance MD:

$$MD(n_1, n_2) = \sqrt{\vec{y}\,\boldsymbol{C}^{-1}\vec{y}^{T}}\ , \tag{C1.1}$$

$$\boldsymbol{C} = [\boldsymbol{C}_{n_1} + \boldsymbol{C}_{n_2}]/2\ , \tag{C1.2}$$

where $\vec{y} = [y_1, y_2, y_3]$, $y_i = \bar{\theta}_{i,n_1} - \bar{\theta}_{i,n_2}$, $\theta_i = \alpha_0, \alpha_1,\ and\ \tau$,for i = 1, 2 and 3, respectively; $\vec{y}^{T}$ is the transpose of $\vec{y}$; $\boldsymbol{C}^{-1}$ is the inverse of the matrix $\boldsymbol{C}$; $\boldsymbol{C}_{n_b}$(b = 1 or 2) is the 3 × 3 covariance matrix of the posterior $p_{t_{n_b}}(\vec{\theta}|r_{n_b}, x_{n_b})$ for which the diagonal elements c_{n_b,a_1a_1} and the off-diagonal elements $c_{n_b,ij}$ (i, $j \in [1, 2, 3]$, $i \neq j$) are defined as:

$$c_{n_b,ii} = Var(\theta_{i,n_b}) = \sum_{\vec{\theta}} \left(\theta_i - \bar{\theta}_{i,n_b}\right)^2 \bullet p_{t_{n_b}}(\vec{\theta}|r_{n_b}, x_{n_b})\ , \tag{C1.3}$$

$$c_{n_b,ij} = Cov(\theta_{i,n_b}, \theta_{j,n_b}) = \sum_{\vec{\theta}} \left(\theta_i - \bar{\theta}_{i,n_b}\right)\left(\theta_j - \bar{\theta}_{j,n_b}\right) \bullet p_{t_{n_b}}(\vec{\theta}|r_{n_b}, x_{n_b})\ . \tag{C1.4}$$

The null hypothesis is that the posterior distributions $p_{t_{n_1}}(\vec{\theta}|r_{n_1}, x_{n_1})$ and $p_{t_{n_2}}(\vec{\theta}|r_{n_2}, x_{n_2})$ are the same. We reject the null hypothesis when

$$MD(n_1, n_2) > MD_0\ , \tag{C2}$$

where MD_0 is a predetermined criterion.

To describe how we partition the posterior distributions of the dark adaptation curve, we first introduce some notations: u_l is the trial number in the experiment, where l and u_l refer to the l^{th} segment and the $u_l{}^{\text{th}}$ trial in the l^{th} segment, respectively. U_l is the total number of trials in the l^{th} segment and L is the number of segment(s) of the entire dark adaptation curve. To partition the dark adaptation curve into segments, we start from the last trial of the entire experiment, U_L. The MD between $p_{t_{U_L}}(\vec{\theta}|r_{U_L}, x_{U_L})$ and $p_{t_{u_L}}(\vec{\theta}|r_{u_L}, x_{u_L})$, the posterior distributions of the last trial and the posterior distributions of the previous trials, are calculated until the last trial of the $L-1^{\text{th}}$ segment, U_{L-1}, is found:

$$MD(p_{t_{U_L}}(\vec{\theta}|r_{U_L}, x_{U_L}), p_{t_{u_L}}(\vec{\theta}|r_{u_L}, x_{u_L})) \leq MD_0, \forall\ u_L \in [1, U_L - 1] \tag{C3.1}$$

$$MD(p_{t_{U_L}}(\vec{\theta}|r_{U_L}, x_{U_L}), p_{t_{U_L-1}}(\vec{\theta}|r_{U_L-1}, x_{U_L-1})) > MD_0 \tag{C3.2}$$

We then repeat the procedure to find all the segments. Therefore, for the l^{th} segment ($l > 1$), we have

$$MD(p_{t_{U_l}}(\vec{\theta}|r_{U_l}, x_{U_l}), p_{t_{u_l}}(\vec{\theta}|r_{u_l}, x_{u_l})) \leq MD_0, \forall\ u_l \in [1, U_l - 1]\ , \tag{C3.3}$$

$$MD(p_{t_{U_l}}(\vec{\theta}|r_{U_l}, x_{U_l}), p_{t_{U_{l-1}}}(\vec{\theta}|r_{U_{l-1}}, x_{U_{l-1}})) > MD_0\,. \tag{C3.4}$$

Because the dark adaptation curve is continuous, we impose the following constraint on the posterior distributions between segments:

$$p_{u_l}(\alpha_0 + \alpha_1) = p_{t_{U_{l-1}}}\left(\alpha_{t_{U_{l-1}}}\right), \tag{C3.5}$$

where $p_{u_l}(\alpha_0 + \alpha_1)$ is the marginal prior distribution of $\alpha_0 + \alpha_1$ at trial u_l (u_lth trial in the lth segment) and $p_{t_{U_{l-1}}}(\alpha_{t_{U_{l-1}}})$ is the probability distribution of the threshold estimate at $t_{n\{U_{l-1},l-1\}}$ of the $l-1$th segment.

Following segmentation, we use the posterior in the last trial of each segment $p_{t_{U_l}}(\vec{\theta}|r_{U_l}, x_{U_l})$ to compute the estimated thresholds in the entire segment.

Appendix D:. Performance of qCD in measuring cascade exponential dark adaptation curves

Method

We simulated the qCD method in characterizing two additional observers whose luminance thresholds change as a cascade exponential function in dark adaptation:

$$\alpha_{cascade}(\vec{\theta}, t) = \begin{cases} \alpha_{01} + \alpha_{11}\exp(-t/\tau_1), & t \le t_c \\ \alpha_{02} + \alpha_{12}\exp(-(t-t_c)/\tau_2), & t > t_c \end{cases}, \tag{D1}$$

where $\alpha_{02} = \alpha_{01} + \alpha_{11}\exp(-t_c/\tau) - \alpha_{12}$, such that the function is continuous at t_c. We set the parameters of Observer 4 to $\vec{\theta}_{observer,4} = (\alpha_{11}, \tau_1, \alpha_{01}, \alpha_{12}, \tau_2, t_c) = (0.0150, 50, 0.0030, 0.0023, 85, 270)$, and Observer 5 to $\vec{\theta}_{observer,5} = (\alpha_{11}, \tau_1, \alpha_{01}, \alpha_{12}, \tau_2, t_c) = (0.0150, 50, 0.0030, 0.0023, 200, 270)$.

For Observers 4, we simulated a 600-s experiment. For Observer 5, we simulated a 1200-s experiment. A 2-s inter-trial-interval (ITI) was implemented in all the simulations (Footnote 3). For each run of the simulated 600-s experiment, there were 300 trials; for each run of the 1200-s experiment, there were 600 trials.

Results

In the trial-by-trial procedure, the posterior distribution after each trial is used to estimate the threshold in that trial. Fig. D1 shows the trial-by-trial threshold estimates from the qCD method for Observers 4 and 5. For Observer 4, the RMSE, SD, and average 68.2% HWCI of the estimated trial-

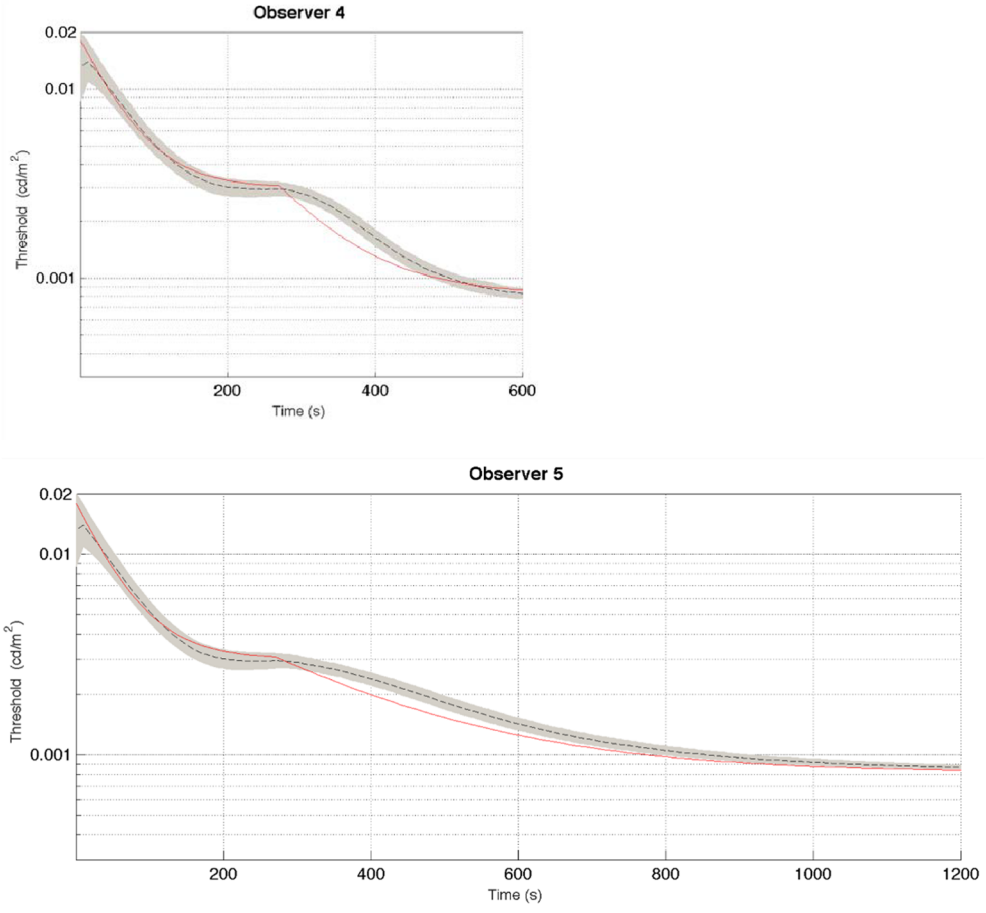


Fig. D1. Trial-by-trial threshold estimates for simulated Observers 4 and 5. Threshold estimates are plotted as functions of time (in s) elapsed since the beginning of dark adaptation. The solid red lines represent the true thresholds of the simulated observers, the black dashed lines represent the average estimates, and the grey shaded areas represent the average 68.2% HWCI.

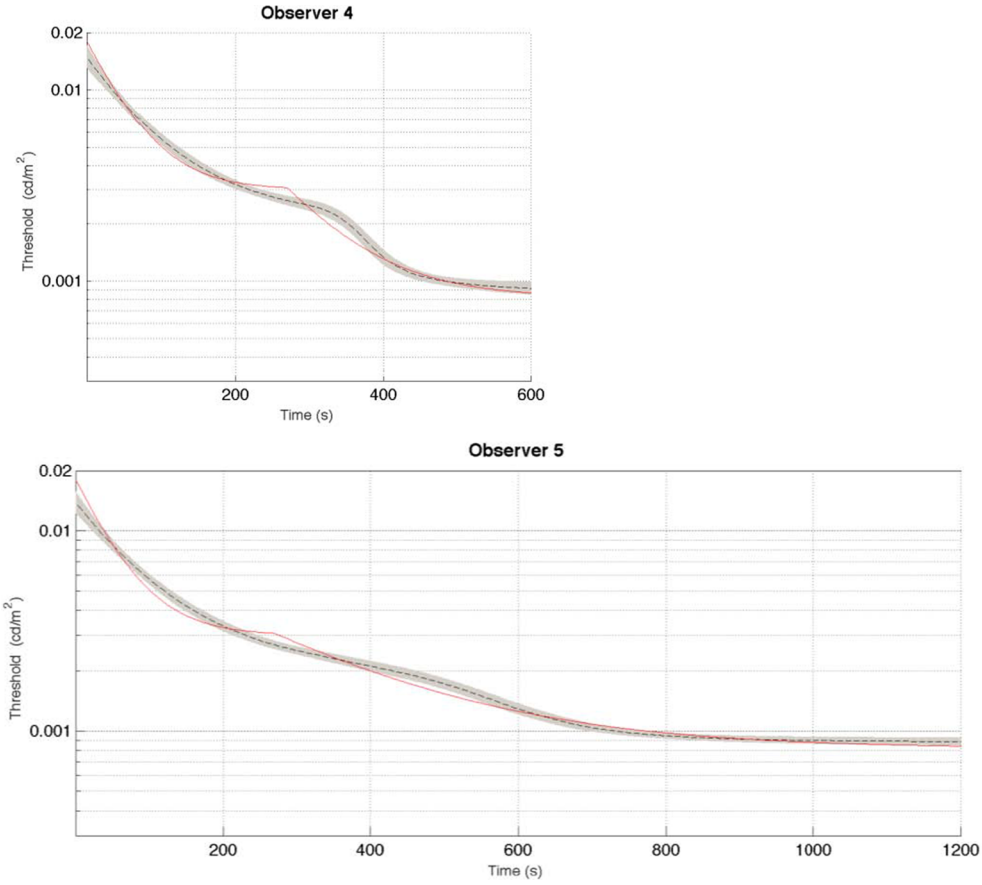


Fig. D2. Post-hoc segment-by-segment analysis of qCD estimates of the dark adaptation curves of Observers 4 and 5. The solid red lines represent the true dark adaptation curves of the simulated observers, the black dashed lines represent the estimated dark adaptation curves and the grey shaded areas represent the average 68.2% HWCI.

by-trial thresholds was 0.054, 0.047, and 0.046 $\log_{10}$ units, respectively. For Observer 5, the RMSE, SD, and average 68.2% HWCI of the estimated trial-by-trial thresholds was 0.042, 0.034, and 0.032 $\log_{10}$ units, respectively.

We then used the post-hoc segment-by-segment analysis to further improve the accuracy and precision of the estimates from the qCD method for simulated Observers 4 and 5. Fig. D2. shows the segmented threshold estimates from the qCD method for Observers 4 and 5. t_c equals to 270 s for Observers 4 and 5. For Observer 4, the number of runs with one, two, and three segments were 133, 864, and 3, respectively. For the 864 two-segment simulation runs, the average t_c is 337 $\pm$ 51.6 s. The RMSE, SD, and average 68.2% HWCI of the segmented threshold estimations was 0.034, 0.040, and 0.030 $\log_{10}$ units, respectively. For Observer 5, the number of runs with one, two, three and four segments were 9, 838, 151, and 2, respectively. For the 838 two-segment simulation runs, the average estimated t_c is 479.9 $\pm$ 79.6 s. For the 151 three-segment simulation runs, the average estimated t_c is 352.6 $\pm$ 64.1 s, with the second $t_c$at 607.4 $\pm$ 64.7 s. The RMSE, SD, and average 68.2% HWCI of the segmented threshold estimations was 0.029, 0.028, and 0.023 $\log_{10}$ units, respectively.

References

Archibald, N. K., Clarke, M. P., Mosimann, U. P., & Burn, D. J. (2009). The retina in Parkinsons disease. *Brain, 132*, 1128–1145. https://doi.org/10.1093/brain/awp068.

Aubert, H. (1865). *Physiologie der Netzhaut* (pp. 394). Breslau: E. Morgenstern.

Baek, J., Lesmes, L. A., & Lu, Z.-L. (2016). qPR: An adaptive partial-report procedure based on Bayesian inference. *Journal of Vision, 16*(10), 25. https://doi.org/10.1167/16.10.25.

Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain, 128*, 2430–2441. https://doi.org/10.1093/brain/awh561.

Braddick, O., & Atkinson, J. (2011). Development of human visual function. *Vision Research, 51*(13), 1588–1609. https://doi.org/10.1016/j.visres.2011.02.018.

Brown, B., Adams, A. J., Coletta, N. J., & Haegerstrom-Portnoy, G. (1986). Dark adaptation in age-related maculopathy. *Ophthalmic and Physiological Optics, 6*(1), 81–84. https://doi.org/10.1111/j.1475-1313.1986.tb00704.x.

Butler, P. D., Martinez, A., Foxe, J. J., Kim, D., Zemon, V., Silipo, G., ... Javitt, D. C. (2007). Subcortical visual dysfunction in schizophrenia drives secondary cortical impairments. *Brain, 130*, 417–430. https://doi.org/10.1093/brain/awl233.

Campbell, F. W. (1983). Why do we measure contrast sensitivity? *Behavioural Brain Research, 10*(1), 87–97. https://doi.org/10.1016/0166-4328(83)90154-7.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology, 197*(3), 551–566.

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precuing of forced-choice discrimination. *The Quarterly Journal of Experimental Psychology Section A, 43*(4), 859–880. https://doi.org/10.1080/14640749108400960.

Cornsweet, T. N. (1962). The staircase-method in psychophysics. *The American Journal of Psychology, 75*(3), 485–491. https://doi.org/10.2307/1419876.

Cornsweet, T. N., & Teller, D. Y. (1965). Relation of increment thresholds to brightness and luminance. *Journal of the Optical Society of America, 55*, 1303–1308.

D'Souza, D. C., Abi-Saab, W. M., Madonick, S., Forselius-Bielen, K., Doersch, A., Braley, G., ... Krystal, J. H. (2005). Delta-9-tetrahydrocannabinol effects in schizophrenia: Implications for cognition, psychosis, and addiction. *Biological Psychiatry, 57*(6), 594–608. https://doi.org/10.1016/j.biopsych.2004.12.006.

Demb, J. B., Boynton, G. M., & Heeger, D. J. (1998). Functional magnetic resonance imaging of early visual pathways in dyslexia. *Journal of Neuroscience, 18*(17), 6939–6951.

Derman, C. (1957). Non-parametric up-and-down experimentation. *The Annals of Mathematical Statistics, 28*(3), 795–798. https://doi.org/10.1214/aoms/1177706895.

Dimitrov, P. N., Guymer, R. H., Zele, A. J., Anderson, A. J., & Vingrys, A. J. (2008). Measuring rod and cone dynamics in age-related maculopathy. *Investigative Ophthalmology & Visual Science, 49*(1), 55–65. https://doi.org/10.1167/iovs.06-1048.

Dimitrov, P. N., Robman, L. D., Varsamidis, M., Aung, K. Z., Makeyeva, G. A., Guymer, R. H., & Vingrys, A. J. (2011). Visual function tests as potential biomarkers in age-related macular degeneration. *Investigative Ophthalmology & Visual Science, 52*(13), 9457–9469. https://doi.org/10.1167/iovs.10-7043.

Dorr, M., Wille, M., Viulet, T., Sanchez, E., Bex, P. J., Lu, Z.-L., & Lesmes, L. (2015). Next-generation vision testing: The quick CSF. *Current Directions in Biomedical Engineering, 1*(1), 131–134. https://doi.org/10.1515/cdbme-2015-0034.

Dosher, B. A., & Lu, Z.-L. (2007). The functional form of performance improvements in perceptual learning learning rates and transfer. *Psychological Science, 18*(6), 531–539. https://doi.org/10.1111/j.1467-9280.2007.01934.x.

Dosher, B., & Lu, Z.-L. (2017). Visual perceptual learning and models. *Annual Review of Vision Science, 3*(1), 343–363. https://doi.org/10.1146/annurev-vision-102016-061249.

Enroth-cugell, & Shapley, R. (1973). Adaptation and dynamics of cat retinal ganglion-cells. *Journal of Physiology-London, 233*(2), 271–309.

Gaffney, A. J., Binns, A. M., & Margrain, T. H. (2011). Topography of cone dark adaptation deficits in age-related maculopathy. *Optometry and Vision Science: Official Publication of the American Academy of Optometry, 88*(9), 1080–1087. https://doi.org/10.1097/OPX.0b013e3182223697.

Gaffney, A. J., Binns, A. M., & Margrain, T. H. (2013). The effect of pre-adapting light intensity on dark adaptation in early age-related macular degeneration. *Documenta Ophthalmologica, 127*(3), 191–199. https://doi.org/10.1007/s10633-013-9400-3.

Gaffney, A. J., Binns, A. M., & Margrain, T. H. (2014). Measurement of cone dark adaptation: A comparison of four psychophysical methods. *Documenta Ophthalmologica, 128*(1), 33–41. https://doi.org/10.1007/s10633-013-9418-6.

Gu, H., Kim, W., Hou, F., Lesmes, L. A., Pitt, M. A., Lu, Z.-L., & Myung, J. I. (2016). A hierarchical Bayesian approach to adaptive vision testing: A case study with the contrast sensitivity function. *Journal of Vision, 16*(6), 15. https://doi.org/10.1167/16.6.15.

Hecht, S., Haig, C., & Chase, A. M. (1937). The influence of light adaptation on subsequent dark adaptation of the eye. *The Journal of General Physiology, 20*(6), 831–850. https://doi.org/10.1085/jgp.20.6.831.

Hecht, S., Haig, C., & Wald, G. (1935). The dark adaptation of retinal fields of different size and location. *The Journal of General Physiology, 19*(2), 321–337. https://doi.org/10.1085/jgp.19.2.321.

Hollins, M., & Alpern, M. (1973). Dark adaptation and visual pigment regeneration in human cones. *The Journal of General Physiology, 62*(4), 430–447. https://doi.org/10.1085/jgp.62.4.430.

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*(10), 5365–5377. https://doi.org/10.1167/iovs.10-5468.

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. https://doi.org/10.1167/15.9.2.

Hou, F., Lesmes, L. A., Kim, W., Gu, H., Pitt, M. A., Myung, J. I., & Lu, Z.-L. (2016). Evaluating the performance of the quick CSF method in detecting contrast sensitivity function changes. *Journal of Vision, 16*(6), 18. https://doi.org/10.1167/16.6.18.

Hou, F., Lesmes, L. A., Kim, W., Gu, H., Pitt, M., Myung, J., & Lu, Z.-L. (2017). Predicting the contrast sensitivity function in different luminance conditions. *Investigative Ophthalmology & Visual Science, 58*(8), 4221.

Jackson, G. R., & Edwards, J. G. (2008). A short-duration dark adaptation protocol for assessment of age-related maculopathy. *Journal of Ocular Biology, Diseases, and Informatics, 1*(1), 7–11. https://doi.org/10.1007/s12177-008-9002-6.

Jackson, G. R., Felix, T., & Owsley, C. (2006). The scotopic sensitivity tester-1 and the detection of early age-related macular degeneration1. *Ophthalmic and Physiological Optics, 26*(4), 431–437. https://doi.org/10.1111/j.1475-1313.2006.00390.x.

Jackson, G. R., Owsley, C., & McGwin, G. (1999). Aging and dark adaptation. *Vision Research, 39*(23), 3975–3982. https://doi.org/10.1016/S0042-6989(99)00092-9.

Jakiela, H., Enroth-cugell, C., & Shapley, R. (1976). Adaptation and dynamics in X-cells and Y-cells of cat retina. *Experimental Brain Research, 24*(4), 335–342.

Jandhyala, V., Fotopoulos, S., MacNeill, I., & Liu, P. (2013). Inference for single and multiple change-points in time series. *Journal of Time Series Analysis, 34*(4), 423–446. https://doi.org/10.1111/jtsa.12035.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics, 49*(3), 227–229.

Kesten, H. (1958). Accelerated stochastic approximation. *The Annals of Mathematical Statistics, 29*(1), 41–59. https://doi.org/10.1214/aoms/1177706705.

Kim, W., Pitt, M. A., Lu, Z.-L., & Myung, J. I. (2017). Planning beyond the next trial in adaptive experiments: A dynamic programming approach. *Cognitive Science, 41*(8), 2234–2252. https://doi.org/10.1111/cogs.12467.

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*(11), 2465–2492. https://doi.org/10.1162/NECO_a_00654.

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.

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*(4), 369–389. https://doi.org/10.1016/j.jmp.2005.12.005.

Kumar, T., & Glaser, D. A. (1993). Initial performance, learning and observer variability for hyperacuity tasks. *Vision Research, 33*(16), 2287–2300. https://doi.org/10.1016/0042-6989(93)90106-7.

Lamb, T. D. (1981). The involvement of rod photoreceptors in dark adaptation. *Vision Research, 21*(12), 1773–1782. https://doi.org/10.1016/0042-6989(81)90211-X.

Lamb, T. D., & Pugh, E. N., Jr. (2004). Dark adaptation and the retinoid cycle of vision. *Progress in Retinal and Eye Research, 23*(3), 307–380. https://doi.org/10.1016/j.preteyeres.2004.03.001.

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics, 63*(8), 1279–1292. https://doi.org/10.3758/BF03194543.

Lesmes, L. A., Bex, P., Lu, Z.-L., Bittner, A. K., Ramulu, P. Y., Stellmann, J.-P., & Dorr, M. (2016). A survey of contrast sensitivity in visual neuropathology. *Investigative Ophthalmology & Visual Science, 57*(12), 5161.

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*(19), 3160–3176. https://doi.org/10.1016/j.visres.2006.04.022.

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), 17.1–17.21. https://doi.org/10.1167/10.3.17.

Lesmes, L. A., Lu, Z.-L., Baek, J., Tran, N., Dosher, B. A., & Albright, T. D. (2015). Developing Bayesian adaptive methods for estimating sensitivity thresholds (d′) in Yes-No and forced-choice tasks. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01070.

Levi, D. M., & Li, R. W. (2009). Perceptual learning as a potential treatment for amblyopia: A mini-review. *Vision Research, 49*(21), 2535–2549. https://doi.org/10.1016/j.visres.2009.02.010.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49*(2), 467–477.

Lu, Z. L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*(9), 1183–1198. https://doi.org/10.1016/S0042-6989(97)00273-3.

Lu, Z.-L., & Dosher, B. (2013). *Visual psychophysics: From laboratory to theory*. Cambridge, MA: The MIT Press.

Maniglia, M., Cottereau, B. R., Soler, V., & Trotter, Y. (2016). Rehabilitation approaches in macular degeneration patients. *Frontiers in Systems Neuroscience, 10*, 107. https://doi.org/10.3389/fnsys.2016.00107.

Mollon, J. D., & Polden, P. G. (1980). A curiosity of light adaptation. *Nature, 286*, 59–62.

Mollon, J. D., Stockman, A., & Polden, P. G. (1987). Transient tritanopia of a second kind. *Vision Research, 27*, 637–650.

Monge, Z. A., & Madden, D. J. (2016). Linking cognitive and visual perceptual decline in healthy aging: The information degradation hypothesis. *Neuroscience and Biobehavioral Reviews, 69*, 166–173. https://doi.org/10.1016/j.neubiorev.2016.07.031.

Moore, B. C. J. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing. *Ear and Hearing, 17*(2), 133–161. https://doi.org/10.1097/00003446-199604000-00007.

Mote, F. A., & Riopelle, A. J. (1951). The effect of varying the intensity and the duration of preexposure upon foveal dark adaptation in the human eye. *The Journal of General Physiology, 34*(5), 657–674. https://doi.org/10.1085/jgp.34.5.657.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual-attention. *Vision Research, 29*(11), 1631–1647. https://doi.org/10.1016/0042-6989(89)90144-2.

Owsley, C., Jackson, G. R., White, M., Feist, R., & Edwards, D. (2001). Delays in rod-mediated dark adaptation in early age-related maculopathy. *Ophthalmology, 108*(7), 1196–1202.

Owsley, C., McGwin, G., Jr., Clark, M. E., Jackson, G. R., Callahan, M. A., Kline, L. B., ... Curcio, C. A. (2016). Delayed rod-mediated dark adaptation is a functional biomarker for incident early age-related macular degeneration. *Ophthalmology, 123*(2), 344–351. https://doi.org/10.1016/j.ophtha.2015.09.041.

Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research, 23*(7), 689–699. https://doi.org/10.1016/0042-6989(83)90210-9.

Pianta, M. J., & Kalloniatis, M. (2000). Characterisation of dark adaptation in human cone pathways: An application of the equivalent background hypothesis. *The Journal of Physiology, 528*(3), 591–608. https://doi.org/10.1111/j.1469-7793.2000.00591.x.

Prins, N. (2013). The psi-marginal adaptive method: How to give nuisance parameters the attention they deserve (no more, no less). *Journal of Vision, 13*(7), https://doi.org/10.1167/13.7.3.

Pugh, E. N. (1975). Rushton's paradox: Rod dark adaptation after flash photolysis. *The Journal of Physiology, 248*(2), 413–431.

Purpura, K., Tranchina, D., Kaplan, E., & Shapley, R. (1990). Light adaptation in the primate retina – analysis of changes in gain and dynamics of monkey retinal ganglion-cells. *Visual Neuroscience, 4*(1), 75–93.

Reuter, T. (2011). Fifty years of dark adaptation 1961–2011. *Vision Research, 51*(21–22), 2243–2262. https://doi.org/10.1016/j.visres.2011.08.021.

Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The Annals of Mathematical Statistics, 22*(3), 400–407. https://doi.org/10.1214/aoms/1177729586.

Robson, J. (1966). Spatial and temporal contrast-sensitivity functions of visual system. *Journal of the Optical Society of America, 56*(8), 1141–2000. https://doi.org/10.1364/JOSA.56.001141.

Rovamo, J., Franssila, R., & Nasanen, R. (1992). Contrast sensitivity as a function of spatial-frequency, viewing distance and eccentricity with and without spatial noise. *Vision Research, 32*(4), 631–637. https://doi.org/10.1016/0042-6989(92)90179-M.

Rushton, W. A. H., & Powell, D. S. (1972). The rhodopsin content and the visual threshold of human rods. *Vision Research, 12*(6), 1073–1081. https://doi.org/10.1016/0042-6989(72)90098-3.

Sagi, D. (2011). Perceptual learning in vision research. *Vision Research, 51*(13), 1552–1566. https://doi.org/10.1016/j.visres.2010.10.019.

Shen, Y., & Richards, V. M. (2013). Bayesian adaptive estimation of the auditory filter. *The Journal of the Acoustical Society of America, 134*(2), 1134–1145. https://doi.org/10.1121/1.4812856.

Shen, Y., Sivakumar, R., & Richards, V. M. (2014). Rapid estimation of high-parameter auditory-filter shapes. *The Journal of the Acoustical Society of America, 136*(4), 1857–1868. https://doi.org/10.1121/1.4894785.

Sloane, M., Owsley, C., & Jackson, C. (1988). Aging and luminance-adaptation effects on

spatial contrast sensitivity. *Journal of the Optical Society of America a-Optics Image Science and Vision, 5*(12), 2181–2190. https://doi.org/10.1364/JOSAA.5.002181.
Song, X. D., Garnett, R., & Barbour, D. L. (2017). Psychometric function estimation by probabilistic classification. *The Journal of the Acoustical Society of America, 141*(4), 2513. https://doi.org/10.1121/1.4979594.
Song, X. D., Wallace, B. M., Gardner, J. R., Ledbetter, N. M., Weinberger, K. Q., & Barbour, D. L. (2015). Fast, continuous audiogram estimation using machine learning. *Ear and Hearing, 36*(6), e326–e335. https://doi.org/10.1097/AUD.0000000000000186.
Steinmetz, R. L., Haimovici, R., Jubb, C., Fitzke, F. W., & Bird, A. C. (1993). Symptomatic abnormalities of dark adaptation in patients with age-related Bruch's membrane change. *The British Journal of Ophthalmology, 77*(9), 549–554.
Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*(5), 13. https://doi.org/10.1167/11.5.13.
Stuart, J. A., & Birge, R. R. (1996). Characterization of the primary photochemical events in bacteriorhodopsin and rhodopsin. *Biomembranes: A Multi-Volume Treatise, 2*, 33–139. https://doi.org/10.1016/S1874-5342(07)80005-5.
Swanson, W., & Birch, E. (1992). Extracting thresholds from noisy psychophysical data. *Perception & Psychophysics, 51*(5), 409–422. https://doi.org/10.3758/BF03211637.
Tahir, H. J., Rodrigo-Diaz, E., Parry, N. R. A., Kelly, J. M. F., Carden, D., & Murray, I. J. (2017). Slowed dark adaptation in older eyes; effect of location. *Experimental Eye Research, 155*, 47–53. https://doi.org/10.1016/j.exer.2016.11.016.
Thomas, M. M., & Lamb, T. D. (1999). Light adaptation and dark adaptation of human rod photoreceptors measured from the a-wave of the electroretinogram. *The Journal of Physiology, 518*(Pt 2), 479–496. https://doi.org/10.1111/j.1469-7793.1999.0479p.x.
Tranchina, D., Gordon, J., & Shapley, R. (1984). Retinal light adaptation – Evidence for a feedback mechanism. *Nature, 310*(5975), 314–316. https://doi.org/10.1038/310314a0.
Treutwein, B., & Strasburger, H. (1999). Fitting the psychometric function. *Perception & Psychophysics, 61*(1), 87–106. https://doi.org/10.3758/BF03211951.
Wald, G., Durell, J., & Stgeorge, R. (1950). The light reaction in the bleaching of Rhodopsin. *Science, 111*(2877), 179–181. https://doi.org/10.1126/science.111.2877.179.
Watson, A. B. (2017). QUEST+: A general multidimensional Bayesian adaptive psychometric method. *Journal of Vision, 17*(3), 10. https://doi.org/10.1167/17.3.10.
Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*, 113–120.
Weil, R. S., Schrag, A. E., Warren, J. D., Crutch, S. J., Lees, A. J., & Morris, H. R. (2016). Visual dysfunction in Parkinson's disease. *Brain, 139*, 2827–2843. https://doi.org/10.1093/brain/aww175.
Winsor, C. P., & Clark, A.-B. (1936). Dark adaptation after varying degrees of light adaptation. *Proceedings of the National Academy of Sciences of the United States of America, 22*(6), 400–404.
Wright, B. A., Bowen, R. W., & Zecker, S. G. (2000). Nonlinguistic perceptual deficits associated with reading and language disorders. *Current Opinion in Neurobiology, 10*(4), 482–486. https://doi.org/10.1016/S0959-4388(00)00119-7.
Zhao, Y., Lesmes, L. A., & Lu, Z.-L. (2017). The quick Change Detection method: Bayesian adaptive assessment of the time course of perceptual sensitivity change. *Investigative Ophthalmology & Visual Science, 58*(8), 5633.