THE CONTRAST SENSITIVITY FUNCTION: FROM LABORATORY TO CLINIC

EDITED BY: Fang Hou, Zhong-Lin Lu, Peter J. Bex and Alexandre Reynaud

PUBLISHED IN: Frontiers in Neuroscience and Frontiers in Medicine

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88971-883-2
DOI 10.3389/978-2-88971-883-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THE CONTRAST SENSITIVITY FUNCTION: FROM LABORATORY TO CLINIC

Topic Editors:
Fang Hou, Wenzhou Medical University, China
Zhong-Lin Lu, New York University, United States
Peter J. Bex, Northeastern University, United States
Alexandre Reynaud, McGill University, Canada

Citation: Hou, F., Lu, Z.-L., Bex, P. J., Reynaud, A., eds. (2021). The Contrast Sensitivity Function: From Laboratory to Clinic. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-883-2

Table of Contents

frontiers in Neuroscience

EDITORIAL
published: 25 October 2021
doi: 10.3389/fnins.2021.783674

Editorial: The Contrast Sensitivity Function: From Laboratory to Clinic

Fang Hou[1]*, *Zhong-Lin Lu*[2,3,4], *Peter Bex*[5] and *Alexandre Reynaud*[6]

[1] School of Ophthalmology & Optometry and Eye Hospital, Wenzhou Medical University, Wenzhou, China, [2] Division of Arts and Sciences, NYU Shanghai, Shanghai, China, [3] Department of Psychology, Center for Neural Science, New York University, New York, NY, United States, [4] NYU-ECNU Institute of Brain and Cognitive Sciences at NYU Shanghai, Shanghai, China, [5] Department of Psychology, Northeastern University, Boston, MA, United States, [6] McGill Vision Research, McGill University, Montréal, QC, Canada

Keywords: contrast sensitivity function, functional vision, spatial frequency, temporal frequency, sensitivity

Editorial on the Research Topic

The Contrast Sensitivity Function: From Laboratory to Clinic

The contrast sensitivity function (CSF) describes how sensitivity (1/contrast threshold) to narrow-band stimuli varies with spatial and/or temporal frequency. It provides a fundamental characterization of spatial or temporal vision, and reflects the combined effects of limiting factors from the optics of the eye all the way to the response properties of cortical neurons. As the boundary between the visible and invisible narrow-band visual stimuli, the CSF correlates better with real-world daily visual functions than visual acuity (VA). This Research Topic, consisting of a collection of 13 fundamental and clinical oriented research articles, extends our knowledge related to the CSF.

How stimulus contrast at different spatial or temporal frequencies affects the response properties of the visual system is the core of the CSF. Camillo et al. found that the contrast and temporal frequency tunings of mouse V1 neurons are inseparable. The divisive contrast-gain normalization model provided a good fit to their data at high temporal frequencies, but not at low temporal frequencies. The results suggest that different normalization mechanisms exist in the mouse visual cortex and may underlie the different relationships between temporal frequency and spatial contrast tunings at different temporal frequencies.

Using psychophysics and functional MRI, Negishi and Shinomori found that weak color (low saturation) had a stronger inhibitory effect on contrast sensitivity than strong color (high saturation), and the interaction between chromatic and achromatic signals happened in V1. Contrast processing can also be improved by perceptual learning. Xi et al. measured ERP contrast response functions before and after 10 days of perceptual learning. They found that training significantly improved VA and contrast sensitivity, and led to changes at different stages of visual processing.

The CSF is an important functional vision assessment in the clinic. A series of studies on various clinical populations further demonstrated its importance. Rosenkranz et al. validated the performance of qCSF, a Bayesian adaptive method for efficient CSF measurement (Lesmes et al., 2010), in patients with multiple sclerosis (MS). The authors suggest that the qCSF method can be used as a tool to evaluate contrast vision for MS patients. Hoffmann et al. measured visual functions and morphologic parameters of the retina in patients with age-related macular degeneration (AMD). They found that among all the measures of visual function, the CSF correlated best with anatomic features of the retina. Li et al. evaluated the effects of different myopia control lenses on the CSF and VA in a group of myopic children. They found that the lenses with concentrically arranged aspherical lenslets had the least influence on the CSF and VA. Min et al. evaluated effects of Eyetronix Flicker Glass (EFG) treatment on children with amblyopia on the CSF and VA, and

OPEN ACCESS

Edited and reviewed by:
Rufin VanRullen,
Centre National de la Recherche Scientifique (CNRS), France

***Correspondence:**
Fang Hou
houf@mail.eye.ac.cn

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *26 September 2021*
Accepted: *27 September 2021*
Published: *25 October 2021*

Citation:
Hou F, Lu Z-L, Bex P and Reynaud A (2021) Editorial: The Contrast Sensitivity Function: From Laboratory to Clinic. Front. Neurosci. 15:783674. doi: 10.3389/fnins.2021.783674

concluded that the EFG can be an additional choice for amblyopia therapy. Cheng et al. showed that 0.01% atropine had no significantly detrimental effect on the CSF in myopic adults, suggesting that 0.01% atropine might be safely used for myopia control. Finally, Stalin et al. assessed visual functions and skiing performance in elite skiers with visual impairments, and found that the CSF did not add any predictive value on skiing performance.

In clinical settings, it is important to measure the CSF quickly, accurately and reliably. Some of the submitted works explored new methods for efficient CSF measurement. The Campbell-Robson chart is widely used to illustrate the CSF as the boundary between the visible and invisible narrow band visual stimuli, and this border might be used to estimate the CSF. However, Tardif et al. examined this putative idea and concluded that the Campbell-Robson chart did not accurately quantify the CSF. Xiong et al. provided a method to simulate visual quality of patients with low vision based on clinical measures of VA and CSF. Their method successfully predicted reading performance and text visibility under a wide range of low vision conditions. Xu et al. developed a Bayesian adaptive method (qVFM) to estimate the contrast sensitivity map over the visual field. They demonstrated that the qVFM method could accurately and efficiently evaluate the contrast sensitivity visual field map in normal observers. Finally, because current CSF tests are too boring to sustain young children's interest, Elfadaly et al. developed a gamified, child-friendly CSF assessment and showed that it successfully assessed CSF in young children.

Altogether this collection of articles emphasizes the importance of measuring the CSF to assess visual function in both basic research and clinical settings. It presents some methods to perform and improve those measures, and considers their interpretation and implications.

AUTHOR CONTRIBUTIONS

All authors contributed to manuscript writing and revision, read, and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (NSFC81600764 to FH) and the Department of Human Resources and Social Security of Zhejiang Province (Qianjiang Talent Project, QJD1803028 to FH).

REFERENCES

Lesmes, L. A., Lu, Z. L., Baek, J., and Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J. Vis.* 10, 17.11–21. doi: 10.1167/10.3.17

Conflict of Interest: PB and Z-LL hold equity shares in Adaptive Sensory Technology, Inc. and have patents related to the qCSF technology.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 07 July 2020
doi: 10.3389/fnins.2020.00665

Mapping the Contrast Sensitivity of the Visual Field With Bayesian Adaptive qVFM

Pengjing Xu[1], *Luis A. Lesmes*[2], *Deyue Yu*[1] *and Zhong-Lin Lu*[3,4,5]*

[1] College of Optometry, The Ohio State University, Columbus, OH, United States, [2] Adaptive Sensory Technology, Inc., San Diego, CA, United States, [3] Division of Arts and Sciences, NYU Shanghai, Shanghai, China, [4] Center for Neural Science and Department of Psychology, New York University, New York, NY, United States, [5] NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, Shanghai, China

Current clinical evaluation, which focuses on central vision, could be improved through characterization of residual vision with peripheral testing of visual acuity, contrast sensitivity, color vision, crowding, and reading speed. Assessing visual functions in addition to light sensitivity, a comprehensive visual field map (VFM) would be valuable for detecting and managing eye diseases. In a previous study, we developed a Bayesian adaptive qVFM method that combines a global module for preliminary assessment of the VFM's shape and a local module for assessment at individual retinal locations. The method was validated in measuring the light sensitivity VFM. In this study, we extended the qVFM method to measure contrast sensitivity across the visual field. In both simulations and psychophysics, we sampled 64 visual field locations (48 x 48 deg) and compared the qVFM method with a procedure that tested each retinal location independently (qFC; Lesmes et al., 2015). In each trial, subjects were required to identify a single optotype (size: 2.5 x 2.5 deg), one of 10 filtered Sloan letters. To compare the accuracy and precision of the two methods, three simulated eyes were tested in 1,280 trials with each method. In addition, data were collected from 10 eyes (5 OS, 5 OD) of five normal observers. For simulations, the average RMSE of the estimated contrast sensitivity with the qVFM and qFC methods were 0.057 and 0.100 after 320 trials, and 0.037 and 0.041 after 1,280 trials [all in log10 units, represent as *log(sensitivity)*], respectively. The average SD of the qVFM and qFC estimates were 0.054 and 0.096 after 320 trials, and 0.032 and 0.041 after 1,280 trials, respectively. The within-run variability (68.2% HWCIs) were comparable to the cross-run variability (SD). In the psychophysics experiment, the average HWCI of the estimated contrast sensitivity from the qVFM and qFC methods across the visual field decreased from 0.33 on the first trial to 0.072 and 0.16 after 160, and to 0.060 and 0.10 after 320 trials. The RMSE between the qVFM and qFC estimates started at 0.26, decreased to 0.12 after 160 and to 0.11 after 320 qVFM trials. The qVFM provides an accurate, precise, and efficient mapping of contrast sensitivity across the entire visual field. The method might find potential clinical applications in monitoring vision loss, evaluating therapeutic interventions, and developing effective rehabilitation for visual diseases.

Keywords: Bayesian adaptive testing, automated perimetry, visual-filed map, peripheral vision, contrast sensitivity, active learning, Sloan letters

OPEN ACCESS

Edited by:
Christopher DiMattina, Florida Gulf Coast University, United States

Reviewed by:
Woojae Kim, Howard University, United States
Jenny C. A. Read, Newcastle University, United Kingdom

***Correspondence:**
Zhong-Lin Lu
zhonglin@nyu.edu

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *14 February 2020*
Accepted: *29 May 2020*
Published: *07 July 2020*

Citation:
Xu P, Lesmes LA, Yu D and Lu Z-L (2020) Mapping the Contrast Sensitivity of the Visual Field With Bayesian Adaptive qVFM. Front. Neurosci. 14:665. doi: 10.3389/fnins.2020.00665

INTRODUCTION

A comprehensive characterization of peripheral vision with assessment of the Visual Filed Map (VFM) is crucial for monitoring the status of vision loss, for developing and providing effective rehabilitation interventions (Sunness et al., 1995; Fletcher and Schuchard, 1997; Markowitz and Muller, 2004), and for obtaining projections of potential benefits from interventions (Massof and Rubin, 2001; Strasburger et al., 2011).

As a part of the clinical ophthalmic diagnostic procedure, the VFM of light sensitivity is assessed by the majority of eye care practitioners, mostly using the standard automated perimetry (SAP) (Dreyer, 1993; Johnson et al., 2011). Assessment of the VFM of many other visual functions, such as contrast sensitivity (Daitch and Green, 1969; Swanson et al., 2014), visual acuity (VA, 1965; Thompson et al., 1982), color vision (Carlow et al., 1976; Hart et al., 1984; Sample and Weinreb, 1990, 1992), reading speed (Ramulu et al., 2009; Yu et al., 2010), and crowding (Balas et al., 2009; Levi and Carney, 2009), is difficult and rarely used in the clinic. In fact, results from the standard automated perimetry (SAP) are noisy (Stewart and Hunt, 1993; Keltner et al., 2000). Precise and accurate VFM assessments of visual functions are time consuming with existing methods (Artes et al., 2002; Weinreb and Kaufman, 2009, 2011). A number of new perimetric methods have been developed and could potentially provide helpful clinical information, but have not sufficiently validated for routine clinical use (Johnson et al., 2011; Strasburger et al., 2011; Keltgen and Swanson, 2012; Swanson et al., 2014).

In a previous study, we developed a novel Bayesian adaptive testing method, the qVFM method, that combines a global module for preliminary assessment of the VFM's shape and a local module for assessing individual visual field locations to provide an efficient and precise assessment of the VFM (Xu et al., 2019). In its first implementation, we applied the qVFM method to assess the light sensitivity visual field map with a Yes/No paradigm. Computer simulations and a psychophysical validation study both showed that the qVFM method could provide an accurate, precise and efficient assessment of light sensitivity VFM.

In this study, we implemented the qVFM method in a 10-alternative forced-choice (10AFC) letter identification paradigm to measure contrast sensitivity (CS) across the visual field to provide an assessment of the contrast sensitivity visual field map.

As a clinical measure, contrast sensitivity predicts functional vision better than many other visual diagnostics (Comerford, 1983; Jindra and Zemon, 1989; Ginsburg, 2003; Faye, 2005). Deficits in contrast sensitivity accompany many visual diseases, including amblyopia (Hess and Howell, 1977; Bradley and Freeman, 1981; Kiorpes et al., 1999; Xu et al., 2006; Qiu et al., 2007), glaucoma (Ross et al., 1984; Stamper, 1984; Hot et al., 2008), optic neuritis (Zimmern et al., 1979; Trobe et al., 1996), diabetic retinopathy (Della Sala et al., 1985; Sokol et al., 1985), Parkinson's disease (Bulens et al., 1986; Bodis-Wollner et al., 1987; Mestre et al., 1990), and multiple sclerosis (Regan et al., 1981, 1982; Hess and Plant, 1985; Travis and Thompson, 1989; Regan and Hamstra, 1991). Such deficits are evident even when acuity and/or light sensitivity perimetry tests appear normal (Jindra and Zemon, 1989; Woods and Wood, 1995). Contrast sensitivity is also an important outcome measure of refractive and cataract surgery (Ginsburg, 1987, 2006; Applegate et al., 1998, 2000; McLeod, 2001; Bellucci et al., 2005), and potential rehabilitation programs for macular degeneration (Loshin and White, 1984), myopia (Tan and Fong, 2008), and amblyopia (Polat et al., 2004; Li et al., 2005, 2009; Zhou et al., 2006; Huang et al., 2008). On the other hand, although the literature has documented the importance of contrast sensitivity test, the current in-clinic contrast sensitivity exams mostly consist of contrast sensitivity measurements in fovea, e.g., the Pelli-Robson chart (Pelli and Robson, 1988), which can only provide a limited contrast sensitivity assessment of residual spatial vision for ophthalmic patients (Elliott and Whitaker, 1992).

Our new implementation of the qVFM method was based on the qFC procedure, originally developed to measure contrast sensitivity with forced-choice paradigms at a single visual location (Hou et al., 2015; Lesmes et al., 2015). Here, we integrated the qFC procedure with the qVFM method to assess contrast sensitivity across the visual field. In the rest of this paper, we first briefly describe the 10AFC implementation of the qVFM method, then computer simulations, and finally a psychophysical validation experiment.

qVFM WITH 10-AFC

The qVFM method consists of three major modules (Xu et al., 2019; see **Appendix C** for more details):

1) The global module, which assesses the shape of the VFM through a Bayesian adaptive procedure to estimate the posterior distributions of the parameters of a tilted elliptic paraboloid function (TEPF):

$$\tau\left(x,y\right)=EPZ-\left(\frac{x}{EPA}\right)^2-\left(\frac{y}{EPB}\right)^2+SLA*x+SLB*y \tag{1}$$

where *EPZ* is the contrast sensitivity at the fovea, *EPA* is the bandwidth (latus rectum) in the horizontal direction, *EPB* is the bandwidth in the vertical direction, *SLA* is the tilt level in the horizontal direction, and *SLB* is the tilt level in the vertical direction. The height of the TEPF, τ(x, y), defines the contrast sensitivity (1/contrast) at a fixed $d' = 1.5$ level at visual field location (x, y).

2) The switch module, which evaluates the rate of information gain in the global module and determines when to switch to the local module, and, at the point of the switch, generates the prior distribution of the visual function (e.g., light sensitivity, contrast sensitivity) at each visual field location based on the posterior from the global module.

3) The local module, which provides independent assessment of visual function at each visual field location using another Bayesian adaptive procedure that determines the location and stimulus parameters of test stimulus in each trial based on the relative information gain across locations.

In the global module, a probability density function, $p(\vec{\theta})$, where $\vec{\theta} = (EPZ, EPA, EPB, SLA, SLB)$, is defined over the parameter space of the TEPF. The initial prior distribution $p_{t=0}(\vec{\theta})$ represents foreknowledge of model parameters before any data collection (trial $t = 0$). A stimulus space, which includes all possible stimulus intensities and stimulus locations (x, y), is also defined in the qVFM procedure. The local module starts with a prior distribution in each retinal location. In both the global and local modules, a one-step-ahead search strategy is used to determine the optimal stimulus in the next trial that would lead to the maximum information gain (equivalent to the minimum expected entropy), and the selection of optimal stimulus location and intensity is always based on the total expected entropy across all the visual field locations. Using Bayesian update and optimal stimulus selection (Kontsevich and Tyler, 1999; Lesmes et al., 2006, 2010, 2015; Lu and Dosher, 2013), the qVFM updates the posterior distribution of the parameters based on subject's response in each trial to estimate the shape of the VFM in the global module or the individual parameters of each location in the local module.

In a previous paper (Xu et al., 2019), we implemented the qVFM method with a Yes/No task. In the new 10AFC implementation, we kept the general algorithm unchanged except the likelihood function, which was based on the d' psychometric function for Yes/No in the earlier implementation of the method. In a 10-AFC task, the d' psychometric function (i.e., perceptual sensitivity for a given stimulus contrast *s)* at each visual field location (x, y), can be modeled as (Foley and Legge, 1981; Legge et al., 1987; Hou et al., 2015):

$$d'(s,x,y) = 1.5\left(\frac{\tau(x,y)}{s}\right)^{\gamma} \tag{2}$$

where s is the reciprocal of the contrast of the stimulus (i.e., 1/contrast), $\tau(x,y)$ is the contrast sensitivity at location (x,y), γ is the steepness of the d' psychometric function. Plotted on log-log axes, this function is linear over the contrast of the stimulus. Following previous studies (Foley and Legge, 1981; Lu and Dosher, 1999; Hou et al., 2015; Lu et al., 2019), we set $\gamma = 2.35$ in the current implementation of the qVFM. Based on signal detection theory (Gu and Green, 1994; Klein, 2001), the probability of correctly identifying the target in an m-alternative forced choice (m-AFC) identification task is a function of the corresponding d' (Hacker and Ratcliff, 1979):

$$P(s,x,y) = \int_{-\infty}^{+\infty} \phi\left(t - d'(s,x,y)\right) \Phi^{m-1}(t)\, dt \tag{3}$$

where $\phi()$ is the probability density function of the standard normal distribution, $\Phi()$ is the cumulative probability density function of the standard normal distribution, m is the number of alternatives in the m-AFC task (which is 10 in this study), and $d'(s, x, y)$ is the d' value for a stimulus s at visual field location (x,y). In an m-AFC task, the observer compares the internal responses of the target with those of the m-1 non-target. The probability density of obtaining an internal response t from the target stimulus is $\phi\left(t - d'(s,x,y)\right)$; the probability density of obtaining an internal response t that is greater than all the m-1 non-target responses is $\Phi^{m-1}(t)$; and, according to the max decision rule, the probability of correctly identifying the target, $P(s,x,y)$, is the probability that all possible internal responses of the target are greater than those from the m-1 non-targets, which is the product of the two probability density functions integrated over all the possible values of t (Lu and Dosher, 2013).

In addition, we assume a fixed lapse rate ε for human observers (Klein, 2001; Wichmann and Hill, 2001; Lesmes et al., 2015):

$$P'(s,x,y) = \frac{1}{10}\varepsilon + (1-\varepsilon)\, P(s,x,y) \tag{4}$$

where $P(s,x,y)$ is the psychometric function without lapse (Equation 3). In the qVFM method, ε is set to 0.03 (Wichmann and Hill, 2001; Lesmes et al., 2010). Equation (4) defines the likelihood function that completely describes the probability of 10AFC target identification across all visual field locations and contrast levels in the qVFM method.

SIMULATIONS

Methods

To evaluate the performance of the qVFM procedure for observers with a range of performance, we simulated three observers asked to perform a 10AFC letter identification task in 64 retinal locations. The parameters of the three simulated observers were chosen to approximate those of the observers in our psychophysical validation study, shown in **Table 1**. The blind spot of all simulated observers was at (−15 degree, −3 degree).

In the qVFM method, the parameter space includes 30 linearly spaced EPA values [from 36.0 to 96.0 degree/$\sqrt{log(sensitivity)}$], 30 linearly spaced EPB values [from 36.0 to 96.0 degree/$\sqrt{log(sensitivity)}$], 50 linearly spaced EPZ values [from 0.25 to 1.4 $log(sensitivity)$], 20 linearly spaced SLA values [from −0.015 to 0.015 $log(sensitivity)$/degree] and 20 linearly spaced SLB values [from −0.016 to 0.016 $log(sensitivity)$/degree]. The broad parameter space ensures robust assessment of a wide range of patient populations and avoids effects of extreme values—the tendency to bias toward the center of the parameter space when the observer's true parameter values are close to the boundary of the space.

TABLE 1 | Parameters of the three simulated observers.

Simulation	EPA	EPB	EPZ	SLA	SLB
Observer 1	72	54	0.60	0.003	0.005
Observer 2	54	48	1.2	0.001	0.003
Observer 3	61	55	0.85	0.002	0.004

The unit of EPA and EPB is degree/$\sqrt{log(sensitivity)}$, unit of EPZ is log(sensitivity), and unit of SLA and SLB is log(sensitivity)/degree.

For each of the five qVFM parameters, the priors were defined by a hyperbolic secant (sech) function (King-Smith and Rose, 1997). For each qVFM parameter, θ_i, for i = 1, 2, 3, 4, 5, the mode of the marginal prior $p(\theta_i)$ was defined by the best guess for that parameter based on a pilot study, $\theta_{i,guess}$, and the width was defined by the confidence, $\theta_{i,confidence}$:

$$P(\theta_i) = sech(\theta_{i,confidence} \times (\theta_i - \theta_{i,\,guess})) \quad (5)$$

The priors were log-symmetric around $\theta_{i,guess}$, whose values for the respective parameters were: EPA = 60 (degree/$\sqrt{log(sensitivity)}$), EPB = 54 (degree/$\sqrt{log(sensitivity)}$), EPZ = 0.90 ($log(sensitivity)$), SLA = 0.002 ($log(sensitivity)$/degree), and SLB = 0.003 ($log(sensitivity)$/degree). For $\theta_{confidence}$ of each parameter, the value was set to 3.1 for EPA, 2.6 for EPB, 3.4 for EPZ, 5.2 for SLA, 4.5 for SLB. The joint prior was defined as the normalized product of the marginal priors, which resulted in a relatively moderate informative prior for the three simulated observers in our study.

The stimulus space includes an 8 x 8 grid of retina locations (48 x 48 degree) and log-linearly spaced contrast values [between 0.05 to 1.0, corresponding to 0 to 1.3 $log(sensitivity)$]: with 60 values in the global module and 120 contrast values in the local module.

We compared the performance of the full qVFM procedure that has all three modules with a reduced qVFM procedure that has only the local module in 1,000 repeated simulations of 1,280 trials each. The priors in the reduced qVFM was generated from the prior of the global module of the full qVFM. In other words, the two methods are equated before the first trial.

Metrics of Evaluation

Accuracy is a measure of how much the estimate deviate from the truth on average, and precision is a measure of the variability of repeated estimates. We quantify accuracy using the root mean squared error (RMSE) of the estimated contrast sensitivities across all 64 visual field locations. The RMSE after the i-th trial can be calculated as:

$$RMSE_i^{simulation} = \sqrt{\frac{\sum_k \sum_j \left(\tau_{ijk} - \tau_k^{true}\right)^2}{J \times K}} \quad (6)$$

where τ_{ijk} is the estimated sensitivity at the k-th VF location after i trials in the j-th run, and τ_k^{true} is the true sensitivity at that location.

Precision is defined as the inverse of the variability of the estimates. Two methods were used to assess the precision of the qVFM method. The first is based on the standard deviation of repeated measures:

$$SD_i = \sqrt{\frac{\sum_k \sum_j \left(\tau_{ijk} - mean(\tau_{ijk})\right)^2}{J \times K}} \quad (7)$$

Another measure of precision is the average half width of the credible interval (HWCI) of the posterior distribution of the estimated sensitivities across retina locations. The 68.2% credible interval represents the range within which the actual value lies with 68.2% probability. Since researchers typically do not repeat an experiment many times for the same observer, the HWCI of the posterior distribution is a very important index of precision that can be obtained with a single run of the qVFM procedure (Hou et al., 2015).

Results

A simulation of the qVFM and qFC methods based on the parameters of the simulated observer 1 is shown in **Supplementary Movie 1**. The simulation program can be downloaded from GitHub (https://github.com/hvxpj/qVFM_Demo/issues/1#issue-604728692). The GUI allows users to adjust the parameters of the simulated observers and prior used in the qVFM method.

The estimated VFMs of the three simulated observers, obtained with the qVFM and qFC methods, are shown in **Figure 1** (simulated observer 1) and **Figures A1, A2** (simulated observers 2 and 3) in **Appendix A**.

In characterizing spatial vision, the area under the log contrast sensitivity function is often used as a summary metric (Applegate et al., 1998, 2000; Oshika et al., 1999, 2006; van Gaalen et al., 2009; Hou et al., 2010; Lesmes et al., 2010; Jia et al., 2014; Dorr et al., 2015; Zheng et al., 2018). Here, we used the volume under the surface of the VFM (VUSVFM) to provide a summary metric of the entire visual field.

Figure 2 shows the $RMSE^{simulation}$, standard deviation, average 68.2% HWCI and VUSVFM of the estimated contrast sensitivities obtained from the qVFM and qFC methods for the three simulated observers over 1,280 trials. In log10 units [represent as $log(sensitivity)$], the average $RMSE^{simulation}$ of the three simulated observers started at 0.24 for both the qVFM and qFC methods. It decreased to 0.057 and 0.10 in the qVFM and qFC methods after the first 320 trials, and to 0.037 and 0.041 in the two methods after 1,280 trials, respectively. The SD of the estimated sensitivities was 0.054 in the qVFM method and 0.096 in the qFC method after 320 trials, which decreased to 0.032 in the qVFM method and 0.041 in the qFC method after 1,280 trials. The average 68.2% HWCI of the estimated sensitivities also decreased with trial number. It started at 0.32 in both the qVFM and qFC methods, decreased to 0.055 in the qVFM method and 0.094 in the qFC method after the first 320 trials, and to 0.033 in the qVFM method and 0.039 in the qFC method after 1,280 trials.

For the qVFM method, the switch from the global module to the local module occurred between 31 and 69 trials, with the mean around 41 trials and standard deviation of 9.3 trials. From **Figure 2**, we can see that the global module acted very efficiently in reducing random errors and uncertainties in the beginning of the measurement.

The simulations showed that the estimated VFM from both the qVFM and qFC method could reach high accuracy and precision in 1,280 trials. The qVFM method could

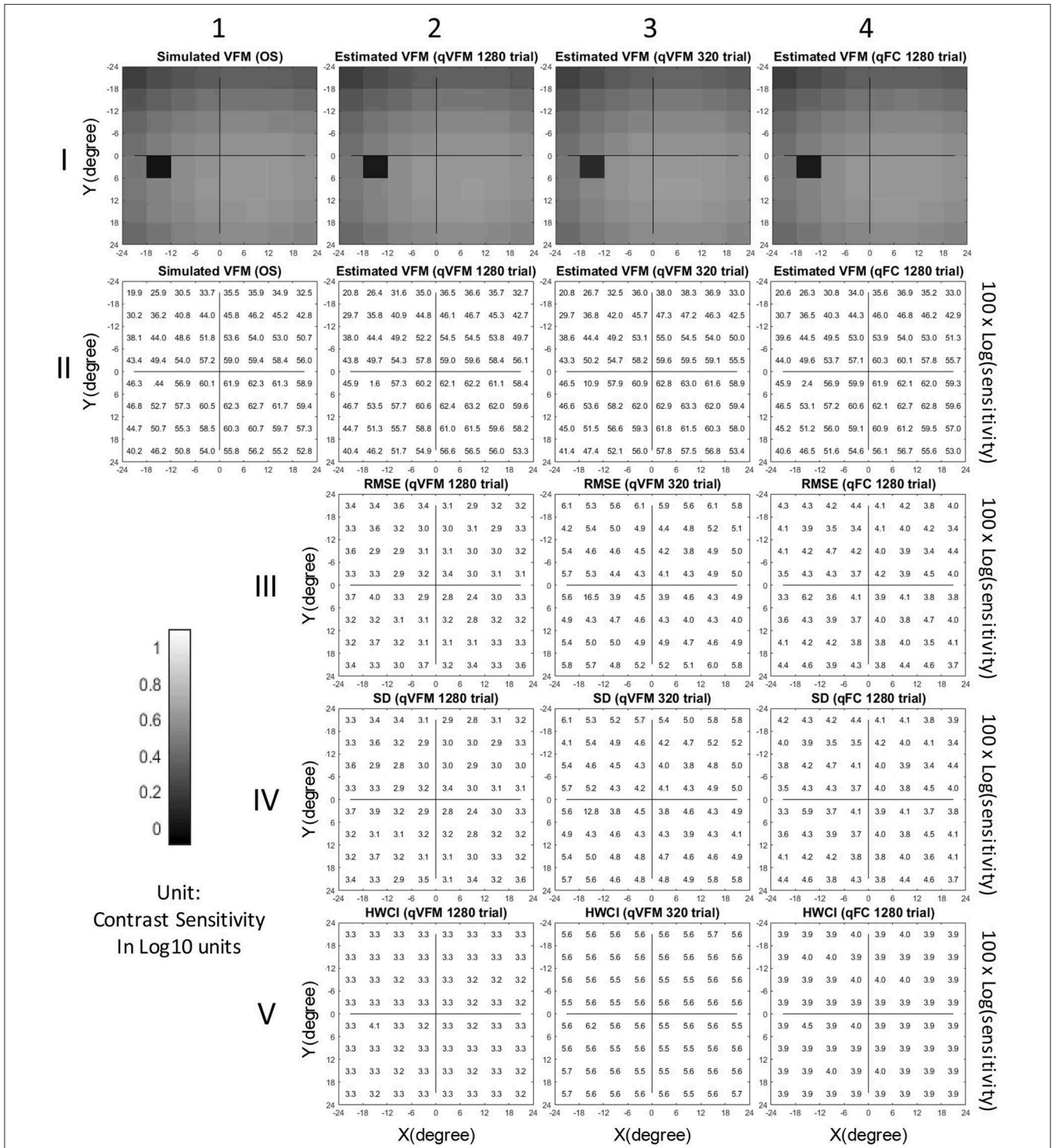


FIGURE 1 | Simulation results I (Observer 1): The contrast sensitivity VFM of the simulated observer (I,1; II,1); The estimated VFM obtained with the qVFM method after 1,280 trials (I,2; II,2) and 320 trials (I,3; II,3); The estimated VFM obtained with the qFC method (I,4; II,4); And the corresponding $RMSE^{simulation}$ (III), standard deviation (IV), and HWCI (V) of the estimates.

however converge much quicker and achieve good accuracy and precision in a much shorter period of time comparing to the qFC method. To achieve 0.1 $\log(sensitivity)$ accuracy and 0.1 $log(sensitivity)$ precision, on average, the qVFM method only took 106 trials, whereas the qFC method needed 334 trials.

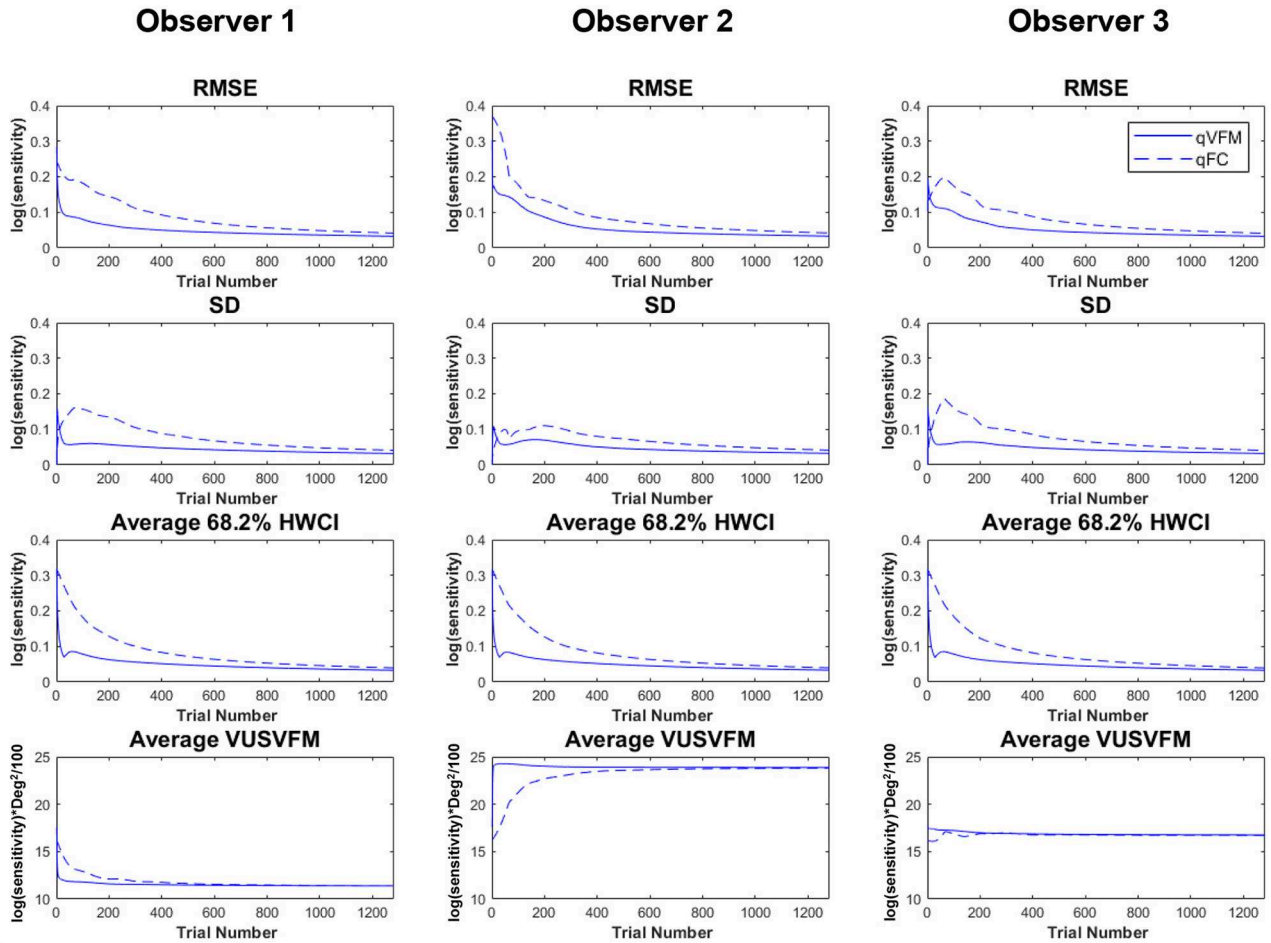


FIGURE 2 | Simulation results II. The first three rows show the $RMSE^{simulation}$, standard deviation, and average 68.2% HWCI of the estimated contrast sensitivities across 64 locations and 1,000 runs. The fourth row shows the average volume under the surface of the VFM (VUSVFM) across 1,000 runs. The true VUSVFM is 11.3 for observer 1, 23.8 for observer 2 and 16.7 for observer 3 (unit: *log(sensitivity)*) x degree^2/100.

PSYCHOPHYSICAL VALIDATION

Methods

Participants

We collected data from ten eyes (5 OS, 5 OD) of 5 normal (3 male and 2 female) subjects, including four naïve observers (Subject 2–Subject 5) and one of the authors (Subject 1). All subjects were between 32 and 39 years of age.

Apparatus

The psychophysical experiment was conducted on an IBM PC compatible computer, running Matlab programs with *PsychToolbox* extensions (Brainard and Vision, 1997; Pelli, 1997). Subjects viewed the stimuli monocularly with natural pupil at a viewing distance of 30 cm in a dimly lighted room. The stimuli were displayed on a Samsung 55-inch monitor [Model: UN55FH6030, Clear Motion Rate (CMR) of 240], with a screen size of 120.6 x 67.8 cm, corresponding to a field of view 127.0 x 97.0 degrees for the subjects, a screen resolution of 1920 x 1080 pixels, a refresh rate of 60 Hz, and a background luminance at 47 cd/m^2. A chin–forehead rest was used to minimize head movements during the experiment.

Stimuli

Ten Sloan letters, filtered with a raised cosine filter and octave bandwidth (central spatial frequency: 1.2 cycles per degree), served as stimuli (**Figure 3**). The contrast of the letters varied between 0.05 to 1, corresponding to 0 to 1.3 *log*(*sensitivity*).

Design and Procedure

In each trial, a single optotype (size: 2.5 x 2.5 degree) was presented for 200 ms in one of the 8 x 8 possible retina locations, evenly distributed in a 48 x 48-degree visual field (**Figure 4**). Subjects were asked to identify the letter. On each trial, the contrast and location of the stimulus was adaptively selected. The inter-trial interval was set to 1.2 s.

Each eye was tested in four sessions, each consisting of an independent 320-trial qVFM assessment and 320 qFC trials, with the two types of trials randomly mixed.

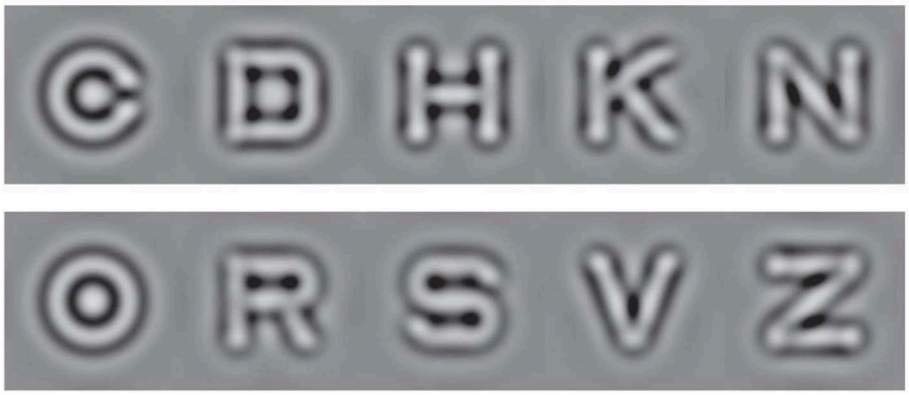


FIGURE 3 | The 10 filtered Sloan letters used in the study.

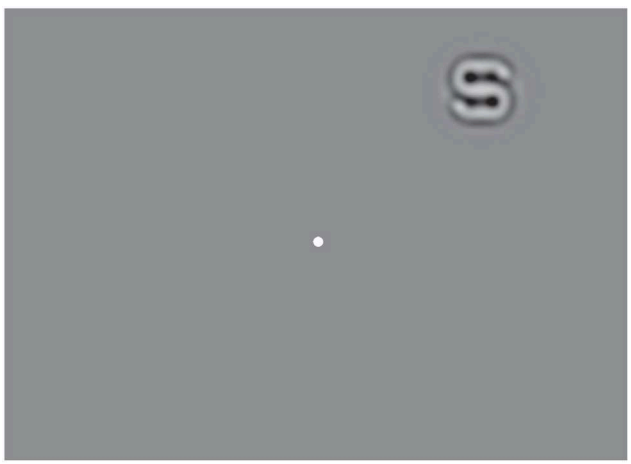


FIGURE 4 | Illustration of the stimulus layout in the psychophysical experiment. A filtered Sloan letter was displayed for 200 ms at one of the 8×8 possible retinal locations on the screen. Subjects were asked to fixate on the center dot and report which letter was present.

Results

The estimated VFMs of the 10 tested eyes from both the qVFM and qFC methods are shown in **Figure 5** (Subject 1) and **Figures B1–B4** (Subject 2-5, in **Appendix B**).

The agreement between the estimated VFMs from the qVFM and qFC was evaluated by the root mean squared error (RMSE) of the estimated contrast sensitivities across all 64 retina locations:

$$RMSE_i^{eyes} = \sqrt{\frac{\sum_l \sum_k \sum_j \left(\tau_{ijkl}^{qVFM} - \tau_{kl}^{qFC}\right)^2}{J \times K \times L}} \quad (8)$$

where τ_{ijkl}^{qVFM} is the estimated contrast sensitivity from the qVFM method in the k-th VF location of the l-th eye after i trials in the j-th session, and τ_{kl}^{qFC} is the estimated contrast sensitivity from the qFC method in the k-th VF location of the l-th eye after 1,280 trials. The average $\text{RMSE}^{\text{eyes}}$ (in $log(sensitivity)$ units) started at 0.26 on the first qVFM trial and decreased to 0.12 after 160 qVFM trials and to 0.11 after 320 qVFM trials across all test sessions and eyes (**Figure 6A**). That the decreasing $\text{RMSE}^{\text{eyes}}$ estimates is a function of trial number suggests that the accuracy of qVFM increased with number of test trials.

The average 68.2% HWCI of the estimated contrast sensitivities [in log10 units, represent as $log(sensitivity)$] across all 10 eyes and 64 retina locations decreased from 0.33 before the first qVFM trial to 0.072 after 160 qVFM trials and 0.060 after 320 qVFM trials. The average 68.2% HWCI of the estimated contrast sensitivities decreased from 0.33 before the first qFC trial to 0.16 after 160 qFC trials, 0.10 after 320 qFC trials, and 0.041 after 1,280 qFC trials (**Figure 6B**). The results suggest that the precision of the estimated sensitivities from the qVFM and qFC methods increased with trial number, and reached 0.1 $log(sensitivity)$ in about 17 and 325 trials, respectively.

For the qVFM method, the switch from the global module to the local module occurred between 31 and 70 trials, with the mean around 41 trials and a standard deviation of 9.8 trials across all 10 eyes, consistent with the simulations. The rapid convergence of the VFM estimates by the global module (the average 68.2% HWCI) is evident in **Figure 6B**.

Figure 6C presents the average estimated VUSVFM of 10 eyes as a function of trial number for qVFM and qFC. The estimated VUSVFM from the two methods was less than 0.6% different after 320 trials. The agreement of these estimates implies that the VUSVFM can be a useful metric of the overall visual filed map.

Test–retest reliability of the qVFM is assessed through analysis of the 4 qVFM runs completed in four sessions. **Figure 7A** plots estimated sensitivities of the paired qVFM runs from the four independent sessions (2 random pairs of qVFM × 10 eyes × 64 locations = 1,280 data points). The average test–retest correlation for the all possible pairs of VFM estimates was 0.971 (SD = 0.001).

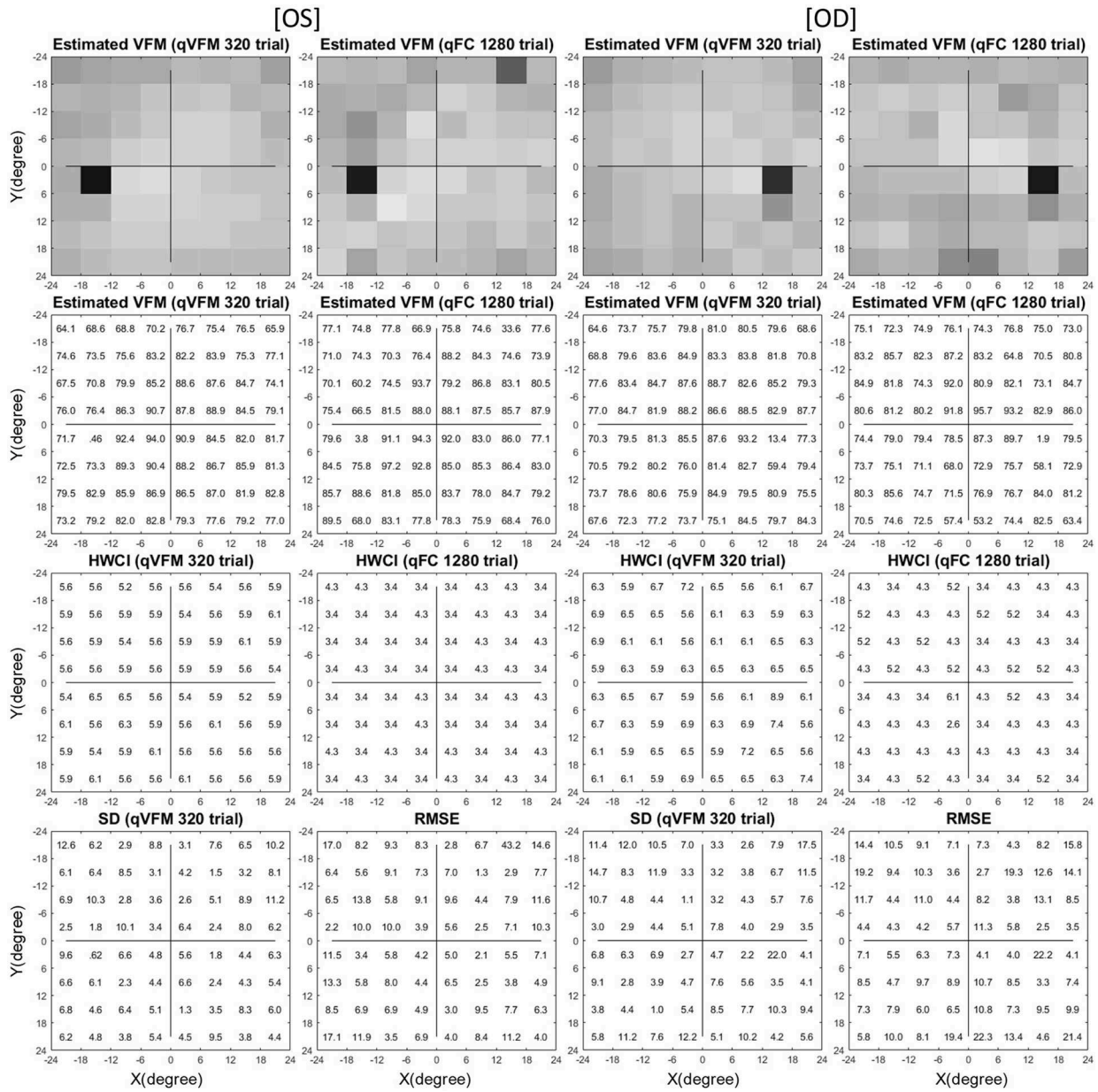


FIGURE 5 | Experimental result I (Subject 1, OS and OD). The estimated contrast sensitivity VFMs are presented in the first row with colormaps and second row with numerical values (unit: 100 x *log(sensitivity)*). For each visual field location of the estimated VFM, the 68.2% HWCI is presented in the third row, the standard deviation of the estimated contrast sensitivity from the 4 repeated qVFM assessments and the $RMSE^{eyes}$ between the qVFM and qFC estimates are presented in the fourth row. The results obtained from OS are displayed in the first and second columns, and OD in the third and fourth columns, respectively. The results from the qVFM and qFC methods are displayed in different columns.

Although test–retest correlation is widely reported as a measure of test–retest reliability, it might not be the most useful way to characterize the reliability of a method (Bland and Altman, 1986). **Figure 7B** presents a Bland–Altman plot of the difference of the qVFM estimates between all possible pairs of repeated measures against their respective means. The mean and standard deviation of the test–retest difference were 1.3×10^{-4} and 0.093 *log(sensitivity)*, respectively. These results suggest that (1) the estimated VFM did not change much over the course of testing sessions, and (2) the test–retest differences between sessions were comparable to the estimated $RMSE^{eyes}$ [0.093 *log(sensitivity)* vs. 0.11 *log(sensitivity)*]. Repeated runs of the qVFM procedure generated quite consistent results, demonstrating its robustness.

To illustrate the convergence of the estimated VUSVFM obtained with the qVFM method, **Figure 7C** presents the

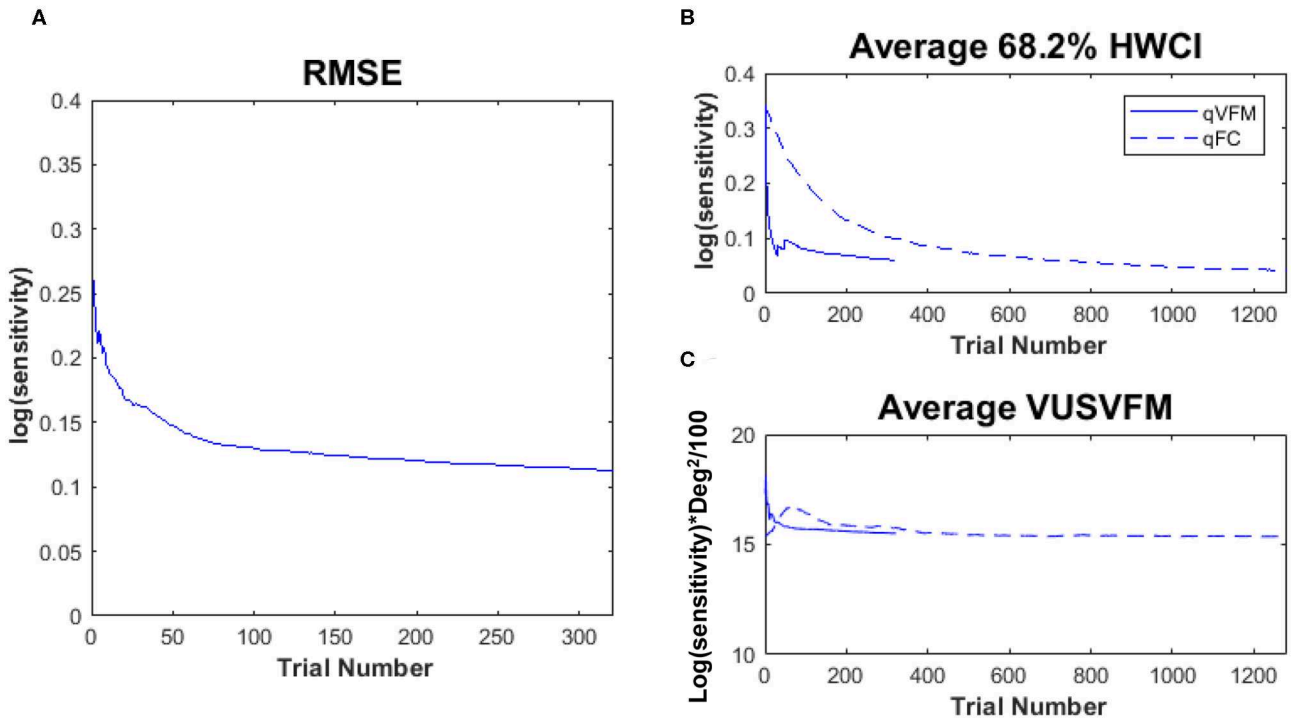


FIGURE 6 | Experimental results II. **(A)** $RMSE^{eyes}$ of the estimated sensitivities from qVFM as a function of trial number, using estimated sensitivities from 1,280 qFC trials as the "truth." **(B)** Average 68.2% HWCI of the estimated sensitivities across 64 locations and 10 eyes. **(C)** Average VUSVFM across 10 eyes. Results from the qVFM method are shown in solid lines, and results from the qFC method are shown in dashed lines.

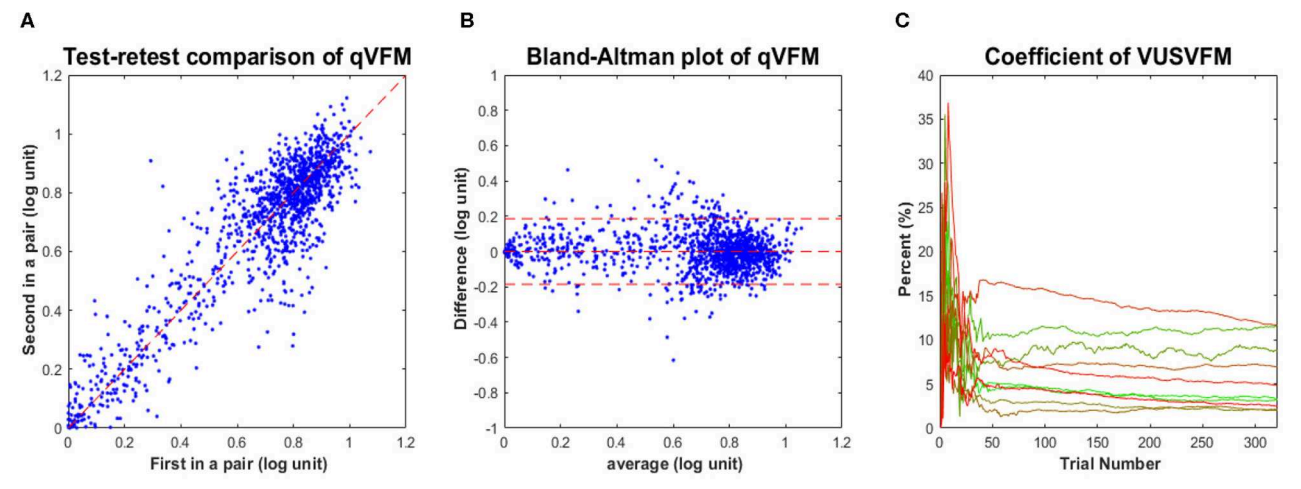


FIGURE 7 | Experimental results III. **(A)** Test-retest comparison of estimated sensitivities from repeated qVFM runs. **(B)** Bland-Altman plot for repeated qVFM runs. **(C)** Coefficient of variability of estimated VUSVFMs (4 runs each) as functions of trial number for the 10 tested eyes.

coefficient of variation of VUSVFM estimates as a function of trial number for each eye. The coefficient of variation, also known as relative standard deviation, is defined as the ratio of the standard deviation to the mean:

$$cv_i = \frac{\sigma_i}{\mu_i} \tag{9}$$

where σ_i is the standard deviation of estimated VUSVFM after the i-th trial across four runs, and μ_i is the mean of the estimated VUSVFMs after the i-th trial across four runs. A consistent pattern, exhibited in each tested eye, is a decrease of variability with trial number: from close to 35% after 20 trials, to less than 12% after 320 trials.

DISCUSSION

Visual filed mapping has undergone revolutionary changes over the past 2000 years, particularly with regard to instrumentation, standardization, quantitative assessment, statistical evaluation, optimization of accuracy, precision and efficiency of testing, and distribution of results (Lascaratos and Marketos, 1988; Walsh, 2010). However, the primary method for performing perimetry tests has remained relatively the same for more than 200 years. Thus, it is both a challenge and an opportunity for us to augment current methods by developing new procedures with novel algorithms that would allow more comprehensive and precise identification of damage to the visual field (Thompson and Wall, 2010; Johnson et al., 2011).

We developed the qVFM method to address this technical challenge in mapping visual functions, based on a hybrid Bayesian adaptive testing framework that combines a global module for preliminary assessment of the VFM's shape and a local module for assessing individual VF locations. We first applied the method to assess light sensitivity VFM in an earlier study. In the current study, we extended the method to assess contrast sensitivity of the visual field, and showed that the method can provide an accurate, precise, efficient assessment. Our simulations showed that the average $RMSE^{simulation}$ and SD of the estimated VFM [in log10 units, represent as *log*(*sensitivity*)] after 1,280 trials were 0.037 and 0.032 by the qVFM, and 0.041 and 0.041 by the qFC, respectively. To achieve 0.1 accuracy and 0.1 precision, on average, it took 106 qVFM trials, and 334 qFC trials. Estimates of within-run variability (68.2% HWCIs) were comparable to cross-run variability (SD). For the subjects in our psychophysical experiment, the average HWCI of the qVFM estimates decreased from 0.33 on the first trial to 0.072 after 160 trials, and to 0.060 after 320 trials. The $RMSE^{eyes}$ of the estimates from the qVFM and qFC methods started at 0.26 on the first trial and decreased to 0.12 after 160 qVFM trials and to 0.11 after 320 trials.

In addition to light sensitivity and contrast sensitivity, the qVFM method can be extended to map many other visual functions, such as visual acuity, binocular vision, color vision, temporal frequency, motion sensitivity, reading speed, and crowding maps, with potential clinical signals for monitoring vision loss, evaluating therapeutic interventions, and developing effective rehabilitation for low vision.

The development of the qVFM and other related methods, such as the qCSF, qVA, and qReading methods (Lesmes et al., 2010; Hou et al., 2018; Lesmes and Dorr, 2019; Shepard et al., 2019; Zhao et al., 2019a), makes it possible for us to identify core deficits of functional vision in visual impairments. By measuring performance in a battery of everyday visual tasks on a large group of subjects, we can model their performance in everyday visual tasks with the candidate metrics provided by the tests (e.g., light sensitivity, contrast sensitivity, acuity, reading speed) and identify the most important core metrics. Such core metrics would allow us to better understand visual deficits, to focus on a reduced set of measures while achieving a thorough assessment of residual vision, and to setup portfolio of effective examinations and rehabilitation interventions.

Mapping Sensitivities With m-AFC Tasks

Earlier adaptive methods focused on targeting pre-defined percent correct performance levels on the empirical psychometric function. Following the development of staircase procedures (Von Békésy, 1947; Wetherill, 1963; Wetherill and Levitt, 1965), the QUEST method (Watson and Pelli, 1983) was a landmark application of Bayesian adaptive inference to measure thresholds. The Bayesian adaptive approach has since been applied to measure empirical thresholds in forced-choice tasks (Watson and Pelli, 1983; King-Smith et al., 1994; King-Smith and Rose, 1997; Snoeren and Puts, 1997; Alcala-Quintana and Garcia-Perez, 2007; García-Pérez and Alcalá-Quintana, 2007).

Previous studies (Leek et al., 1992; Leek, 2001; Alcalá-Quintana and García-Pérez, 2004; Hou et al., 2010) have revealed that the shape of the psychometric function could have a profound impact on the efficiency of adaptive procedures that search optimal stimuli in a two-dimensional stimulus space. In a particular experimental setting, the slope of the d' psychometric function is related to the internal noise distribution and transducer of the observer (Dosher and Lu, 1998; Lu and Dosher, 2008, 2013) and is not easy to manipulate. However, for a single d' psychometric function, it is possible to reduce the guessing rate and increase the slope of the percent correct psychometric function by increasing the number of alternatives in an m-AFC task, and therefore increase the efficiency of the adaptive procedure. The benefit of more alternatives in m-AFC tasks was documented in association with the qCSF method (Hou et al., 2015), and has been extended to the qVFM procedure in this study. It can also be extended to other Bayesian adaptive testing procedures such as QUEST, ZEST, Psi, quick TvC, quick Partial Report, qReading and quick Change-Detection, most of which are based on d' psychometric functions (King-Smith et al., 1994; Kontsevich and Tyler, 1999; Kujala and Lukka, 2006; Lesmes et al., 2006; Baek et al., 2016; Hou et al., 2018; Shepard et al., 2019; Zhang et al., 2019; Zhao et al., 2019b).

Effects of the Prior

It is well-known that the initial prior probability distribution could change the starting point of parameter estimation

TABLE 2 | The parameters of four prior settings.

Priors	θ_i	EPA	EPB	EPZ	SLA	SLB
WP	$\theta_{i,guess}$	72	54	0.6	0.002	0.003
	$\theta_{confidence}$	1.7	1.3	1.2	7.0	6.4
WI	$\theta_{i,guess}$	66	48	0.9	0.003	0.001
	$\theta_{confidence}$	1.7	1.3	1.2	7.0	6.4
SP	$\theta_{i,guess}$	72	54	0.6	0.002	0.003
	$\theta_{confidence}$	8.4	7.8	6.8	28	25
SI	$\theta_{i,guess}$	66	48	0.9	0.003	0.001
	$\theta_{confidence}$	8.4	7.8	6.8	28	25

Unit of EPA, EPB is degree/$\sqrt{log(sensitivity)}$, unit of EPZ is log(sensitivity), unit of SLA, SLB is log(sensitivity)/degree.

and the efficiency of the estimation process (Baek et al., 2016; Gu et al., 2016). For the three simulated observers in the current study, the prior distributions were moderately informative. To illustrate the effects of the prior, we conducted an additional set of simulations with four different prior settings (**Table 2**): (a) a weakly informative proper prior, (b) a weakly informative but improper prior, (c) a strong informative proper prior, and (d) a strong informative but improper prior.

The parameters of the simulated observer were: EPA = 72 (degree/$\sqrt{log(sensitivity)}$), EPB = 54 (degree/$\sqrt{log(sensitivity)}$), EPZ = 0.6 ($log(sensitivity)$), SLA = 0.002 ($log(sensitivity)$/degree), SLB = 0.003 ($log(sensitivity)$/degree). The parameter space and the stimulus space remained the same.

Here, we introduce the absolute bias as the index of accuracy. The average absolute bias of the estimated threshold across all locations after the i-th trial can be calculated as:

$$abs\ Bias_i = \frac{\sum_k \left|\sum_j \left(\tau_{ijk} - \tau_k^{true}\right)\right|}{J \times K} \tag{10}$$

where τ_{ijk} is the estimated contrast sensitivity in the k-th retina location after i trials in the j-th run, and τ_k^{true} is the true sensitivity of that location.

Figure 8 shows the performance of the qVFM and qFC procedures with the four different prior settings. In both the qVFM and qFC procedures, the strong informative proper prior led to the best performance in terms of the average absolute

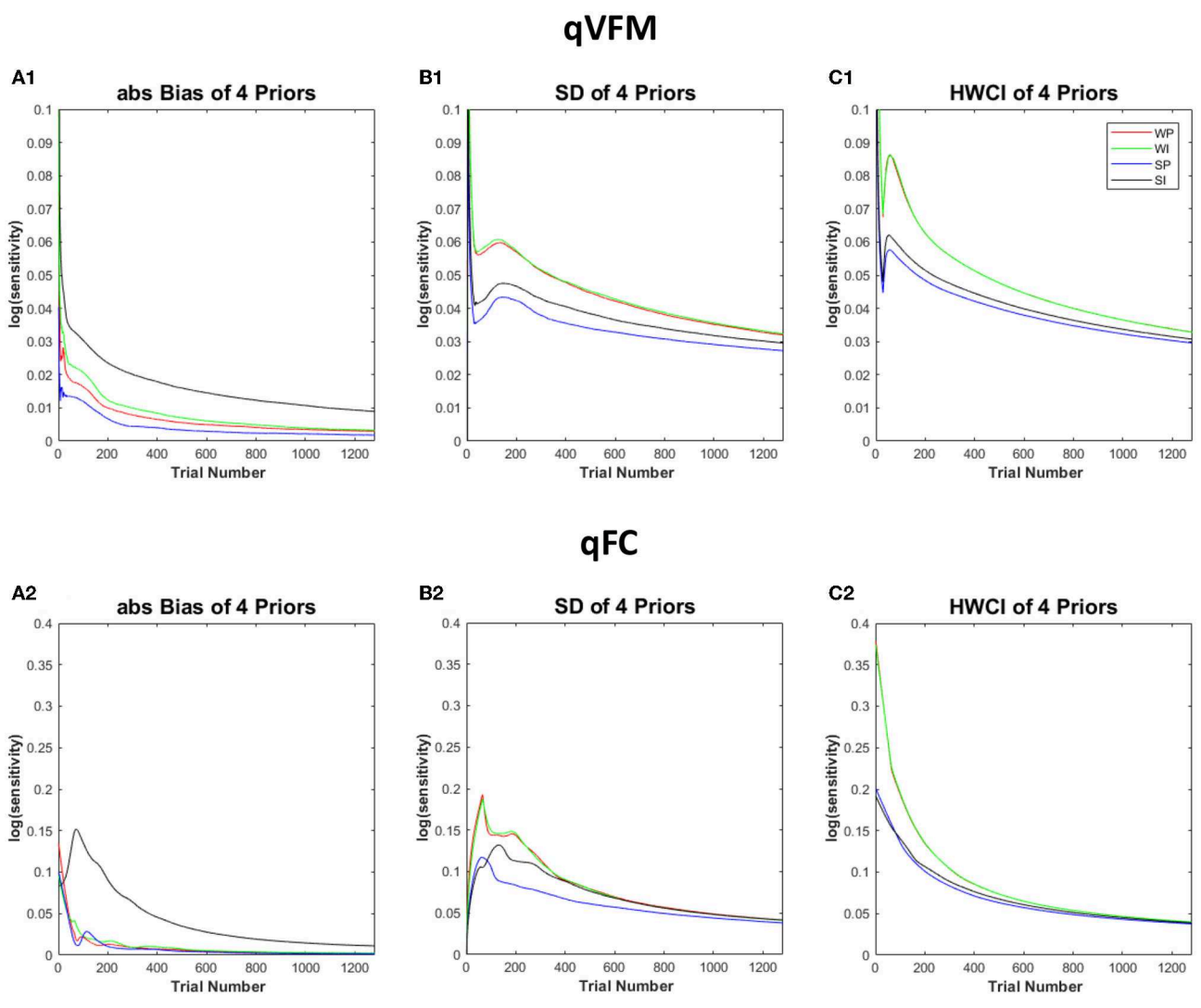


FIGURE 8 | Simulation results III. The performance of the qVFM and qFC procedures in four prior settings. WP, weakly informative proper prior; WI, weakly informative but improper prior; SP, strong informative proper prior; SI, strong informative but improper prior. **(A1, B1, C1)** The average absolute Bias, SD, the average 68.2% HWCI for qVFM procedures in each of four prior settings. **(A2, B2, C2)** The average absolute Bias, SD, the average 68.2% HWCI for qFC procedures in each of four prior settings.

bias, SD and the average 68.2% HWCI. The strong informative improper prior led to the worst average absolute bias and slightly better precision than the weakly informative priors. The weakly informative proper and improper priors exhibited similar performance in all measures, with accuracy between those of the strong informative proper and improper priors, and worse precision comparing to them.

In all four prior settings, the qVFM procedure led to better performance than the qFC procedure. Especially with the strong informative priors, the improper prior made the accuracy of the qFC estimates much worse. The difference of the average absolute bias between the strong informative proper and improper priors was 0.05 for qFC and 0.016 for qVFM after 320 trials. The results suggest that the qVFM method was more robust than the qFC method when the prior was improper.

These results suggest that proper informative prior can speed up the estimation process of the qVFM procedure. We can inform the prior with previous knowledge or pilot data, such as the representative parameters from a particular patient population, or priors derived with the hierarchical adaptive method (Kim et al., 2014; Gu et al., 2016).

CONCLUSION

In this study, we implemented the qVFM method to measure contrast sensitivity VFM with a 10-alternative forced-choice paradigm. Detailed assessment of contrast sensitivity across the visual field and other core metrics of functional visual is critical for quantifying the effectiveness of new drugs and rehabilitation therapies. We have tested our method on 10 eyes of five normal observers. Applications of our method to clinical populations may require additional development. Further integrating with other measurements, such as fundus or OCT images, may further improve the efficiency of the qVFM method. The broad adoption of the qVFM method can potentially improve both clinical research and clinical care.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of the Ohio State University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Z-LL, PX, LL, and DY designed the qVFM algorithms. PX performed simulations, carried out the experiment, and analyzed the data. PX and Z-LL wrote the manuscript with input from all authors. Z-LL and DY supervised the project. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by NIH grants EY025658 to DY and EY021553 to Z-LL.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2020.00665/full#supplementary-material

Supplementary Movie 1 | A simulated 320-trial sequence of the qVFM and qFC procedures based on the parameters of simulated observer 1. The first row shows the contrast sensitivity VFM of the simulated observer, and the estimated contrast sensitivity VFM's obtained with the qVFM and qFC methods. The second and third rows show the estimated VUSVFM and average HWCI of the estimated contrast sensitivities.

REFERENCES

Alcalá-Quintana, R., and García-Pérez, M. A. (2004). The role of parametric assumptions in adaptive Bayesian estimation. *Psychol. Methods* 9, 250. doi: 10.1037/1082-989X.9.2.250

Alcala-Quintana, R., and Garcia-Perez, M. A. (2007). A comparison of fixed-step-size and Bayesian staircases for sensory threshold estimation. *Spat. Vis.* 20, 197–218. doi: 10.1163/156856807780421174

Applegate, R. A., Hilmantel, G., Howland, H. C., Tu, E. Y., Starck, T., and Zayac, E. J. (2000). Corneal first surface optical aberrations and visual performance. *J. Refract. Surg.* 16, 507–514.

Applegate, R. A., Howland, H. C., Sharp, R. P., Cottingham, A. J., and Yee, R. W. (1998). Corneal aberrations and visual performance after radial keratotomy. *J. Refract. Surg.* 14, 397–407. doi: 10.3928/1081-597X-19980701-05

Artes, P. H., Iwase, A., Ohno, Y., Kitazawa, Y., and Chauhan, B. C. (2002). Properties of perimetric threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies. *Invest. Ophthalmol. Vis. Sci.* 43, 2654–2659.

Baek, J., Lesmes, L. A., and Lu, Z.-L. (2016). qPR: an adaptive partial-report procedure based on Bayesian inference. *J. Vis.* 16:25. doi: 10.1167/16.10.25

Balas, B., Nakano, L., and Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *J. Vis.* 9:13. doi: 10.1167/9.12.13

Bellucci, R., Scialdone, A., Buratto, L., Morselli, S., Chierego, C., Criscuoli, A., et al. (2005). Visual acuity and contrast sensitivity comparison between Tecnis and AcrySof SA60AT intraocular lenses: a multicenter randomized study. *J. Cataract Refract. Surg.* 31, 712–717. doi: 10.1016/j.jcrs.2004.08.049

Bland, J. M., and Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327, 307–310. doi: 10.1016/S0140-6736(86)90837-8

Bodis-Wollner, I., Marx, M. S., Mitra, S., Bobak, P., Mylin, L., and, Yahr, M. (1987). Visual dysfunction in Parkinson's disease: Loss in spatiotemporal contrast sensitivity. *Brain* 110, 1675–1698. doi: 10.1093/brain/110.6.1675

Bradley, A., and Freeman, R. D. (1981). Contrast sensitivity in anisometropic amblyopia. *Invest. Ophthalmol. Vis. Sci.* 21, 467–476.

Brainard, D. H., and Vision, S. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357

Bulens, C., Meerwaldt, J. D., Van der Wildt, G. J., and Keemink, C. J. (1986). Contrast sensitivity in Parkinson's disease. *Neurology* 36, 1121–1121. doi: 10.1212/WNL.36.8.1121

Carlow, T. J., Flynn, J. T., and Shipley, T. (1976). Color perimetry. *Arch. Ophthalmol* 94, 1492–1496. doi: 10.1001/archopht.1976.03910040326007

Comerford, J. P. (1983). Vision evaluation using contrast sensitivity functions. *Am. J. Optom. Physiol. Opt.* 60, 394–398. doi: 10.1097/00006324-198305000-00009

Daitch, J. M., and Green, D. G. (1969). *Contrast sensitivity of the human peripheral* retina. *Vision Res.* 9:947–952. doi: 10.1016/0042-6989(69)90100-X

Della Sala, S., Bertoni, G., Somazzi, L., Stubbe, F., and Wilkins, A. J. (1985). Impaired contrast sensitivity in diabetic patients with and without retinopathy: a new technique for rapid assessment. *Br. J. Ophthalmol.* 69, 136–142. doi: 10.1136/bjo.69.2.136

Dorr, M., Wille, M., Viulet, T., Sanchez, E., Bex, P. J., Lu, Z.-L., et al. (2015). Next-generation vision testing: the quick CSF. *Curr. Dir. Biomed. Eng.* 1, 131–134. doi: 10.1515/cdbme-2015-0034

Dosher, B. A., and Lu, Z.-L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci. U.S.A.* 95, 13988–13993. doi: 10.1073/pnas.95.23.13988

Dreyer, E. B. (1993). Automated static perimetry. *Arch. Ophthalmol.* 111:310. doi: 10.1001/archopht.1993.01090030028017

Elliott, D. B., and Whitaker, D. (1992). Clinical contrast sensitivity chart evaluation. *Ophthalmic Physiol. Opt.* 12, 275–280. doi: 10.1111/j.1475-1313.1992.tb00397.x

Faye, E. E. (2005). "Contrast sensitivity tests in predicting visual function," in *International Congress Series* (London, UK: Elsevier), 521–524. doi: 10.1016/j.ics.2005.05.001

Fletcher, D. C., and Schuchard, R. A. (1997). Preferred retinal loci relationship to macular scotomas in a low-vision population. *Ophthalmology* 104, 632–638. doi: 10.1016/S0161-6420(97)30260-7

Foley, J. M., and Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Res.* 21, 1041–1053. doi: 10.1016/0042-6989(81)90009-2

García-Pérez, M. A., and Alcalá-Quintana, R. (2007). The transducer model for contrast detection and discrimination: formal relations, implications, and an empirical test. *Spat. Vis.* 20, 5–43. doi: 10.1163/156856807779369724

Ginsburg, A. P. (1987). Contrast sensitivity, drivers' visibility, and vision standards. *Transp. Res. Rec.* 1149, 32–39.

Ginsburg, A. P. (2003). Contrast sensitivity and functional vision. *Int. Ophthalmol. Clin.* 43, 5–15. doi: 10.1097/00004397-200343020-00004

Ginsburg, A. P. (2006). Contrast sensitivity: determining the visual quality and function of cataract, intraocular lenses and refractive surgery. *Curr. Opin. Ophthalmol.* 17, 19–26. doi: 10.1097/01.icu.0000192520.48411.fa

Gu, H., Kim, W., Hou, F., Lesmes, L. A., Pitt, M. A., Lu, Z.-L., et al. (2016). A hierarchical Bayesian approach to adaptive vision testing: a case study with the contrast sensitivity function. *J. Vis.* 16, 15–15. doi: 10.1167/16.6.15

Gu, X., and Green, D. M. (1994). Further studies of a maximum-likelihood yes–no procedure. *J. Acoust. Soc. Am.* 96, 93–101. doi: 10.1121/1.410378

Hacker, M. J., and Ratcliff, R. (1979). A revised table of d'for M-alternative forced choice. *Atten. Percept. Psychophys.* 26, 168–170. doi: 10.3758/BF03208311

Hart, W. M., Hartz, R. K., Hagen, R. W., and Clark, K. W. (1984). Color contrast perimetry. *Invest. Ophthalmol. Vis. Sci.* 25, 400–413.

Hess, R. F., and Howell, E. R. (1977). The threshold contrast sensitivity function in strabismic amblyopia: evidence for a two type classification. *Vision Res.* 17, 1049–1055. doi: 10.1016/0042-6989(77)90009-8

Hess, R. F., and Plant, G. T. (1985). Temporal frequency discrimination in human vision: evidence for an additional mechanism in the low spatial and high temporal frequency region. *Vision Res.* 25, 1493–1500. doi: 10.1016/0042-6989(85)90227-5

Hot, A., Dul, M. W., and Swanson, W. H. (2008). Development and evaluation of a contrast sensitivity perimetry test for patients with glaucoma. *Invest. Ophthalmol. Vis. Sci.* 49, 3049–3057. doi: 10.1167/iovs.07-1205

Hou, F., Huang, C., Lesmes, L. A., Feng, L., Tao, L., Zhou, Y., et al. (2010). qCSF in clinical application: efficient characterization and classification of contrast sensitivity functions in amblyopia. *Invest. Ophthalmol. Vis. Sci.* 51, 5365–5377. doi: 10.1167/iovs.10-5468

Hou, F., Lesmes, L. A., Bex, P. J., Dorr, M., and Lu, Z.-L. (2015). Using 10AFC to further improve the efficiency of the quick CSF method. *J. Vis.* 15:2. doi: 10.1167/15.9.2

Hou, F., Zhao, Y., Lesmes, L. A., Bex, P., Yu, D., and Lu, Z.-L. (2018). Bayesian adaptive assessment of the reading function for vision: the qReading method. *J. Vis.* 18:6. doi: 10.1167/18.9.6

Huang, C., Zhou, Y., and Lu, Z.-L. (2008). Broad bandwidth of perceptual learning in the visual system of adults with anisometropic amblyopia. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4068–4073. doi: 10.1073/pnas.0800824105

Jia, W., Yan, F., Hou, F., Lu, Z.-L., and Huang, C.-B. (2014). qCSF in clinical applications: efficient characterization and classification of contrast sensitivity functions in aging. *Invest. Ophthalmol. Vis. Sci.* 55:762. doi: 10.1167/iovs.10-5468

Jindra, L. F., and Zemon, V. (1989). Contrast sensitivity testing: a more complete assessment of vision. *J. Cataract Refract. Surg.* 15, 141–148. doi: 10.1016/S0886-3350(89)80002-1

Johnson, C. A., Wall, M., and Thompson, H. S. (2011). A history of perimetry and visual field testing. *Optom. Vis. Sci.* 88, E8–E15. doi: 10.1097/OPX.0b013e3182004c3b

Keltgen, K. M., and Swanson, W. H. (2012). Estimation of spatial scale across the visual field using sinusoidal stimuli. *Invest. Ophthalmol. Vis. Sci.* 53, 633–639. doi: 10.1167/iovs.10-6674

Keltner, J. L., Johnson, C. A., Quigg, J. M., Cello, K. E., Kass, M. A., and Gordon, M. O. (2000). Confirmation of visual field abnormalities in the Ocular Hypertension Treatment Study. *Arch. Ophthalmol.* 118, 1187–1194. doi: 10.1001/archopht.118.9.1187

Kim, W., Pitt, M. A., Lu, Z.-L., Steyvers, M., and Myung, J. I. (2014). A hierarchical adaptive approach to optimal experimental design. *Neural Comput.* 26, 2465–2492. doi: 10.1162/NECO_a_00654

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., and Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision Res.* 34, 885–912. doi: 10.1016/0042-6989(94)90039-6

King-Smith, P. E., and Rose, D. (1997). Principles of an adaptive method for measuring the slope of the psychometric function. *Vision Res.* 37, 1595–1604. doi: 10.1016/S0042-6989(96)00310-0

Kiorpes, L., Tang, C., and Movshon, J. A. (1999). Factors limiting contrast sensitivity in experimentally amblyopic macaque monkeys. *Vision Res.* 39, 4152–4160. doi: 10.1016/S0042-6989(99)00130-3

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: a commentary. *Percept. Psychophys.* 63, 1421–1455. doi: 10.3758/BF03194552

Kontsevich, L. L., and Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Res.* 39, 2729–2737. doi: 10.1016/S0042-6989(98)00285-5

Kujala, J. V., and Lukka, T. J. (2006). Bayesian adaptive estimation: The next dimension. *J. Math. Psychol.* 50, 369–389. doi: 10.1016/j.jmp.2005.12.005

Lascaratos, J., and Marketos, S. (1988). "A historical outline of Greek ophthalmology from the Hellenistic period up to the establishment of the first universities," in *History of Ophthalmology 1*, eds H. E. Henkes and Cl. Zrenner (Dordrecht: Springer), 157–169. doi: 10.1007/978-94-009-1307-3_17

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Percept. Psychophys.* 63, 1279–1292. doi: 10.3758/BF03194543

Leek, M. R., Hanna, T. E., and Marshall, L. (1992). Estimation of psychometric functions from adaptive tracking procedures. *Percept. Psychophys.* 51, 247–256. doi: 10.3758/BF03212251

Legge, G. E., Kersten, D., and Burgess, A. E. (1987). Contrast discrimination in noise. *JOSA A* 4, 391–404. doi: 10.1364/JOSAA.4.000391

Lesmes, L. A., and Dorr, M. (2019). "Active learning for visual acuity testing," *in Proceedings of the 2nd International Conference on Applications of Intelligent Systems*, (Las Palmas de Gran Canaria) 1–6. doi: 10.1145/3309772.3309798

Lesmes, L. A., Jeon, S.-T., Lu, Z.-L., and Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: the quick TvC method. *Vision Res.* 46, 3160–3176. doi: 10.1016/j.visres.2006.04.022

Lesmes, L. A., Lu, Z.-L., Baek, J., and Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J. Vis.* 10, 17–17. doi: 10.1167/10.3.17

Lesmes, L. A., Lu, Z.-L., Baek, J., Tran, N., Dosher, B. A., and Albright, T. D. (2015). Developing Bayesian adaptive methods for estimating sensitivity thresholds (d') in Yes-No and forced-choice tasks. *Front. Psychol.* 6:1070. doi: 10.3389/fpsyg.2015.01070

Levi, D. M., and Carney, T. (2009). Crowding in peripheral vision: why bigger is better. *Curr. Biol.* 19, 1988–1993. doi: 10.1016/j.cub.2009.09.056

Li, R., Polat, U., Makous, W., and Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nat. Neurosci.* 12, 549. doi: 10.1038/nn.2296

Li, R. W., Young, K. G., Hoenig, P., and Levi, D. M. (2005). Perceptual learning improves visual performance in juvenile amblyopia. *Invest. Ophthalmol. Vis. Sci.* 46, 3161–3168. doi: 10.1167/iovs.05-0286

Loshin, D. S., and White, J. (1984). Contrast sensitivity: the visual rehabilitation of the patient with macular degeneration. *Arch. Ophthalmol.* 102, 1303–1306. doi: 10.1001/archopht.1984.01040031053022

Lu, Z.-L., and Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *JOSA A* 16, 764–778. doi: 10.1364/JOSAA.16.000764

Lu, Z.-L., and Dosher, B. A. (2008). Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychol. Rev.* 115, 44. doi: 10.1037/0033-295X.115.1.44

Lu, Z.-L., Zhao, Y., Lesmes, L. A., Dorr, M., and Bex, P. (2019). Unbiased threshold estimates in Bayesian Adaptive qCSF and qFC with mismatched psychometric function slopes. *Invest. Ophthalmol. Vis. Sci.* 60:3908.

Lu, Z.-L., and Dosher, B. A. (2013). *Visual Psychophysics: From Laboratory to Theory*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/9780262019453.001.0001

Markowitz, S. N., and Muller, C. (2004). Macular perimetry in low vision. *Can. J. Ophthalmol.* 39, 56–60. doi: 10.1016/S0008-4182(04)80053-X

Massof, R. W., and Rubin, G. S. (2001). Visual function assessment questionnaires. *Surv. Ophthalmol.* 45, 531–548. doi: 10.1016/S0039-6257(01)00194-1

McLeod, S. D. (2001). Beyond snellen acuity: the assessment of visual function after refractive surgery. *Arch. Ophthalmol.* 119, 1371–1373. doi: 10.1001/archopht.119.9.1371

Mestre, D., Blin, O., Serratrice, G., and Pailhous, J. (1990). Spatiotemporal contrast sensitivity differs in normal aging and Parkinson's disease. *Neurology* 40:1710. doi: 10.1212/WNL.40.11.1710

Oshika, T., Klyce, S. D., Applegate, R. A., and Howland, H. C. (1999). Changes in corneal wavefront aberrations with aging. *Invest. Ophthalmol. Vis. Sci.* 40, 1351–1355.

Oshika, T., Okamoto, C., Samejima, T., Tokunaga, T., and Miyata, K. (2006). Contrast sensitivity function and ocular higher-order wavefront aberrations in normal human eyes. *Ophthalmology* 113, 1807–1812. doi: 10.1016/j.ophtha.2006.03.061

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366

Pelli, D. G., and Robson, J. G. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clin .Vision Sci.* 2, 187–199.

Polat, U., Ma-Naim, T., Belkin, M., and Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6692–6697. doi: 10.1073/pnas.0401200101

Qiu, Z., Xu, P., Zhou, Y., and Lu, Z.-L. (2007). Spatial vision deficit underlies poor sine-wave motion direction discrimination in anisometropic amblyopia. *J. Vis.* 7:7. doi: 10.1167/7.11.7

Ramulu, P. Y., West, S. K., Munoz, B., Jampel, H. D., and Friedman, D. S. (2009). Glaucoma and reading speed: the Salisbury Eye Evaluation project. *Arch. Ophthalmol.* 127, 82–87. doi: 10.1001/archophthalmol.2008.523

Regan, D., Bartol, S., Murray, T. J., and Beverley, K. I. (1982). Spatial frequency discrimination in normal vision and in patients with multiple sclerosis. *Brain* 105, 735–754. doi: 10.1093/brain/105.4.735

Regan, D., and Hamstra, S. (1991). Shape discrimination for motion-defined and contrast-defined form: Squareness is special. *Perception* 20, 315–336. doi: 10.1068/p200315

Regan, D., Raymond, J., Ginsburg, A. P., and Murray, T. J. (1981). Contrast sensitivity, visual acuity and the discrimination of Snellen letters in multiple sclerosis. *Brain* 104, 333–350. doi: 10.1093/brain/104.2.333

Ross, J. E., Bron, A. J., and Clarke, D. D. (1984). Contrast sensitivity and visual disability in chronic simple glaucoma. *Br. J. Ophthalmol.* 68, 821–827. doi: 10.1136/bjo.68.11.821

Sample, P. A., and Weinreb, R. N. (1990). Color perimetry for assessment of primary open-angle glaucoma. *Invest. Ophthalmol. Vis. Sci.* 31, 1869–1875.

Sample, P. A., and Weinreb, R. N. (1992). Progressive color visual field loss in glaucoma. *Invest. Ophthalmol. Vis. Sci.* 33, 2068–2071.

Shepard, T. G., Hou, F., Bex, P. J., Lesmes, L. A., Lu, Z.-L., and Yu, D. (2019). Assessing reading performance in the periphery with a Bayesian adaptive approach: the qReading method. *J. Vis.* 19:5. doi: 10.1167/19.5.5

Snoeren, P. R., and Puts, M. J. (1997). Multiple parameter estimation in an adaptive psychometric method: MUEST, an extension of the QUEST method. *J. Math. Psychol.* 41, 431–439. doi: 10.1006/jmps.1997.1188

Sokol, S., Moskowitz, A., Skarf, B., Evans, R., Molitch, M., and Senior, B. (1985). Contrast sensitivity in diabetics with and without background retinopathy. *Arch. Ophthalmol.* 103, 51–54. doi: 10.1001/archopht.1985.01050010055018

Stamper, R. L. (1984). The effect of glaucoma on central visual function. *Trans. Am. Ophthalmol. Soc.* 82:792.

Stewart, W. C., and Hunt, H. H. (1993). Threshold variation in automated perimetry. *Surv. Ophthalmol.* 37, 353–361. doi: 10.1016/0039-6257(93)90065-F

Strasburger, H., Rentschler, I., and Jüttner, M. (2011). Peripheral vision and pattern recognition: a review. *J. Vis.* 11, 13–13. doi: 10.1167/11.5.13

Sunness, J. S., Schuchard, R. A., Shen, N., Rubin, G. S., Dagnelie, G., and Haselwood, D. M. (1995). Landmark-driven fundus perimetry using the scanning laser ophthalmoscope. *Invest. Ophthalmol. Vis. Sci.* 36, 1863–1874.

Swanson, W. H., Malinovsky, V. E., Dul, M. W., Malik, R., Torbit, J. K., Sutton, B. M., et al. (2014). Contrast sensitivity perimetry and clinical measures of glaucomatous damage. *Optom. Vis. Sci.* 91:1302. doi: 10.1097/OPX.0000000000000395

Tan, D. T., and Fong, A. (2008). Efficacy of neural vision therapy to enhance contrast sensitivity function and visual acuity in low myopia. *J. Cataract Refract. Surg.* 34, 570–577. doi: 10.1016/j.jcrs.2007.11.052

Thompson, H. S., Montague, P., Cox, T. A., and Corbett, J. J. (1982). The relationship between visual acuity, pupillary defect, and visual field loss. *Am. J. Ophthalmol.* 93, 681–688. doi: 10.1016/0002-9394(82)90460-3

Thompson, H. S., and Wall, M. (2010). Imaging and Perimetry Society (IPS). A history of perimetry. Available online at: http://perimetry.org/index.php/history

Travis, D., and Thompson, P. (1989). Spatiotemporal contrast sensitivity and colour vision in multiple sclerosis. *Brain* 112, 283–303. doi: 10.1093/brain/112.2.283

Trobe, J. D., Beck, R. W., Moke, P. S., and Cleary, P. A. (1996). Contrast sensitivity and other vision tests in the optic neuritis treatment trial. *Am. J. Ophthalmol,.* 121, 547–553. doi: 10.1016/S0002-9394(14)75429-7

VA, V. A. (1965). Visual acuity.

van Gaalen, K. W., Jansonius, N. M., Koopmans, S. A., Terwee, T., and Kooijman, A. C. (2009). Relationship between contrast sensitivity and spherical aberration: Comparison of 7 contrast sensitivity tests with natural and artificial pupils in healthy eyes. *J. Cataract Refract. Surg.* 35, 47–56. doi: 10.1016/j.jcrs.2008.09.016

Von Békésy, G. (1947). Uber ein neues Audiometer. *Arch Elektr Ubertragung* 1:13.

Walsh, T. (2010). *Visual Fields: Examination and Interpretation*. Oxford, NY: Oxford University Press.

Watson, A. B., and Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* 33, 113–120. doi: 10.3758/BF03202828

Weinreb, R. N., and Kaufman, P. L. (2009). The glaucoma research community and FDA look to the future: a report from the NEI/FDA CDER Glaucoma Clinical Trial Design and Endpoints Symposium. *Invest. Ophthalmol. Vis. Sci.* 50, 1497–1505. doi: 10.1167/iovs.08-2843

Weinreb, R. N., and Kaufman, P. L. (2011). Glaucoma research community and FDA look to the future, II: NEI/FDA Glaucoma Clinical Trial Design and Endpoints Symposium: measures of structural change and visual function. *Invest. Ophthalmol. Vis. Sci.* 52, 7842–7851. doi: 10.1167/iovs.11-7895

Wetherill, G. B. (1963). Sequential estimation of quantal response curves. *J. R. Stat. Soc. Ser. B Methodol.* 25, 1–38. doi: 10.1111/j.2517-6161.1963.tb00481.x

Wetherill, G. B., and Levitt, H. (1965). Sequential estimation of points on a psychometric function. *Br. J. Math. Stat. Psychol.* 18, 1–10. doi: 10.1111/j.2044-8317.1965.tb00689.x

Wichmann, F. A., and Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* 63, 1293–1313. doi: 10.3758/BF03194544

Woods, R. L., and Wood, J. M. (1995). The role of contrast sensitivity charts and contrast letter charts in clinical practice. *Clin. Exp. Optom.* 78, 43–57. doi: 10.1111/j.1444-0938.1995.tb00787.x

Xu, P., Lesmes, L. A., Yu, D., and Lu, Z.-L. (2019). A novel Bayesian adaptive method for mapping the visual field. *J. Vis.* 19:16. doi: 10.1167/19.14.16

Xu, P., Lu, Z.-L., Qiu, Z., and Zhou, Y. (2006). Identify mechanisms of amblyopia in Gabor orientation identification with external noise. *Vision Res.* 46, 3748–3760. doi: 10.1016/j.visres.2006.06.013

Yu, D., Cheung, S.-H., Legge, G. E., and Chung, S. T. (2010). Reading speed in the peripheral visual field of older adults: Does it benefit from perceptual learning? *Vision Res.* 50, 860–869. doi: 10.1016/j.visres.2010.02.006

Zhang, P., Zhao, Y., Dosher, B. A., and Lu, Z.-L. (2019). Assessing the detailed time course of perceptual sensitivity change in perceptual learning. *J. Vis.* 19:9. doi: 10.1167/19.5.9

Zhao, Y., Lesmes, L., and Lu, Z.-L. (2019b). Efficient assessment of the time course of perceptual sensitivity change. *Vision Res.* 154, 21–43. doi: 10.1016/j.visres.2018.10.009

Zhao, Y., Lesmes, L. A., Dorr, M., Bex, P., and Lu, Z.-L. (2019a). Accuracy and Precision of the ETDRS Chart, E-ETDRS and Bayesian qVA Method. *Invest. Ophthalmol. Vis. Sci.* 60:5908.

Zheng, H., Wang, C., Cui, R., He, X., Shen, M., Lesmes, L. A., et al. (2018). Measuring the contrast sensitivity function using the qCSF method with 10 Digits. *Transl. Vis. Sci. Technol.* 7:9. doi: 10.1167/tvst.7.6.9

Zhou, Y., Huang, C., Xu, P., Tao, L., Qiu, Z., Li, X., et al. (2006). Perceptual learning improves contrast sensitivity and visual acuity in adults with anisometropic amblyopia. *Vision Res.* 46, 739–750. doi: 10.1016/j.visres.2005.07.031

Zimmern, R. L., Campbell, F. W., and Wilkinson, I. M. (1979). Subtle disturbances of vision after optic neuritis elicited by studying contrast sensitivity. *J. Neurol. Neurosurg. Psychiatry* 42, 407–412. doi: 10.1136/jnnp.42.5.407

Conflict of Interest: Z-LL, PX, LL, and DY own intellectual property rights on the qVFM technology and have a pending patent on it. LL and Z-LL have equity interest in Adaptive Sensory Technology, Inc. LL holds employment at AST.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 25 August 2020
doi: 10.3389/fnins.2020.00868

Contrast-Dependence of Temporal Frequency Tuning in Mouse V1

*Daniela Camillo†, Mehran Ahmadlou† and J. Alexander Heimel**

Cortical Structure and Function Group, Netherlands Institute for Neuroscience, Institute of the Royal Academy of Arts and Sciences, Amsterdam, Netherlands

The perception of speed is influenced by visual contrast. In primary visual cortex (V1), an early stage in the visual perception pathway, the neural tuning to speed is directly related to the neural tuning to temporal frequency of stimulus changes. The influence of contrast on speed perception can be caused by the joint dependency of neural responses in V1 on temporal frequency and contrast. Here, we investigated how tuning to contrast and temporal frequency in V1 of anesthetized mice are related. We found that temporal frequency tuning is contrast-dependent. V1 was more responsive at lower temporal frequencies than the dLGN, consistent with previous work at high contrast. The temporal frequency tuning moves toward higher temporal frequencies with increasing contrast. The low half-maximum temporal frequency does not change with contrast. The Heeger divisive normalization equation provides a good fit to many response characteristics in V1, but does not fit the dependency of temporal frequency and contrast with set of parameters for all temporal frequencies. Different mechanisms for normalization in the visual cortex may predict different relationships between temporal frequency and contrast non-linearity. Our data could help to make a model selection.

Keywords: V1, contrast, temporal frequency, divisive normalization, mouse, visual cortex

OPEN ACCESS

Edited by:
Fang Hou,
Wenzhou Medical University, China

Reviewed by:
Dajun Xing,
Beijing Normal University, China
Rémy Allard,
Université Pierre et Marie Curie, France
Riccardo Storchi,
University of Modena and Reggio Emilia, Italy

***Correspondence:**
J. Alexander Heimel
heimel@nin.knaw.nl

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *09 April 2020*
Accepted: *27 July 2020*
Published: *25 August 2020*

Citation:
Camillo D, Ahmadlou M and Heimel JA (2020) Contrast-Dependence of Temporal Frequency Tuning in Mouse V1. Front. Neurosci. 14:868. doi: 10.3389/fnins.2020.00868

INTRODUCTION

While the signals that are produced by an image and leave the retina are dependent on the overall level of contrast, the interpretation of an image is largely independent of the overall contrast (Avidan et al., 2002). Reducing the contrast makes an image harder to see, but does not change its interpretation. Although we have some insight on how this independence of contrast arises by thresholding, we have no detailed understanding of this process even at the first stage of cortical visual processing. In the primary visual cortex (V1), neurons are responsive to local differences in image contrast, edges in particular (Hubel and Wiesel, 1959). In a good approximation, V1 neurons operate as spatiotemporal filters of the image contrast. Most investigations have focused on the interaction of spatial frequency filtering and contrast of grating stimuli. Initially, responses of V1 neurons were thought to be separable for contrast and spatial frequency, meaning that responses are the product of a function depending on stimulus contrast and a function depending on the spatial frequency (Albrecht and Hamilton, 1982). Later, it became clear that spatial frequency tuning and contrast are not completely inseparable in V1 in cat (Skottun et al., 1986), monkey (Sceniak et al., 2002; Priebe et al., 2006), and mouse (Heimel et al., 2010).

Likewise, the temporal frequency tuning and the contrast response of neurons in early visual cortical areas were first considered to be independent (Foster et al., 1985), but later found to depend on each other in macaque, cat and ferret V1 (Albrecht, 1995; Alitto and Usrey, 2004;

Priebe et al., 2006) and macaque MT (Krekelberg et al., 2006; Pawar et al., 2019). In V1, temporal frequency tuning and speed are directly linked, because spatial and temporal frequency dependencies are separable in most of the individual neuronal responses (Tolhurst and Movshon, 1975). The interdependency of temporal frequency and contrast could thus underlie the so-called Thompson effect (Thompson, 1982; Thompson and Stone, 1997) that our perception of speed and temporal frequency is different at low contrast (Krekelberg et al., 2006). We wanted to understand the nature of the interaction of contrast and temporal frequency, and were interested to learn if this interaction is universal across mammals. In the mouse, V1 temporal frequency tuning has been measured at high contrast (Niell and Stryker, 2008; Gao et al., 2010; Durand et al., 2016), but the relationship between contrast and temporal frequency on responses has not been studied yet. Due to their small eye size, mice have about 100 times lower spatial acuity than humans (0.5 vs 60 cycles per degree; Prusky et al., 2000), but their temporal frequency tuning is more similar. In photopic conditions, contrast sensitivity in mice peaks at 1.5 Hz, six fold below humans (Burr and Ross, 1982; Umino et al., 2008), and mouse psychophysics of temporal contrast shares fundamental properties with human psychophysics (Umino et al., 2018).

We studied the responses in V1 of anesthetized mice to gratings of different temporal frequencies and contrasts. We found that responses do not factorize in contrast and temporal frequency dependencies, and that temporal frequency tuning moves to higher frequencies with higher contrast. V1 responses to many stimuli can be fitted by a divisive normalization model (Albrecht and Geisler, 1991; Heeger, 1992; Carandini and Heeger, 2011). Divisive normalization also describes the interdependency of contrast and spatial frequency of grating responses (Heimel et al., 2010). We investigated if divisive normalization also explains the relationship between contrast and temporal frequency in the responses, and if normalization operates equally across temporal frequencies. We found that, while the normalization model with a single saturation constant and exponent can approximately match V1 population responses for all combinations of temporal frequency and contrast, it does not describe the change in temporal frequency tuning with contrast for low and intermediate temporal frequencies.

MATERIALS AND METHODS

Animals

We used male, 2–4 month old, calb2-cre mice bred on a C57BL6/J background (Strain #010774, Jackson laboratory), which we also used for investigating calretinin-positive cortical interneurons (Camillo et al., 2018). All animals were kept in a 12 h day/night cycle with access to food and water *ad libitum*. The experiments were carried out during the day cycle. All experiments were approved by the animal care and use committee of the Royal Netherlands Academy of Arts and Sciences. The experiments were performed in accordance with relevant guidelines and regulations.

Extracellular Electrophysiology

Mice were injected with urethane (1.2 g per kg of mouse body weight, intraperitoneally) and chlorprothixene (8 mg per kg, subcutaneous). We injected atropine sulfate (0.1 mg per kg) to reduce mucous secretions. We maintained body temperature at 36.5–37°C with a heating pad and rectal probe. Additional doses of urethane were injected when a toe-pinch response was observed. The head was fixated with ear bars and a bite bar. During surgery, the eyes were protected from light by black stickers and from drying by Cavasan eye ointment. The scalp above visual cortex was removed and a very small craniotomy was made around 2,900–3,000 μm lateral and 300–500 μm anterior to Lambda. Laminar silicon electrodes (A1 × 16–5 mm-50–177-A16, 16 channels spaced 50 μm apart, Neuronexus) for extracellular recordings were inserted in the binocular region of V1. The signals were digitized at 24 kHz and band pass filtered between 0.5 and 10 kHz using a Tucker-Davis Technologies RX5 Pentusa. Signals were thresholded at 3× standard deviation to isolate spikes, and spikes were sorted after a principal component extraction by KlustaKwik (Harris et al., 2000) and custom-written Matlab (Mathworks) scripts.

Visual Stimuli

Stimuli were back projected by a gamma-corrected Plus U2-X1130 85 Hz DLP projector onto a screen (Macada Innovision) placed 18 cm in front of the animal. Full screen size was 60 × 42 cm. Stimuli were produced by scripts using Psychophysics Toolbox 3 (Kleiner et al., 2007) running on Matlab. We first mapped the receptive fields of the units at the recording sites by presenting a 5 min movie (5 frames per second) of small white squares (approximately 5 degrees wide) in random positions on black background (ratio of white to black area: 1:30) (Ahmadlou et al., 2018). These receptive field positions were used to ascertain that we were recording in binocular V1. The next visual stimuli were full-screen, sine-wave, drifting gratings of 0.05 cycles per degree. Drift frequencies were 0.5, 1, 2, 4, 8, and 16 Hz. This corresponded to speeds of 10, 20, 40, 80, 160 and 320 degree per second. Grating contrasts were 10, 30, 50, 70, and 90%. In each 2 s long stimulus presentation, a grating was drifting in one of the eight cardinal and oblique directions. The stimuli were shown in pseudorandom order (i.e., shuffled per block). Each combination of contrast, temporal frequency and direction was shown five times for each recording. The screen was an equiluminant gray (10 cd m^{-2}) for 1.5 s between the stimuli.

Data Analysis

Analysis was done using Matlab scripts. For all stimuli, we computed the evoked visual responses, averaged over the duration of the stimulus, minus the spontaneous rate. The spontaneous rate was defined as the mean rate in the last 0.5 s before stimulus onset. We averaged the response for each combination of contrast and temporal frequency over all drift directions. Only units were included that had a minimum response (i.e., the evoked responses minus the spontaneous rate) of 1 spikes per second for at least one combination of temporal frequency and contrast. The response dependence on

the temporal frequency was fitted with a difference of Gaussians (d.o.G.), i.e., $R(f) = R_e \exp(-1/2\, f^2/{w_e}^2) - R_i \exp(-1/2\, f^2/{w_i}^2)$, where f is the temporal frequency, R_e and w_e are the gain and width of the positive Gaussian and R_i and w_i of the wider negative Gaussian. The fits were made by minimizing the summed squared error of the fit to the mean responses for all temporal frequencies, using the Matlab fminsearch implementation of the Nelder-Mead simplex algorithm. The fit was rejected if the optimal fit was found for $R_e < 10^{-4}$ or $R_i < 10^{-4}$ or $w_e > 10 \times 16$ Hz. The d.o.G. fit was used to calculate the optimal temporal frequency and the low and high half-maximum temporal frequencies at which the responses were half of the interpolated maximum response.

The response dependence on the Michelson contrast of the stimulus was fitted with a Naka-Rushton function, i.e., $R(c) = R_m\, c^n/(\sigma^n + c^n)$, where c is the contrast and R_m, σ, and n are fitting parameters (Albrecht and Hamilton, 1982). The fits were made by minimizing the summed squared error of the fit to all responses for all contrasts plus very small contributions of σ^2 and $(n-2)^2$ to reduce the degeneracy in fitting sometimes nearly linear data, using Matlab fminsearch. From the fit, the C50 value was interpolated as the contrast at which the response would be half of the response at 100% contrast.

The explained variance per unit for the temporal frequency and contrast response fits were calculated as $1 - \{\Sigma_i (F(i) - R(i))^2\}/\{\Sigma_i (R(i) - R_m)^2\}$, where $F(i)$ are the fitted values for each frequency or contrast i, and R_m is the mean of all $R(i)$'s.

The population models in Section "Divisive Normalization" were fit by minimizing the norm of the difference between all the measured values and the fit values over all parameters, using Matlab fminsearch. For the normalization model, the optimal parameters were $\sigma = 6.9$ and $n = 0.87$. For the shunting-extended model, the optimal parameters were $\sigma = 0.5$, $\tau = 0.11$/Hz, and $n = 0.99$.

An approximation for the LGN population tuning in the mouse was made by taking the values reported in the literature for the optimal temporal frequency, and low and high half-maximum responses for the LGN population (Tang et al., 2016), respectively, 3.2, 1.5, and 6.0 Hz and fitting a d.o.G. function with the same values (there were only band pass cells in the LGN). The fit was made by a stochastic search for the d.o.G. parameters that minimized the difference between the optimal, low half-maximum and high-maximum values of the d.o.G. with the literature values.

Experimental Design and Statistics

The number of mice used for this study was determined before its start and was based on previous experience with determining feature tuning with mouse extracellular electrophysiology. The mice were randomly selected from the breeding stock. We used the Shapiro–Wilk test to test the C50 values and the optimal, low and high half-maximum values for normality. Most of these populations were not normally distributed at the 95% significance level and therefore we used non-parametric statistics for comparison and use the median and the bootstrapped standard deviation of the median as its standard error to describe the data. For comparisons of multiple populations, we used the Kruskal–Wallis test. For paired comparisons of two measurements of one population, we used the Wilcoxon signed-rank test. For comparing the fraction of units in two categories, we used the chi-square test. For testing a non-zero slope, we used the Matlab fitlm function, which applies a linear regression and computes the p-value for the t-statistic of the hypothesis test that the corresponding coefficient is equal to zero or not.

Software Accessibility

The scripts for visual stimulus display and analysis of the data are available online at https://github.com/heimel/InVivoTools.

RESULTS

We measured the response of neurons in V1 of anesthetized mice using linear silicon electrodes to drifting full-screen gratings of different contrasts and temporal frequencies (**Figures 1A–D**). For this report, we studied the 59 units in four mice that had a response larger than 1 spikes per second for at least one combination of contrast and temporal frequency. We first analyzed the temporal frequency dependence at 90% contrast, the highest stimulus contrast that we used, and reproduced previous findings (Niell and Stryker, 2008; Gao et al., 2010; Durand et al., 2016). The temporal frequency tuning of these cells could be well-fitted with a difference-of-Gaussian (d.o.G.) curve (median explained variance was 97%, examples in **Figures 1A,C**). The median optimal temporal frequency was 2.83 ± 0.14 Hz (bootstrapped standard deviation) (**Figure 1E**). The median high half-maximum temporal frequency was 7.7 ± 0.3 Hz (**Figure 1F**). Units were termed band-pass cells if they responded to stimuli shown at 0.5 Hz, the lowest temporal frequency that we tested, at less than half the interpolated response to the optimal temporal frequency. About half, 30 of 59 (51%) of the units were band-pass cells. These band-pass cells had a median low half-maximum temporal frequency of 1.09 ± 0.07 Hz (**Figure 1G**). The other half were considered low-pass cells, although the histogram of the ratio of the response at 0.5 Hz over the response at the optimal temporal frequency shows that there is no strict division between low-pass and band-pass cells (**Figure 1H**).

Varying Temporal Frequency and Contrast

Next, we considered the temporal frequency tuning across all the presented contrasts of 10, 30, 50, 70, and 90%. For 46 of the 59 units, the temporal frequency tuning could be well fitted with a d.o.G. for all contrasts. The median explained variances were 97, 97, 96, 91, and 82% for contrasts of 90, 70, 50, 30, and 10%, respectively. For the other 13 cells, the fit was too poor, because the response at the 10% was completely absent or so low that the number of repetitions that we used was insufficient to provide an accurate measurement of the response. Cells that were band-pass in temporal frequency responses for high contrasts often became low-pass at lower contrasts (**Figure 2A**). Low-pass cells at high contrasts were always also low-pass at low contrasts (**Figure 2B**). The mean temporal frequency responses for the whole population were also well described by different d.o.G. functions for the different contrasts (**Figure 2C**). The

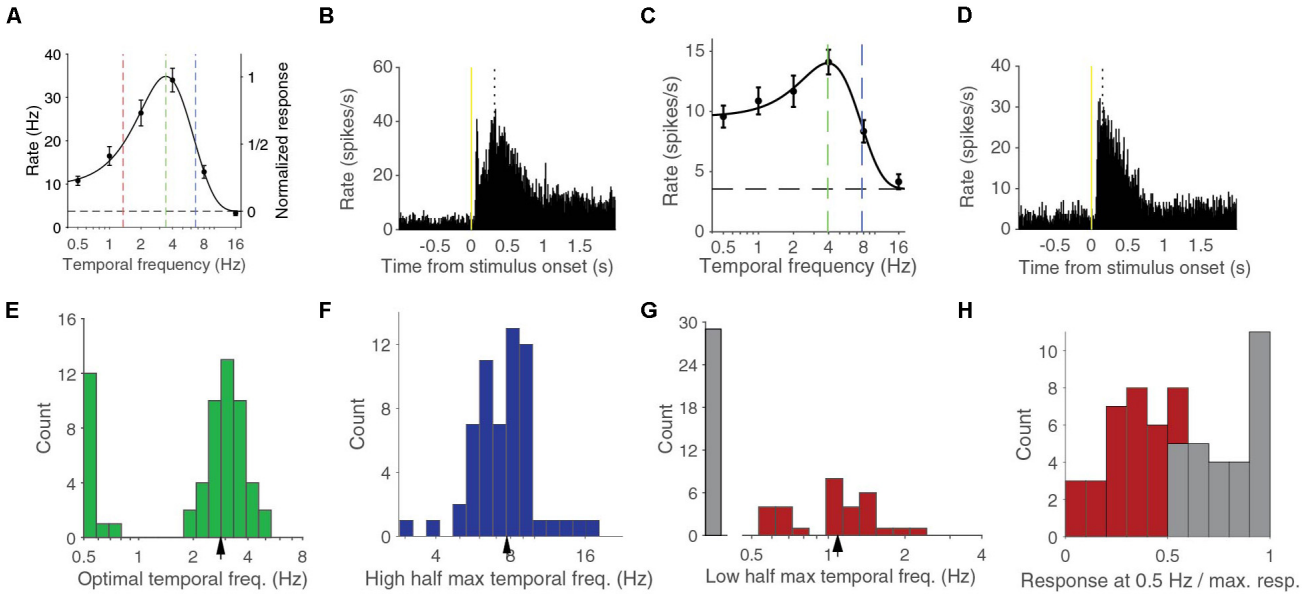


FIGURE 1 | Temporal frequency response at 90% contrast can be band pass or low pass. **(A)** Temporal frequency tuning of an example band pass neuron. Error bars indicate s.e.m. around the mean rate during the stimulus presentation. Black dashed line indicates the spontaneous firing rate. Curve is a difference-of-Gaussian (d.o.G.) fit to the data. Green dashed line indicates the optimal temporal frequency at which the neuron is predicted to give most response. Red line indicates the lower half-maximum response frequency. Blue line indicates the higher half-maximum response frequency. **(B)** Peristimulus time histogram (PSTH) of the activity for all stimuli of the same neuron shown in **(A)**. Dotted line indicates the peak response time. **(C)** Temporal frequency tuning of an example low pass neuron. **(D)** PSTH for the same neuron as **(C)**. **(E)** Histogram of optimal temporal frequencies from d.o.G. fits. Arrowhead indicates median. **(F)** Histogram of high half-maximum temporal frequencies. Arrowhead indicates median. **(G)** Histogram of low pass cells and the lower half-maximum temporal frequencies of band pass cells. Arrowhead indicates the median of the band pass cells. **(H)** Histogram of the response at 0.5 Hz over the response at the optimal temporal frequency. In gray the low-pass cells, in red the band-pass cells.

shape of the curve and the optimal temporal frequency and high half-maximum temporal frequencies of the population response were fairly constant above 30% contrast except for a gain change. Looking at the values of individually fitted temporal tuning curves, we noticed that there was significant difference in the optimal temporal frequency across contrasts (p = 0.00004, Kruskal–Wallis test, d.f. = 225, χ^2 = 25.6, 46 units in four mice; **Figure 2D**). A similar, but not significant, trend was also present in the high half-maximum temporal frequency across contrasts (p = 0.10, Kruskal–Wallis test, d.f. = 225, χ^2 = 7.7; **Figure 2D**). More specifically, the optimal temporal frequency at 10% was much lower than the optimal temporal frequency at 90% (median ± s.e.m. at 10% contrast: 1.23 ± 0.40 Hz; at 90% contrast: 2.90 ± 0.19 Hz; p = 0.00012 Wilcoxon signed-rank, z = 3.8, 46 units in four mice; **Figure 2E**). The high half-maximum temporal frequency was also significantly lower at 10% contrast than it was at 90% contrast (median ± s.e.m. at 10% contrast: 6.08 ± 0.99 Hz; at 90% contrast: 8.07 ± 0.27 Hz; p = 0.033 Wilcoxon signed-rank, z = 2.1). The response at 8 Hz relative to the maximum response increased with increasing contrast (p = 0.029, non-zero slope test, F = 4.9; **Figure 2F**). This did not happen at 0.5 Hz. If anything, there was a drop of relative response with increasing contrast and the slopes of the 0.5 and 8 Hz curve were significantly different (p = 0.0017, non-zero slope test on difference, F = 9.99; **Figure 2F**). The low half-maximum temporal frequency was smaller at low contrast (at 10% contrast: 0.89 ± 0.11 Hz; at 90% contrast: 1.14 ± 0.09 Hz), but this was just a trend (p = 0.10, Wilcoxon signed-rank, statistic = 79). At 10% contrast, there were only 16 band-pass neurons out of the 46 fitted by a d.o.G. (low-pass were 29 of 59 units at 90% contrast, 30 of 46 at 10% contrast, p = 0.10, chi-square test; **Figure 2G**). Overall, the temporal frequency tuning in V1 shifts toward higher temporal frequencies with increasing contrast.

Other than looking at how the temporal frequency tuning changed with contrast, we can also look at how the contrast response function changed with temporal frequency. The contrast dependence at a single temporal frequency can be fitted by a Naka-Rushton function (Albrecht and Hamilton, 1982). For higher contrasts, the Naka-Rushton curves have a decreasing steepness, referred to as saturation (**Figure 3A**). Occasionally, neurons were "super-saturated" and response decreased for the highest contrasts, but generally there was a very good fit (median explained variance was 98%). The C50 value, i.e., the interpolated contrast at which the cell responds at half the extrapolated 100% contrast response, is a common way to give an indication of the range of contrast where a cell is most sensitive. A high or low C50 indicates that the cell is most sensitive to, respectively, high or low contrasts. The population response curves showed differences in C50 across temporal frequencies (**Figure 3B**). The population contrast curves of **Figure 3B**, however, are more linear than the individual tuning curves and do not necessarily accurately reflect the changes in the individual neuronal contrast

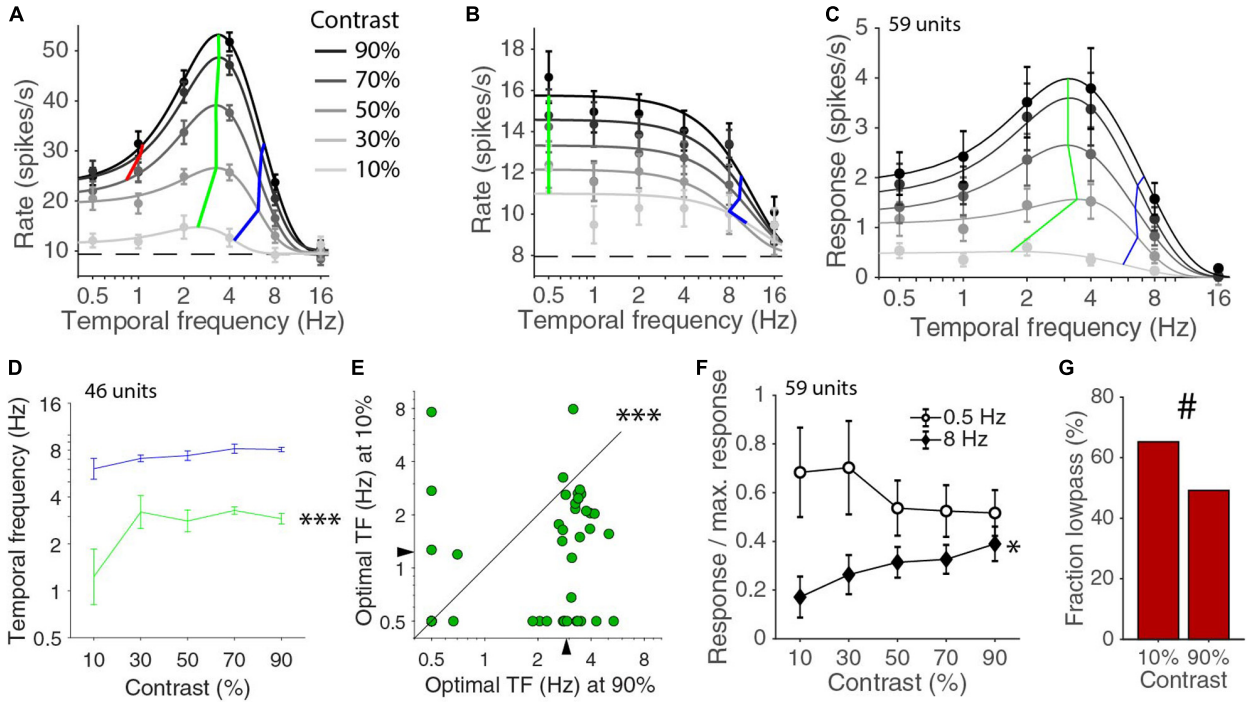


FIGURE 2 | Relationship of temporal frequency and contrast. **(A)** Temporal frequency tuning for different contrasts for an example band pass neuron. Error bars indicate s.e.m. around the mean rate during the stimulus presentation. Dashed line indicates the spontaneous rate. Black curves are d.o.G. fits, independently made for each contrast. Green lines connect the optimal temporal frequencies for each contrast. Blue lines connect high half-maximum responses. Red lines connect low half-maximum response when they exist. **(B)** Example like **(A)** for a low pass neuron. **(C)** Temporal frequency tuning for the mean population response for all units. Green and blue lines are connecting the optimal and high half-maximum responses for the fits of the population averages. **(D)** Median optimal (green) and high half-maximum temporal frequency (blue) for different contrasts for all individual units. Error bars are bootstrapped s.e.m. There is a difference across contrasts for optimal frequency (***p = 0.00004, Kruskal–Wallis test, 46 units in four mice), but only a trend for a change in the high half-maximum (p = 0.10, Kruskal–Wallis test). **(E)** Optimal temporal frequency is lower at 10% contrast than at 90% contrast (***p = 0.00012, Wilcoxon signed-rank). **(F)** Response at 0.5 and 8 Hz normalized by the maximum response at each contrast. The relative response at 8 Hz increases with contrast (*p = 0.029, non-zero slope test). **(G)** Fraction of low pass cells at 10 and 90% contrast (#, 29/59 at 90% contrast, 30/46 at 10%, p = 0.10, chi-square test).

response curves. The medians of the individual units C50, however, showed similar differences across temporal frequencies (p = 0.016, 59 units in four mice, Kruskal–Wallis test, d.f. = 294, χ^2 = 12.1, **Figure 3C**) and the C50 at 8 Hz is higher than at 0.5 Hz (p = 0.0057, Wilcoxon signed-rank, z = 2.8; **Figure 3D**). This again illustrates that response functions did not factorize in separate contrast and temporal frequency dependent functions. The C50 value alone does not completely describe the contrast response function. In particular, it does not capture whether the response changes over the full range of contrasts or only over a narrow range. For this reason, we also computed the dynamic range, i.e., the difference between the contrasts that evoke one quarter and three quarters of the maximum response (**Figure 3A**). A cell with a C50 of 50% contrast with a response that grows linearly with contrast has a dynamic range of 50%. A cell that has a very steep increase in response around the C50 contrast has a much lower dynamic range. The median dynamic range had a dependence on temporal frequency (p = 0.011, Kruskal–Wallis test, d.f. = 294, χ^2 = 13.0; **Figure 3E**), and peaked at 2 Hz. The distributions of the dynamic range at 0.5 and 8 Hz, however, were not different from each other (p = 0.51, Wilcoxon signed-rank, z = 0.66; **Figure 3F**).

Divisive Normalization

Our results thus clearly show that the mouse V1 population response is not a product of a function dependent on the temporal frequency and a function dependent on the stimulus contrast. The same is the case for the combination of spatial frequency and contrast (Heimel et al., 2010). The interdependence of response on contrast and spatial frequency was accurately described by divisive normalization. Divisive normalization is characterized by the normalization equation, describing the response R_i of neuron i by:

$$R_i = D_i^n \Big/ \left(\sigma^n + \sum_k D_k^n\right) \qquad (1)$$

where the enumerator D_i describes the driving input into the neuron and the denominator is a saturation constant σ plus the sum of a large number of driving inputs D_k, the normalization pool (**Figure 4A**; Heeger, 1992; Carandini and Heeger, 2011). The exponent n is a parameter signifying the rectification stage of the model. If we consider the population response $P = \sum_i R_i$, we find $P = \sum_i D_i^n \Big/ \left(\sigma^n + \sum_k D_k^n\right)$. We have established experimentally that the population response does not factorize, but let us assume

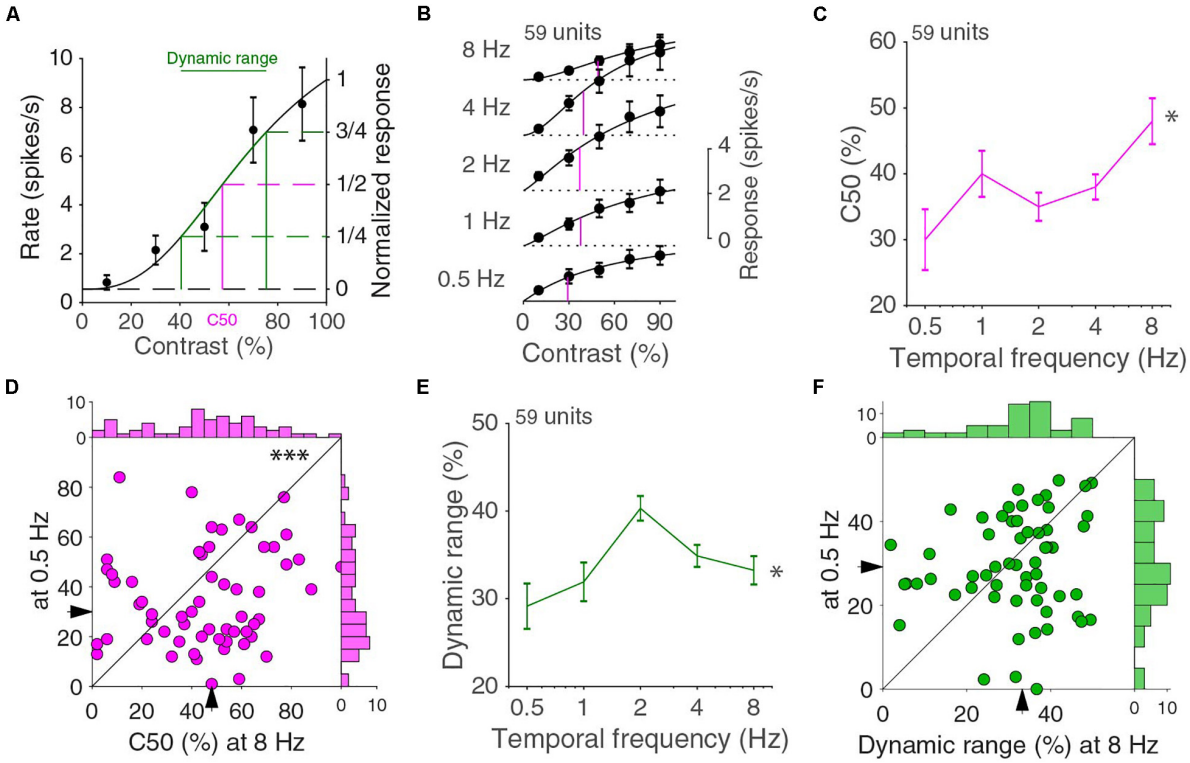


FIGURE 3 | Contrast response function for different temporal frequencies. **(A)** Contrast response function for an example neuron. Error bars indicate s.e.m. around the mean. The curve is a fit with a Naka-Rushton function. C50 is the contrast at which the fitted response is half the response at 100% contrast. The dynamic range is the difference in contrasts that evoke one and three quarters of the 100% contrast response **(B)**. Population averaged contrast responses across temporal frequencies. In magenta, the C50s of the Naka-Rushton fits of the population responses are indicated. Independent fits are made for all curves. **(C)** Median C50 varies across temporal frequencies (*p = 0.016, Kruskal–Wallis test, 59 units in four mice). Error bars indicate bootstrapped s.e.m. **(D)** C50 at 0.5 and 8 Hz (***p = 0.0057 Wilcoxon signed-rank). Histograms are shown on top and to the right. **(E)** The dynamic range differs across temporal frequencies (*p = 0.011, Kruskal–Wallis test). **(F)** Dynamic range at 0.5 and 8 Hz (p = 0.51, Wilcoxon signed-rank). Histograms are shown on top and to the right.

the driving inputs into V1 approximately do, i.e., there exist functions $d_i(f)$ such that $D_i = c\ d_i(f)^{1/n}$ with contrast c and temporal frequency f. We can then define $T(f) = \Sigma_i\ d_i(f)$, and find

$$P(c,f) = c^n / \left(\sigma^n / T(f) + c^n\right) \qquad (2)$$

This is a Naka-Rushton function, just as those that were used to fit the contrast responses of individual neurons, but the shape of the function depends on the temporal frequency. From Eq. (2) follows that there should be a single temporal frequency tuning function $T(f)$ and two parameters σ and n to fit the V1 population response at all contrasts and temporal frequencies. We find that indeed the 30 population responses (six temporal frequencies at five contrasts) can be well fitted by the normalization model with seven parameters (n, σ and the five d.o.G. parameters for T). An example fit (with T described by a difference-of-Gaussians) explaining 98% of the variance in the means is shown in **Figure 4B**, but there is a large range of parameter values with a similar goodness of fit. For all good fits, $\sigma^n/T(f)$ is much larger than 1 and n is close to 1. In those cases, the population response is approximately equal to $c\ T(f)/\sigma$. In fact, fitting the joint temporal frequency and contrast tuning with a function $c\ T(f)$ that is just linear in contrast, also explained 97% of the variance, and is thus an equally good fit with two parameters less (n, σ). The contrast-temporal frequency curves of individual neurons, however, are much more poorly fit by fits that are linear in contrast, than by the normalizing model with n and σ optimized for each unit (median explained variance of linear model: 83%, normalization model: 92%; $p = 10^{-11}$ Wilcoxon, $z = -6.8$; **Figures 4C,D**). The normalization model thus provides a much better, but still not perfect fit. If we look at the population fits in **Figure 4B** more closely, we see that they are relatively poor at low temporal frequencies. The normalization model explained only 92 and 94% of the variance at 0.5 and 1 Hz, while it explained 99, 99, and 97% at 2, 4, and 8 Hz. The fits undershoot for the lowest contrasts, and overshoot for the highest contrasts. Furthermore, while the data showed that the optimal temporal frequency shifts with contrast, **Figure 4B** shows that the optimal temporal frequency does not change with contrast. Indeed, taking the derivative with respect to f of the population response in Eq. (2), we find

$$\frac{dP(c,f)}{df} = \frac{\sigma^n c^n}{\left[\sigma^n + c^n T(f)\right]^2} \frac{dT(f)}{df}.$$

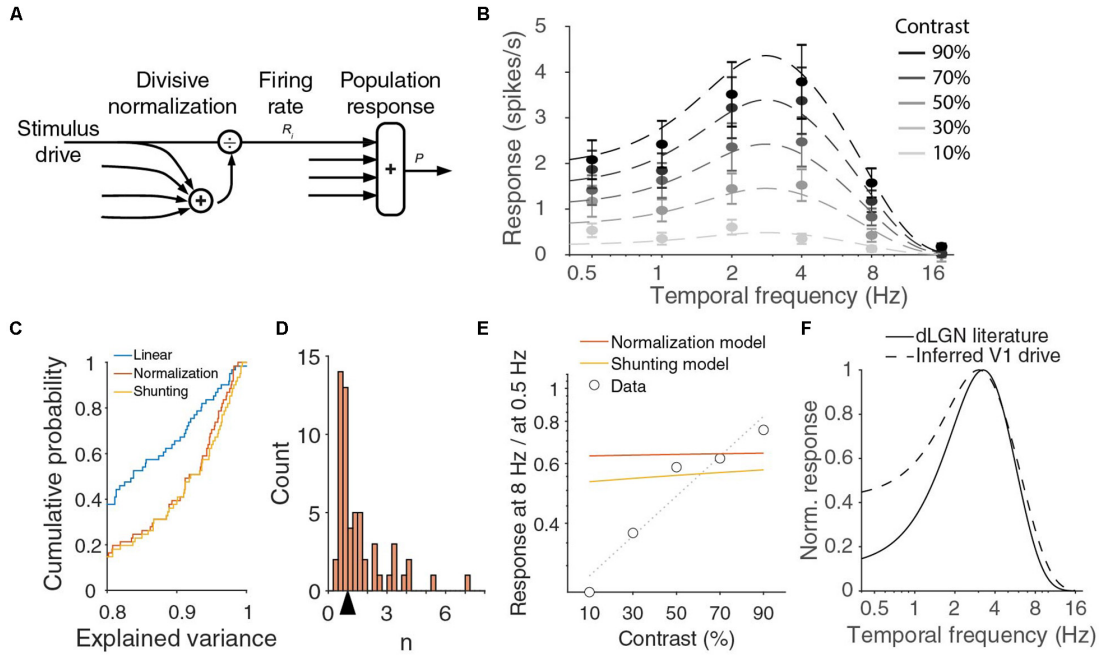


FIGURE 4 | V1 output can be fit with normalization model, but inferred drive is lower pass than dLGN output. **(A)** In the normalization model, the response of a neuron is described by a stimulus drive divided by the sum of a pool of such inputs. **(B)** Population responses across temporal frequencies and contrasts can be fit with divisive normalization of d.o.G. temporal frequency tuning of the input. Red curve is a fit with R_m = 37 spikes/s, σ = 6.9, n = 0.87, R_e = 1 spikes/s, w_e = 4.7 Hz, R_i = 0.66 spikes/s, w_i = 1.5 Hz. Error bars denote s.e.m. around mean. **(C)** Cumulative histogram of the explained variances when all individual units are fit with a linear model, the normalization model, or the shunting-extension of the normalization model. **(D)** The histogram of n for fitted to all individual units. The arrow indicates the median 0.99. **(E)** The ratio of the response at 8 Hz over the response at 0.5 Hz strongly varies with contrast in the data, but this is not true in the optimal fit of the normalization model or for the shunting model. Dotted line is linear fit to the data. **(F)** Full contrast dLGN output from literature (Tang et al., 2016) (solid line) does not match the stimulus drive from the normalization equation (dashed line).

Therefore, the maximum of the population response P will be at the maximum of $T(f)$ at $f = f_{\text{opt}}$ where $dT(f)/df = 0$, independently of the contrast. Furthermore, taking the derivative of $P(c,f)/P(c,f_{\text{opt}})$ with respect to contrast, we find

$$\frac{d}{dc}\frac{P(c,f)}{P(c,f_{opt})} = \frac{nc^{n-1}\sigma^n\left\{T\left(f_{opt}\right) - T(f)\right\}}{T\left(f_{opt}\right)T(f)\left\{\frac{\sigma^n}{T(f)} + c^n\right\}^2} > 0,$$

because $T(f_{\text{opt}}) > T(f)$. This means that the response at a suboptimal temporal frequency will grow relatively faster with contrast than the response at the optimal frequency f_{opt}. Thus the responses at 0.5 and 8 Hz, divided by the maximum response, should grow with contrast. In **Figure 2F**, we presented that this was true at 8 Hz, but not at 0.5 Hz. Indeed, if we plot the ratio of the response at 8 Hz and at 0.5 Hz, we see that the experimental data show a strong increase with contrast, while the optimal fit of the normalization model predicts the ratio to be almost independent of the stimulus contrast (**Figure 4E**).

One assumption implicitly made in the normalization Eq. (1) is that there is a single pair of saturation constant σ and exponent n for different stimuli. Any neural implementation of normalization, however, will be dynamic and there could be an interaction of the stimulus dynamics and normalization. Indeed, one suggested implementation of divisive normalization by shunting inhibition (Carandini and Heeger, 1994; Carandini et al., 1997) predicts that σ grows with the temporal frequency of the stimulus. This would allow $P(c,f)/P(c,f_{\text{opt}})$ to decrease with contrast for small frequencies (again can be shown by inserting Eq. (2) and taking the derivative), and potentially fit the decreasing relative response at 0.5 Hz (**Figure 2F**). However, if we substitute $\sigma + \tau f$ for σ, and optimize the fit for all data, we find essentially the same fit as for the model with fixed σ and do not get a better fit for the ratio of the responses to 8 and 0.5 Hz (**Figure 4E**, shunting model). If we reconsider Eq. (2), however, we see that any frequency-dependency of σ can be absorbed by choosing a different $T(f)$. This explains why we get almost the same fit. The reason for the small difference between the models in **Figure 4E** is because we constrained $T(f)$ to be a difference-of-Gaussians. If we release this requirement, the shunting model and the regular model give exactly the same fits, with an explained variance of 98%. Fitting individual units with this shunting model slightly improves the fits, as it has an extra degree of freedom and includes the static normalization equation ($p = 8 \times 10^{-6}$, Wilcoxon, $z = -4.4$; **Figure 4C**), but the median explained variance is lower for the extended model when we adjust it for the extra model parameter (median adjusted explained variance normalization model: 0.89, shunting model: 0.88).

The normalization model predicts the temporal frequency tuning of the driving input to V1 to be equal to $T(f)^{1/n}$ (**Figure 4F**). It is difficult to measure the driving input into

V1, but we expect it to be close to the signal leaving the dLGN. The temporal frequency dependence of the mouse dLGN has been measured with full contrast stimuli (Grubb and Thompson, 2003; Tang et al., 2016). From the reported optimal temporal frequencies (median 3.2 Hz) and low and high half-maximum frequencies (medians of 1.4 and 6.0 Hz, respectively), we can infer the dLGN population response (**Figure 4F**). There is a large difference between the prediction from the normalization model and the measured dLGN response at the lower temporal frequencies.

We conclude that our findings are consistent with divisive normalization operating at the higher temporal frequencies, but that either divisive normalization is not operating as described by Eq. (1) at the lower temporal frequencies, or that the dependencies on contrast and temporal frequency in the driving inputs into V1 do not factorize.

DISCUSSION

We found that the shape of contrast tuning curves of mouse V1 neurons depended on the temporal frequency of the stimulus and vice versa, the shape of the temporal frequency tuning curve depended on the stimulus contrast. We measured this with drifting gratings, as commonly done (Foster et al., 1985; Priebe et al., 2006; Gao et al., 2010). The gratings had a fixed spatial frequency, and changing temporal frequency was equivalent to changing the drifting speed. It has been established, however, that in V1 dependencies on spatial and temporal frequency are mostly separable (Tolhurst and Movshon, 1975), and rather than looking at speed tuning, it is natural to focus on spatial and temporal frequency tuning.

Our findings at 90% contrast are very comparable to previous measurements of temporal frequency tuning at 80–100% contrast in mouse V1 (Niell and Stryker, 2008; Gao et al., 2010; Durand et al., 2016). In close agreement to our finding that 49% of the cells were low-pass tuned, Gao et al. (2010) found that 44% were low-pass. The two early studies found a peak response for temporal frequencies between 1 and 2 Hz. This is slightly below our median peak frequency (2.83 Hz), which exactly matches the 2.8 Hz found by Durand et al. (2016). We cannot give a reason for this difference. Perhaps there was a difference in the depth of anesthesia across the studies, although there was no significant difference in peak frequency between the awake and anesthetized condition (2.8 vs 2.99 Hz; Durand et al., 2016). Niell and Stryker (2008) used counter-phase changing gratings, different from our drifting gratings, but Gao et al. (2010) used drifting gratings. At low luminance, temporal frequency tuning is strongly dependent on luminance (Ferry-Porter law; Bex and Langley, 2007; Umino et al., 2008), but the previous studies and ours were done at similar, photopic, light levels, where tuning is independent of luminance. Our median optimal temporal frequency is below the 4 Hz found for rats (Girman et al., 1999), but in line with other nocturnal and crepuscular animals that are strongly dependent on rod vision (Heimel et al., 2005).

The peak frequency in V1 is below the peak frequency in the mouse dLGN (3–4 Hz; Grubb and Thompson, 2003; Durand et al., 2016; Tang et al., 2016). Furthermore, while half of the cells in mouse V1 are low-pass for temporal frequency, the cells in the dLGN are consistently band-pass (Grubb and Thompson, 2003). The temporal frequency tuning thus substantially moves toward lower frequencies in the mouse like it does in the monkey (Hawken et al., 1996). This shift has been hypothesized to be caused by a combination of intracortical inhibition and thalamocortical NMDA receptors (Krukowski and Miller, 2001).

We found that the temporal frequency tuning characteristics were dependent on contrast, in particular at low contrast. Responses at high temporal frequencies were relatively more reduced at lower contrasts than responses at more optimal temporal frequencies (**Figure 2F**). Relative response at low temporal frequencies, on the other hand, were not more reduced at lower contrasts. Reducing contrast also slightly lowered the preferred temporal frequency (**Figure 2D**). In carnivores, the responses at low and high temporal frequencies also grew more with increasing contrast than responses at the optimal temporal frequency (Holub and Morton-Gibson, 1981). In particular, the responses at high temporal frequencies grow faster at higher contrasts, and thus also leading to a shift toward a higher temporal frequency preference with increasing contrast (Albrecht, 1995; Alitto and Usrey, 2004). This influence of contrast on temporal frequency tuning is different from its influence on spatial frequency tuning. In the macaque, the optimal spatial frequency is independent of contrast and both the low and high half-maximum frequencies move away from the optimal frequency with increasing contrast (Sceniak et al., 2002). In the mouse, the contrast dependence has not been measured down to the typical low half-maximum frequency (which requires very large stimuli due to the low visual acuity of mice), but the response to high spatial frequency gratings was also much more sensitive to contrast than the response to gratings of the optimal spatial frequency (Heimel et al., 2010). The effect of contrast on spatial frequency tuning follows from the divisive normalization behavior of visual cortex (Heimel et al., 2010). According to the normalization model, the response of a neuron is well described by dividing its driving input by a constant and the sum of the activity of a normalizing pool of inputs (Heeger, 1992; Carandini and Heeger, 2011). For spatial frequencies that evoke much response, the activity level of the normalizing pool is higher and the driving input is divided by a larger number than for spatial frequencies that evoke little response. The effect is thus a widening of the tuning curve with increasing contrast. At high temporal frequencies, the effect of contrast on the population response is as predicted by the normalization model (see **Figure 2F**). Higher temporal frequencies cause less activity in the normalization pool than more optimal temporary frequencies, resulting in less normalization and thus more dependence on contrast. At low temporal frequencies, the stationary normalization Eq. (1) predicts the same effect, but the data do not show this (**Figure 2F**). One of the assumptions implicitly made in Eq. (1) is that the saturation constant σ and the exponent n are independent of the stimulus. This assumption holds across orientations and spatial frequencies (Heeger, 1992). At the introduction of the normalization model, the validity of this

assumption was not asserted for different temporal frequencies, but it was noted that the model could explain the contrast-dependent changes in temporal frequency tuning that were found in the cat (Holub and Morton-Gibson, 1981; Heeger, 1992; Albrecht, 1995). The stationary normalization equation, however, does not accurately describe the responses across temporal frequencies and contrasts with a single set of σ and n (Carandini et al., 1997). It could still be, however, that a single divisive normalization mechanism is operating, but that the interaction of a changing stimulus and the normalization mechanism leads to a different set of parameters in the equation describing the stationary state. Currently, we do not know how divisive normalization is implemented in the visual cortex. A number of mechanisms, such as (shunting) inhibition (Carandini and Heeger, 1994), excitation (Sato et al., 2016) and short-term synaptic depression (Carandini et al., 2002) have been proposed to underlie the normalization phenomenon in visual cortex, but there is not one mechanism that is consistent with all effects. Carandini et al. (1997) have shown how shunting inhibition could lead to divisive normalization with a saturation constant σ that grows with increasing stimulus frequency. The resulting stationary normalization equation, however, produces the same fit to the data as the model without this extra degree of freedom, if we also optimize the frequency-dependency of the driving input.

Even without knowing the mechanism underlying normalization, looking at normalization as a dynamic process offers a possible explanation for the relatively poor fit of the data at lower temporal frequencies. Consider that the normalization mechanism is operating on a time scale that is faster than that of the low temporal frequencies (0.5 Hz) that we used. The activity of many neurons in V1 is modulated by the stimulating temporal frequency. A simple cell would be very responsive during one period of the stimulus cycle. Perhaps the effects of normalization diminish quickly during the responsive period of a cell. For low temporal frequencies, normalization would thus have no effect during most of the responsive period of the cell and would not change the average firing rate as much as it would do for higher temporal frequencies.

One more assumption that was implicitly made in deriving the stationary normalization Eq. (1) was that the activity in the normalization pool is equal to the average of the population activity. Of course, this can only be an approximation, because the neurons certainly do not have instantaneous access to the population activity across the entire visual field (Reynaud et al., 2012). In our case, however, we have used full screen gratings and recorded from neurons that had their receptive field away from the screen borders. The stimulus input would thus at least have been relatively homogeneous for nearby neurons which have similar receptive fields. The normalization pool, however, could have a different temporal frequency tuning than the population activity. This extra freedom in the fit of the model would certainly produce a more accurate fit to the data. It could also be that the normalization pool has a local polarity (dark/light) preference, and therefore oscillate with the stimulus frequency. In this case, it also becomes necessary to estimate the time averaging by the normalization mechanism. To fit human steady state visual evoked potentials of masking stimuli with the normalization model, a temporal averaging window of 26 ms provided the best fit (Tsai et al., 2012). If one, however, allows the normalization pool and the population pool to vary independently for different stimulus parameters, the predictive power of the normalization model disappears.

An entirely different explanation is that normalization does operate also at lower temporal frequencies, but that the assumption that the contrast and temporal frequency dependencies of the input driving V1 factorize does not hold. We do not know the precise input into V1, but we do know that indeed in the macaque this assumption fails. Changes in temporal frequency altered the contrast tuning in dLGN (Derrington and Lennie, 1984; Dhruv et al., 2009). In the ferret, the difference in contrast gain at low and high temporal frequency was not higher in V1 than the difference already present in the LGN (Alitto and Usrey, 2004). Furthermore, even in the retina (of the cat) contrast and temporal frequency do not completely separate for low temporal frequencies (Shapley and Victor, 1978). Responses of X and Y retinal ganglion cells to low frequencies of modulation (<1 Hz) grew less than proportionally with contrast. Response amplitudes at higher modulation frequencies scaled approximately proportionally with contrast. The source of the interdependency of temporal frequency and contrast in mouse V1 responses may thus already lie in the LGN or the retina.

This explanation why the normalization model poorly fits the data, leaves open the possibility that divisive normalization is operating in the V1 exactly as predicted. However, along the visual hierarchy in human cortex, the temporal frequency tuning becomes progressively more low-pass (Mullen et al., 2010). This would not follow from a normalization model working the same at all temporal frequencies. As discussed previously, dynamic implementations of divisive normalization may lead to frequency dependence of the saturation constant (Carandini et al., 1997), but there could also be other mechanisms operating in the visual system to change temporal frequency tuning to lower frequencies. This may be the combination of intracortical inhibition and NMDA receptor signaling hypothesized to be responsible for the change in temporal frequency from dLGN to V1 (Krukowski and Miller, 2001). It will be interesting to understand if this would correctly predict our data on the interplay between temporal frequency and contrast in the responses. Furthermore, it may give a mechanism underlying the Thompson effect that contrast and perceived speed and flicker are not completely separated (Thompson, 1982; Thompson and Stone, 1997).

More than 50 years after the first psychophysical measurements of the dependence of contrast sensitivity on temporal frequency (Robson, 1966), we find that we still do not know how the limits set by the retina are changed into the limits of our perception. Our measurements could help to select between candidate neural implementations of the normalization model linking visual input to perception.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors on request, without undue reservation.

ETHICS STATEMENT

The animal study was reviewed and approved by DEC and IVD KNAW.

AUTHOR CONTRIBUTIONS

DC and MA performed the experiments. DC, MA, and JH devised the experiments. JH wrote the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by NWO VIDI Grant 864.10.010.

ACKNOWLEDGMENTS

We thank Christiaan Levelt and Pieter Roelfsema for sharing equipment, Chris Klink for critical reading of the manuscript, and Matteo Carandini for helpful suggestions.

REFERENCES

Ahmadlou, M., Zweifel, L. S., and Heimel, J. A. (2018). Functional modulation of primary visual cortex by the superior colliculus in the mouse. *Nat. Commun.* 9:3895.

Albrecht, D. G. (1995). Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions. *Vis. Neurosci.* 12, 1191–1210. doi: 10.1017/s0952523800006817

Albrecht, D. G., and Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function. *J. Neurophysiol.* 48, 217–237. doi: 10.1152/jn.1982.48.1.217

Albrecht, D. G., and Geisler, W. S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vis. Neurosci.* 7, 531–546. doi: 10.1017/s0952523800010336

Alitto, H. J., and Usrey, W. M. (2004). Influence of contrast on orientation and temporal frequency tuning in ferret primary visual cortex. *J. Neurophysiol.* 91, 2797–2808. doi: 10.1152/jn.00943.2003

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., and Malach, R. (2002). Contrast sensitivity in human visual areas and its relationship to object recognition. *J Neurophysiol.* 87, 3102–3116. doi: 10.1152/jn.2002.87.6.3102

Bex, P. J., and Langley, K. (2007). The perception of suprathreshold contrast and fast adaptive filtering. *J. Vis.* 7, 1–23. doi: 10.1167/7.12.1

Burr, D. C., and Ross, J. (1982). Contrast sensitivity at high velocities. *Vision Res.* 22, 479–484. doi: 10.1016/0042-6989(82)90196-1

Camillo, D., Ahmadlou, M., Saiepour, M. H., Yasaminshirazi, M., Levelt, C. N., and Heimel, J. A. (2018). Visual processing by calretinin expressing inhibitory neurons in mouse primary visual cortex. *Sci. Rep.* 8:12355.

Carandini, M., and Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science* 264, 1333–1336. doi: 10.1126/science.8191289

Carandini, M., Heeger, D. J., and Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* 17, 8621–8644. doi: 10.1523/jneurosci.17-21-08621.1997

Carandini, M., and Heeger, D. J. (2011). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13, 51–62. doi: 10.1038/nrn3136

Carandini, M., Heeger, D. J., and Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *J. Neurosci.* 22, 10053–10065. doi: 10.1523/jneurosci.22-22-10053.2002

Derrington, A. M., and Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J. Physiol.* 357, 219–240. doi: 10.1113/jphysiol.1984.sp015498

Dhruv, N. T., Tailby, C., Sokol, S. H., Majaj, N. J., and Lennie, P. (2009). Nonlinear signal summation in magnocellular neurons of the macaque lateral geniculate nucleus. *J. Neurophysiol.* 102, 1921–1929. doi: 10.1152/jn.00331.2009

Durand, S., Iyer, R., Mizuseki, K., de Vries, S., Mihalas, S., and Reid, R. C. (2016). A comparison of visual response properties in the lateral geniculate nucleus and primary visual cortex of awake and anesthetized mice. *J. Neurosci.* 36, 12144–12156. doi: 10.1523/jneurosci.1741-16.2016

Foster, K. H., Gaska, J. P., Nagler, M., and Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *J. Physiol.* 365, 331–363. doi: 10.1113/jphysiol.1985.sp015776

Gao, E., DeAngelis, G. C., and Burkhalter, A. (2010). Parallel input channels to mouse primary visual cortex. *J. Neurosci.* 30, 5912–5926. doi: 10.1523/jneurosci.6456-09.2010

Girman, S. V., Sauvé, Y., and Lund, R. D. (1999). Receptive field properties of single neurons in rat primary visual cortex. *J. Neurophysiol.* 82, 301–311. doi: 10.1152/jn.1999.82.1.301

Grubb, M. S., and Thompson, I. D. (2003). Quantitative characterization of visual response properties in the mouse dorsal lateral geniculate nucleus. *J. Neurophysiol.* 90, 3594–3607. doi: 10.1152/jn.00699.2003

Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H., and Buzsáki, G. (2000). Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J. Neurophysiol.* 84, 401–414. doi: 10.1152/jn.2000.84.1.401

Hawken, M. J., Shapley, R. M., and Grosof, D. H. (1996). Temporal-frequency selectivity in monkey visual cortex. *Vis. Neurosci.* 13, 477–492. doi: 10.1017/s0952523800008154

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* 9, 181–197. doi: 10.1017/s0952523800009640

Heimel, J. A., Van Hooser, S. D., and Nelson, S. B. (2005). Laminar organization of response properties in primary visual cortex of the gray squirrel (*Sciurus carolinensis*). *J. Neurophysiol.* 94, 3538–3554. doi: 10.1152/jn.00106.2005

Heimel, J. A., Saiepour, M. H., Chakravarthy, S., Hermans, J. M., and Levelt, C. N. (2010). Contrast gain control and cortical TrkB signaling shape visual acuity. *Nat. Neurosci.* 13, 642–648. doi: 10.1038/nn.2534

Holub, R. A., and Morton-Gibson, M. (1981). Response of visual cortical neurons of the cat to moving sinusoidal gratings: response-contrast functions and spatiotemporal interactions. *J. Neurophysiol.* 46, 1244–1259. doi: 10.1152/jn.1981.46.6.1244

Hubel, D. H., and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J Physiol.* 148, 574–591. doi: 10.1113/jphysiol.1959.sp006308

Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in psychtoolbox-3?. *Perception* 36, 1–16.

Krekelberg, B., van Wezel, R. J., and Albright, T. D. (2006). Interactions between speed and contrast tuning in the middle temporal area: implications for the neural code for speed. *J. Neurosci.* 26, 8988–8998. doi: 10.1523/jneurosci.1983-06.2006

Krukowski, A. E., and Miller, K. D. (2001). Thalamocortical NMDA conductances and intracortical inhibition can explain cortical temporal tuning. *Nat. Neurosci.* 4, 424–430. doi: 10.1038/86084

Mullen, K. T., Thompson, B., and Hess, R. F. (2010). Responses of the human visual cortex and LGN to achromatic and chromatic temporal modulations: an fMRI study. *J. Vis.* 10:13. doi: 10.1167/10.13.13

Niell, C. M., and Stryker, M. P. (2008). Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* 28, 7520–7536. doi: 10.1523/jneurosci.0623-08.2008

Pawar, A. S., Gepshtein, S., Savel'ev, S., and Albright, T. D. (2019). Mechanisms of spatiotemporal selectivity in cortical area MT. *Neuron* 101, 1–14.

Priebe, N. J., Lisberger, S. G., and Movshon, J. A. (2006). Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *J. Neurosci.* 26, 2941–2950. doi: 10.1523/jneurosci.3936-05.2006

Prusky, G. T., West, P. W., and Douglas, R. M. (2000). Behavioral assessment of visual acuity in mice and rats. *Vision Res.* 40, 2201–2209. doi: 10.1016/s0042-6989(00)00081-x

Reynaud, A., Masson, G. S., and Chavane, F. (2012). Dynamics of local input normalization result from balanced short- and long-range intracortical interactions in area V1. *J. Neurosci.* 32, 12558–12569. doi: 10.1523/jneurosci.1618-12.2012

Robson, J. G. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. *J. Opt. Soc. Am.* 56, 1141–1142. doi: 10.1364/josa.56.001141

Sato, T. K., Haider, B., Häusser, M., and Carandini, M. (2016). An excitatory basis for divisive normalization in visual cortex. *Nat. Neurosci.* 19, 568–570. doi: 10.1038/nn.4249

Sceniak, M. P., Hawken, M. J., and Shapley, R. (2002). Contrast-dependent changes in spatial frequency tuning of macaque V1 neurons: effects of a changing receptive field size. *J. Neurophysiol.* 88, 1363–1373. doi: 10.1152/jn.2002.88.3.1363

Shapley, R. M., and Victor, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *J. Physiol.* 285, 275–298. doi: 10.1113/jphysiol.1978.sp012571

Skottun, B. C., Bradley, A., and Ramoa, A. S. (1986). Effect of contrast on spatial frequency tuning of neurones in area 17 of cat's visual cortex. *Exp. Brain Res.* 63, 431–435.

Tang, J., Ardila Jimenez, S. C., Chakraborty, S., and Schultz, S. R. (2016). Visual receptive field properties of neurons in the mouse lateral geniculate nucleus. *PLoS One* 11:e0146017. doi: 10.1371/journal.pone.0146017

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Res.* 22, 377–380. doi: 10.1016/0042-6989(82)90153-5

Thompson, P., and Stone, L. S. (1997). Contrast affects flicker and speed perception differently. *Vision Res.* 37, 1255–1260. doi: 10.1016/s0042-6989(96)00302-1

Tolhurst, D. J., and Movshon, J. A. (1975). Spatial and temporal contrast sensitivity of striate cortical neurones. *Nature* 257, 674–675. doi: 10.1038/257674a0

Tsai, J. J., Wade, A. R., and Norcia, A. M. (2012). Dynamics of normalization underlying masking in human visual cortex. *J. Neurosci.* 32, 2783–2789.

Umino, Y., Solessio, E., and Barlow, R. B. (2008). Speed, spatial, and temporal tuning of rod and cone vision in mouse. *J. Neurosci.* 28, 189–198. doi: 10.1523/jneurosci.3551-07.2008

Umino, Y., Pasquale, R., and Solessio, E. (2018). Visual temporal contrast sensitivity in the behaving mouse shares fundamental properties with human psychophysics. *eNeuro* 5:e181-18.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Medicine

ORIGINAL RESEARCH
published: 26 August 2020
doi: 10.3389/fmed.2020.00469

Can Psychophysics Be Fun? Exploring the Feasibility of a Gamified Contrast Sensitivity Function Measure in Amblyopic Children Aged 4–9 Years

*Doaa Elfadaly[1,2], Sahar Torky Abdelrazik[2], Peter B. M. Thomas[1,3], Tessa M. Dekker[4], Annegret Dahlmann-Noor[1,3] and Pete R. Jones[3,4,5]**

[1] Moorfields Eye Hospital NHS Foundation Trust, London, United Kingdom, [2] Department of Ophthalmology, Faculty of Medicine, Minia University, Minia, Egypt, [3] NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom, [4] Child Vision Laboratory, Institute of Ophthalmology, University College London (UCL), London, United Kingdom, [5] Division of Optometry and Visual Sciences, City, University of London, London, United Kingdom

OPEN ACCESS

Edited by:
Alexandre Reynaud, McGill University, Canada

Reviewed by:
Tamara S. Oechslin, University of Alabama at Birmingham, United States
Anna Kosovicheva, Northeastern University, United States

***Correspondence:**
Pete R. Jones
p.r.jones@ucl.ac.uk

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: *13 May 2020*
Accepted: *13 July 2020*
Published: *26 August 2020*

Citation:
Elfadaly D, Abdelrazik ST, Thomas PBM, Dekker TM, Dahlmann-Noor A and Jones PR (2020) Can Psychophysics Be Fun? Exploring the Feasibility of a Gamified Contrast Sensitivity Function Measure in Amblyopic Children Aged 4–9 Years. Front. Med. 7:469. doi: 10.3389/fmed.2020.00469

Routine assessments of the Contrast Sensitivity Function [CSF] could be useful for the diagnosis and monitoring of amblyopia. However, current CSF measures are not clinically practical, as they are too slow, too boring, and too uncomfortable to sustain a young child's interest. Here we assess the feasibility of a more gamified approach to CSF testing, in which a maximum likelihood psychophysical algorithm (QUEST+) is combined with a largely unconstrained user interface (no fixation target, head restraints, or discrete trials). Twenty-five amblyopes (strabismic, anisometropic, or mixed) aged 4.0–9.2 years performed the gamified CSF assessment monocularly (once per eye). The test required the child to "pop" (press) grating stimuli as they "bounced" around a tablet screen. Head tracking via the tablet's front-facing camera was used to adjust for variations in viewing distance *post hoc*. CSFs were fitted for each eye, and Area Under the CSF (AUCSF) computed as a summary measure of sensitivity. The results showed that AUCSF measurements were able to separate moderately and severely amblyopic eyes from fellow eyes (case-control effect), and to distinguish individuals with varying degrees of vision loss (dose effect). Even the youngest children exhibited no difficulties completing the test or comprehending what to do, and most children appeared to find the test genuinely enjoyable. Informal feedback from a focus group of older children was also positive, although potential shortcomings with the present design were identified. This feasibility study indicates that gamified, child-friendly vision assessments have promise as a future means of pediatric clinical assessment. Such measures could be particularly valuable for assessing children outside of conventional eye-care facilities (e.g., home-monitoring, school screening).

Keywords: amblyopia, psychophysics, children, contrast sensitivity function, QUEST+, quick CSF, OpenFace, gabor

INTRODUCTION

Precise measures of spatial vision (acuity, contrast sensitivity) are important for the diagnosis and monitoring of amblyopia. Ideally, the entire spatial contrast sensitivity function (1) [CSF] should be measured, since sensitivity to low spatial frequency information can be affected in amblyopia, independent of acuity (2–4). Unfortunately, there is at present no effective clinical solution for the routine assessment of CSFs in young children.

One key difficulty is that conventional CSF assessments are too slow to be performed routinely in young children. Thus, while letter charts [e.g., Pelli Robson charts (5)] can provide a rapid summary measure of overall contrast sensitivity, to measure contrast detection thresholds precisely, and to do so across multiple, specific spatial frequencies, typically requires a protracted psychophysical procedure composed of several hundreds of trials (10 min) (6).

Recently, the problem of long test durations has been mitigated by the development of more efficient psychophysical algorithms, such as the "quick CSF" (qCSF) (7–12) or QUEST+ (13, 14). These "maximum likelihood" (ML) algorithms evaluate all previous trials, along with all possible outcomes to any subsequent stimulus, in order to determine the most informative stimulus to present next. This makes ML assessments faster than conventional psychophysical procedures [e.g., adaptive staircases (15, 16)], allowing the whole CSF to be measured in around 30–100 trials (3–10 mins) (6, 8, 10). Furthermore, in situations such as home monitoring, where the same individual undergoes repeated testing, data from any previous assessments can be entered as "prior information," further reducing any subsequent test durations.

However, while ML assessments are faster than conventional psychophysical methods, current implementations remain inappropriate for the routine assessment of young children. They are often uncomfortable due to the use of chin rests and fixation targets. Furthermore, children often perceive the tests as boring: consisting as they do of a protracted sequence of monotonous, independent "trials" — each following an identical and highly regimented format (e.g., "was the target on the left, or the right?").

Most adults and older children are willing to tolerate a degree of boredom or discomfort. In younger children, however, such "human factors" can lead to a loss of motivation, often resulting in visual function being critically underestimated (17, 18). Furthermore, even with close monitoring and constant encouragement, it is not uncommon for psychophysical procedures to have to be abandoned in young children. And while high attrition rates, aberrant data points, and extensive supervision can sometimes be accommodated in scientific research, none is sustainable clinically.

In principle, CSF assessments could be made more fun and engaging by adopting a "gamified" approach to vision testing. For example, in the present study we invited children to "pop" bubbles (Gabor patches) by pressing them as they "bounced" around a tablet screen: a task that even very young children tended to find intuitive and engaging (and a response mechanism that forms the basis of many commercial, tablet-based games, targeted at young children). Other research groups are also exploring similar "gamification" strategies. For example, Bosten et al. (19) describe a tablet-based test to screen for color vision deficiency (CVD) in preliterate children (2–6 years), in which the child reveals characters by correctly selecting colored targets. While, Hosokawa et al. (20) have developed a battery of game-like tests designed to probe various visual functions in a hospital waiting area (contrast sensitivity, visual fields, crowding, multiple object tracking).

It is important to recognize, however, that gamification carries a potential cost in terms of empirical rigor. Thus, in designing the test the way we did, we knowingly introduced many potential confounds into our CSF measure, including: criterion effects (i.e., the level of confidence observers felt was necessary before responding); viewing strategies (e.g., whether the user actively scanned the scene or fixated passively in one region); perceptual crowding effects between the multiple targets [i.e., targets shown close together can be harder to detect (21)]; response error (e.g., children pressing too slowly, or in the wrong location); and various stimulus artifacts (e.g., due to screen non-uniformities, variations in viewing angle, or smudges on the screen from repeated pressing).

All of these factors are potential sources of measurement error. Whether they actually prevent the collection of useful data is, however, an empirical question. It may be that a less constrained, gamified procedure is simply unable to produce sensible CSF measurements. At the other extreme, it may be that conventional psychophysical assessments are limited almost entirely by the child's interest and concentration, in which case a gamified test may produce *more* accurate and reliable data than a conventional assessment. Finally, the truth may lie somewhere in between these two extremes – with a gamified test providing a tolerable loss of accuracy, that may be offset by more practical benefits (e.g., higher completion rates or ease-of-use).

To begin to explore the feasibility of a gamified CSF assessment we invited 25 young children (4.0–9.2 years) with diagnosed amblyopia to complete a novel assay that we informally dubbed the "pop CSF" (pCSF) test (see *Methods* for test details). We also asked an advisory group of children and young people with lived experience of various eye and vision conditions to give informal feedback on the test, and to consider its potential pros and cons.

The goal, at this preliminary stage, was not to formally validate a new test or medical device, but to investigate the potential merit of a more gamified approach to psychophysics. In particular, we examined whether children actually found such a test fun and engaging ("usability"), and whether it is capable of producing sensible results ("accuracy"). If so, this would provide grounds for formally validating such a test in a larger sample of amblyopic children in future.

Amblyopia constituted a particularly good test case because, while there is no "gold standard" CSF measure, it is non-contentiously the case that the fellow eye should present as more sensitive (higher CSF) than the affected eye ("case-control effect") and that children with more severe amblyopia should exhibit a greater between-eye difference in their CSF than children with less severe amblyopia ("dose effect").

METHODS

Participants

Participants were 25 children aged 4.0–9.2 years (median {iqr} age 5.8 {1.6} years), with an established clinical diagnosis of strabismic ($N = 7$), anisometropic ($N = 9$) or mixed amblyopia ($N = 9$). Participants were classified as severely (> 0.6 logMAR), moderately (0.3 – 0.6 logMAR), or mildly (0.2 – 0.3 logMAR) amblyopic, based on best-corrected logMAR acuity in their most affected eye. Acuity was assessed by crowded Kay optotypes (< 5.0 years) or Thompson crowded letter chart (≥ 5.0 years).

Participants were required to have an interocular difference in best-corrected visual acuity (BCVA) of 0.2 logMAR or greater (22), and to be receiving or starting treatment (either patching or atropine). Exclusion criteria were: (1) other ocular abnormalities; (2) neurological abnormalities including cerebral visual impairment; (3) developmental disorders; or (4) deprivation amblyopia.

Children were recruited from children's clinics (orthoptic and consultant-led clinics at Moorfields Eye Hospital, London, UK) as part of a wider research project studying changes in suppression and visual function during conventional treatment for childhood amblyopia (to be reported elsewhere). The research was carried out in accordance with the tenets of the Declaration of Helsinki, and was approved by the UK Health Research Authority (REC ID #18/SC/0700; IRAS ID #248985).

The "pCSF" Test

Our novel "pop contrast sensitivity function" (pCSF) test is shown in **Figure 1A**. Users were simply required to "pop bubbles" (circular Gabor patches), by touching them as they "bounced" around a tablet screen. This was a truly gamified procedure (i.e., not a cosmetic wrapper for a conventional psychophysical test), although it was a fairly rudimentary implementation, which included relatively few audiovisual features and was primarily coded over a single weekend. The MATLAB source-code for the pCSF test is freely available at: https://github.com/petejonze/pCSF.

Hardware

The hardware consisted of a Microsoft Surface Pro 4 tablet, featuring a 12.3 inch, 2,736 × 1,824-pixel touchscreen display (Microsoft, Redmond, Washington, U.S.). This display is only 8-bit, so bit-stealing was used to obtain >10-bit luminance precision (23). The screen was calibrated (linearized) using central measurements from a CRS ColorCal Mk 2 colorimeter (Cambridge Research Systems, Cambridge, UK). It was not corrected for spatial non-uniformity. The screen's front-facing camera was used to perform real-time head pose estimation (see below), and ran at a spatial resolution of 640 × 640-pixels.

Software

The test was programmed in MATLAB 2016b, using Psychtoolbox v3 (24). Head tracking was performed by OpenFace 2.2.0 (25, 26), which is open-source software composed of compiled C++ code with various third-party dependencies, including OpenCV (27).

Stimuli

The stimuli consisted of horizontal Gabor patches, presented against an isoluminant gray background (the mean {range} luminance across a 4 × 6 grid of uniformly spaced screen locations was 94.9 {88.1–102.4} cd/m^2). Up to a maximum of five Gabors could be present simultaneously. On each frame, the probability of one new Gabor appearing was 1/60·N, where N was the number of Gabors already present and 1/60 represents the refresh rate of the screen. This meant that that the probability of a new Gabor appearing was inversely proportional to the number of Gabors already present. In practice, a new stimulus appeared on average every 3.1 s (median). During the test, each Gabor traveled independently across the screen, changing direction when reaching the screen edge or when touching

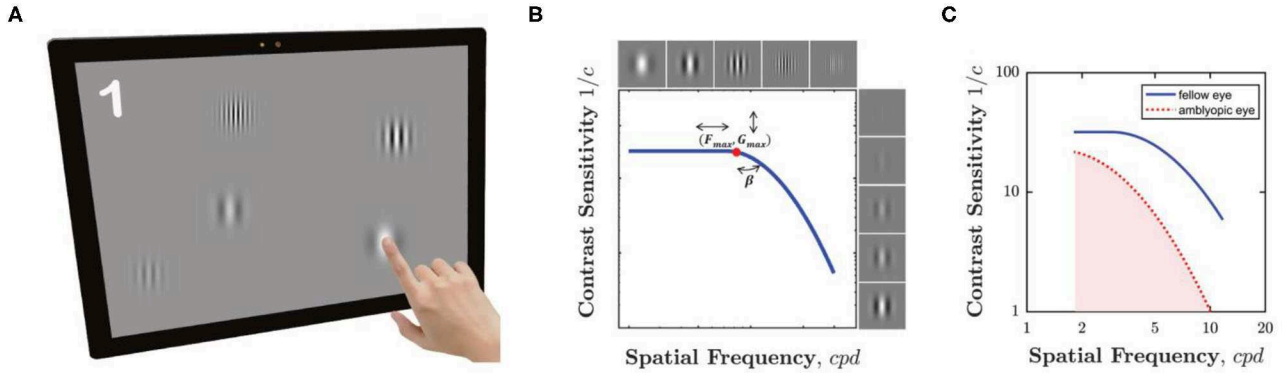


FIGURE 1 | The pCSF test. **(A)** Hardware. Participants pressed equiluminant Gabor patches of variable frequency and contrast as they bounced around the screen. The maximum of five Gabors is shown here for effect, but typically only one or two (or zero) were displayed at any one time, and often some Gabors were suprathreshold. **(B)** Psychophysical algorithm. Stimulus selection and model fitting was performed using a Maximum Likelihood (QUEST+) algorithm, which attempted to fit the three-parameter model shown graphically here (and described formally in Equation 1). **(C)** Example CSFs for a single child aged 5.2 years, for both their amblyopic eye (red dashed line) and fellow eye (blue solid line). The shaded region indicates the area under the contrast sensitivity function (AUCSF) summary measure. A higher AUCSF value indicates greater overall contrast sensitivity.

another Gabor. Their direction and velocity was determined by a simple approximation of molecular dynamics (the Lennard-Jones potential, see source code for details). In practice median {IQR} velocity across all trials was 151 {104–191} pixels/sec. Phase and orientation were fixed at 0 and 90°, respectively. Michelson contrast and spatial frequency were free parameters, controlled by the psychophysical algorithm (*see below*). The range of spatial frequencies was limited such that every Gabor contained at least four visible bands, with a minimum of two pixels per band. The standard deviation of the Gaussian hull, SD_{px}, was fixed at 52 pixels. At a nominal viewing distance of 50 cm, 52 pixels corresponds to 0.55° visual angle, meaning that 99% of stimulus energy fell within a diameter of 2.83°. For drawing purposes, the total spatial support (diameter) of each Gabor was 350 pixels. Each Gabor was visible for a maximum of 6 s (or until pressed), and to avoid hard temporal edges, the onset/offset of each Gabor was temporally ramped with 1-s cosine filters. The starting location of each Gabor was random, but constrained so that Gabors never overlapped or appeared outside the area of the screen.

Task

A Gabor was considered "Hit" if the participant pressed within 221 pixels of its center ($3 \cdot \sqrt{2} \cdot SD_{px}$) within 7 s of its onset (i.e., its 6 s visible duration, plus a 1 s grace period to allow for any ongoing motor responses to complete after stimulus offset). Stimuli not pressed within 7 s were removed and considered a "Miss." Following each Hit, a "pop" sound played, and the Gabor was replaced with an image of a coin, 1 s in duration. To discourage guessing, a negative buzzer sound played after each False Alarm, though False Alarms were not entered into the psychophysical algorithm (see *next*). A running score was visible at the top left of the screen. This score began at zero, and increased/decreased by 1 after each Hit/False-Alarm (minimum: zero). There was no feedback or loss of points following a Miss. Note that the score was for motivational purposes only. These data are not reported, and because of the adaptive nature of the design all children would be expected to attain a similar score, irrespective of their CSF.

Psychophysics

The core psychophysical algorithm (the "back end") consisted of a QUEST+ (13, 14) (Maximum Likelihood) procedure, similar to the qCSF (7–12). It was the same algorithm that we have described in detail previously (6). However, in previous works it received input from a conventional four-alternative forced choice (4AFC) psychophysical task, whereas here the "front end" input was provided from the unconstrained, gamified procedure described above. In brief, the algorithm attempted to fit the 3 parameter model illustrated graphically in **Figure 1B**, and which is given formally by:

$$\alpha = \begin{cases} 1/exp_{10}\left(log_{10}\left(G_{max}\right) - log_{10}\left(2\right)\left(\frac{log_{10}(f)-log_{10}(F_{max})}{log_{10}(2\beta)/2}\right)^2\right) & \text{if } f > F_{max} \\ log_{10}\left(G_{max}\right) & \text{otherwise} \end{cases}, \quad (1)$$

where G_{max} represents peak gain (contrast sensitivity), F_{max} peak spatial frequency, and β the rate of fall-off in sensitivity at high frequencies (full width half maximum, in octaves). Note that this formulation of the CSF represents a modified version of the log-parabola model recommended previously by Lesmes (11) and others (28). For simplicity, however, no fall-off at low spatial frequencies was included, allowing us to reduce the free parameters in our model to 3 (plus one for lapse rate, see below). This modification is unlikely to have had a substantive detrimental impact on the present results, since no stimuli below ∼1.8 cycles per degree (cpd) were presented.

The stimulus domain consisted of 15 Michelson Contrast values log-spaced from 0.01 to 100%, and 10 spatial frequency values log-spaced from 0.019 to 0.125 cycles per pixel (1.8 to 11.7 cpd, assuming a nominal viewing distance of 50 cm). The parameter domain consisted of 15 G_{max} values log-spaced from 3 to 300; 10 F_{max} values log-spaced from 1 to 10; and 9 β values linearly spaced from 1 to 9: all with uniform priors. The underlying psychometric function was assumed to be a Weibull psychometric with a fixed slope of 3, a fixed lower asymptote (guess rate) of 0.05, and a variable (fitted) upper asymptote (lapse rate) of 0.05, 0.1, or 0.2. The model was updated after a Hit or a Miss (but not after a False Alarm).

Maximum likelihood algorithms are typically terminated after either a fixed number of trials, or when a given level of statistical confidence has been reached. As this was an initial feasibility assessment, however, the experimenter (author DE) terminated the test manually after the child had made ∼30 correct responses. The medan {IQR} N trials, including misses, was 51 {43–63}.

Analysis

Following standard practice, the final estimates of G_{max}, F_{max} and β were computed as the mean of the QUEST+ posterior probability distribution. This distribution was refitted *post hoc* for greater fidelity. When doing so, the G_{max} and F_{max} parameter domains were increased to 40 elements each (NB: resulting in much larger search space, that could not have been processed in real time during the experiment itself). Furthermore, the spatial frequency stimulus domain was increased to 30 values log-spaced between 1 and 15 cpd. When performing this refitting, the spatial frequency of each Gabor patch was also recomputed, based on the presented stimulus value (in pixels), and the estimated viewing distance at stimulus offset, as estimated by OpenFace (see *next*).

Head Pose Estimation

The location of the observer's head was monitored continuously by the tablet's front-facing camera, via OpenFace 2.2.0: a free machine-learning tool for facial landmark detection, head pose estimation, facial action unit recognition, and eye-gaze estimation (25). Estimates of viewing distance were made using a speed-optimized Convolutional Experts Constrained Local Model (CE-CLM). This yielding one vector of $\langle x, y, z \rangle$ location coordinates, in millimeters, per video frame. Estimates were made at ∼29 Hz, although the sampling rate varied, depending on CPU availability. Note that OpenFace makes various

assumptions in order to estimate viewing distance (e.g., regarding interpupillary distance). Ideally, these assumptions would be replaced by empirical measurements from each individual [or, alternatively, distance estimates could be calibrated by having observeørs wear/hold an object of known size; see Ref∼(29)]. None of this was not done in the present work, however, as we wanted the test to remain as simple and pragmatic as possible. This may have contributed to random or systematic measurement error in the final CSF estimates.

Ideally, information regarding viewing distance would have been factored into the psychophysical algorithm live, during testing [i.e., as we have done previously when estimating visual fields (30)]. We did not do so here, however, as integrating live measurements is a non-trivial task (i.e., requires extensive pre-processing, and additional code), and if performed incorrectly can be counterproductive. This information was therefore factored in *post hoc*, based on estimated viewing distance at stimulus offset (i.e., at the point when the trial was scored a Hit or a Miss). Across all trials, median {IQR} estimated viewing distance was 531 {378–755} mm.

Procedure

Testing was performed monocularly (once per eye), in a controlled research space. During testing, children wore their habitual best-corrected glasses, and the non-test eye was patched. The starting eye was randomized. In nine children (36%), the child's eyes had been dilated with tropicamide as part of their prior clinical appointment. However, the presence or absence of dilation did not appear to have any substantive impact on the results (see *Results*).

Children were not given any practice prior to testing, and were told simply to press any black and white stripes that they saw. Before the first trial, a tape measure was used to position the participant's head ∼50 cm from the screen, and they were asked to keep their head still during testing. However, viewing distance was not strictly enforced, and children were observed to move considerably during testing (movements that were corrected for *post hoc* using head tracking). Note that 50 cm represents a tradeoff. Farther viewing distances allow higher spatial frequencies to be presented/tested (i.e., given the limited pixel density of the screen), and mean that a given head movement (in cm) has a relatively smaller effect on stimulus size (in degrees visual angle). Conversely, a shorter viewing distance (e.g., ∼30 cm) would likely have been closer to the child's natural comfortable working distance ("Harmon distance"), and so may have reduced head movements. The choice of 50 cm was based on informal piloting, and may not have been optimal, particularly for young children.

Testing took place in a single session, and lasted no more than 10 min total. This testing took place after the child had completed a routine clinical appointment (∼60 min), and after a further ∼45 min of conventional psychophysical testing, as part of a wider research project (data collection ongoing). The pCSF test was the last task before children were discharged, and some of the younger children were visibly fatigued at this point.

Key outcome measures included: estimated contrast sensitivity (AUCSF), completion rates and test durations.

User Feedback

Informal user feedback was obtained in two ways. First, the final 12 participants were asked to rate their enjoyment of the test, from 1 ("very low") to 5 ("very high"). For reference, they were then asked to rate their enjoyment of a conventional psychophysical procedure performed earlier that day [specifically: a four-alternative forced choice visual crowding task: "Vac-Man," described elsewhere previously (21)]. Further informal feedback was sought from the February 2020 meeting of the Moorfields Young Person's Advisory Group ("Eye-YPAG": https://generationr.org.uk/eye-ypag). The Eye-YPAG is a group of older children (8–16 years), most of whom have first- or second-hand lived experience of various eye/vision conditions (not limited to amblyopia). Many members have experience of a wide range of clinical eye tests, and have taken part in clinical trials. During a 45-min session, these children were shown the pCSF test, and were invited to try it and provide unstructured feedback.

RESULTS

Figure 2 shows area under the CSF (AUCSF) scores for each individual. Ideally, all data points should fall above the unity line (i.e., indicating that the fellow eye is more sensitive than the amblyopic eye; "case-control effect"). This was the case for 100% of severe case, 70% of moderate cases, but only 50% mild cases (chance). Accordingly, the amblyopic eye was significantly less sensitive than the fellow eye in the moderate and severe cases, but not in the mild cases [*Wilcoxon signed-rank test for paired differences in AUCSF*; $P_{mild} = 0.843$; $P_{moderate} = 0.028$; $P_{severe} = 0.016$].

To explore whether a dose effect was also present, **Figure 3** shows pCSF performance (ratio of affected eye AUCSF to fellow eye AUCSF) as a function of disease severity. One way of analyzing the data is to divide individuals into discrete severity groups (mild/moderate/severe), based on their logMAR acuity in their worse eye. This analysis is shown in **Figure 3A**, and indicated that children with more severe amblyopia had poorer pCSF performance than children with less severe amblyopia [*Kruskal-Wallis non-parametric one-way analysis of variance*; $\chi^2 = 9.57$, $P = 0.008$]. An alternative, more nuanced approach is to instead plot pCSF ratios against the ratio in logMAR scores between the two eyes. This more continuous approach is shown in **Figure 3B**, and gave qualitatively a similar result, with children with poorer (higher) logMAR ratios exhibiting poorer (lower) pCSF ratios [*Spearman's Rho*; $r_{23} = -0.62$, $P = 0.001$].

So far we have only considered AUCSF (a summary measure of overall contrast sensitivity). If the analysis in **Figure 3B** was instead repeated using each of the three individual CSF parameters in equation 1, no significant associations with amblyopia severity were observed ($P_{\mathrm{Gmax}} = 0.187$; $P_{\mathrm{Fmax}} = 0.400$; $P_{\beta} = 0.464$). The AUCSF effect was highly conserved, however, both in terms of significance and effect size, if we instead replaced AUCSF with a scalar measure of high-frequency cutoff (log sensitivity at maximum spatial frequency; *Spearman's Rho*; $r_{23} = -0.62$, $P < 0.001$). This may suggest that any differences in

CSF were primarily related to changes at high spatial frequencies only; and/or it may simply reflect the fact that the reference variable (i.e., the x-axis in **Figure 3B**) is based only on logMAR acuity [i.e., and changes at low spatial frequencies may manifest independently (2–4)].

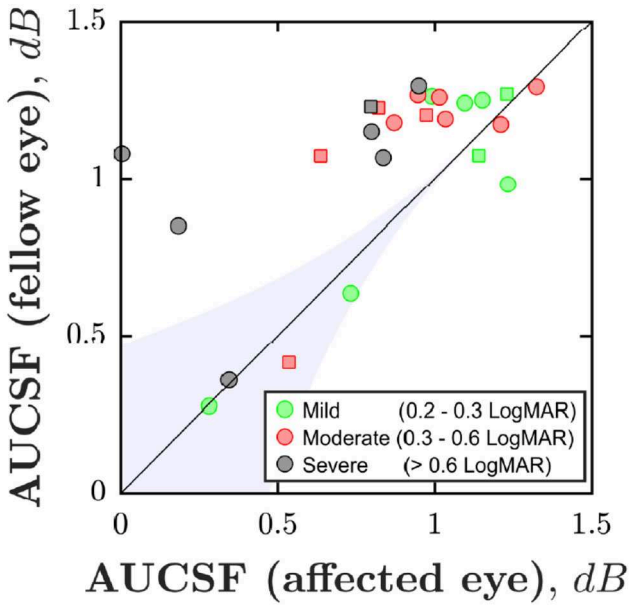


FIGURE 2 | Case-control effect: Area Under the CSF (AUCSF) for each of the 25 individuals, broken down by eye. Marker color denotes amblyopia severity (as quantified by logMAR acuity in worse eye). Squares indicate cases of strabismus without anisometropia. The black line indicates unity (performance similar in both eyes). The blue shaded region of **Figure 2** is for illustration only. However, points falling in a region such as this likely represent general non-compliance (poor performance in both eyes).

Median test duration did not significantly differ between the amblyopic and fellow eye [*Wilcoxon signed-rank test*; $Z = 0.61$, $P = 0.545$], and was 2.7 {CI_{95}: 2.3–2.9} min for amblyopic eyes, and 2.5 {CI_{95}: 2.2–3.1} min for fellow eyes. Across all eyes, median test duration was 2.6 {2.3–2.9} min.

The presence of pupil dilation (mydriatics) did not appear to affect performance, with no differences observed in mean test duration [*Wilcoxon ranked sum test*; $Z = 0.08$, $P = 0.932$], or in the AUCSF ratio [*Wilcoxon ranked sum test*; $Z = 0.37$, $P = 0.713$].

There was a small but significant difference in False Alarm rate between the two eyes [*Wilcoxon signed-rank test*; $Z = 2.62$, $P = 0.009$], with children more frequently pressing the screen incorrectly under their amblyopic (Median: 2.2 per min) vs. fellow eye (Median: 1.8 per min). In amblyopic eyes, the absolute number of False Alarms (which ranged from 1 to 25; Median: 5) also varied as a function of AUCSF [*Spearman's Rho*; $r_{23} = -0.51$, $P = 0.009$], with eyes with the lowest estimated sensitivity associated with the greatest number of False Alarms (note, False Alarm trials were not used when estimating sensitivity). There was, however, no association between estimated sensitivity (AUCSF) and the *total* number of screen presses (both correct and incorrect) [$r_{23} = -0.09$, $P = 0.676$], suggesting that children may have been attempting to maintain a relatively constant rate/number of responses between eyes.

As shown in **Figure 4**, all of the 12 children questioned rated the pCSF test as "enjoyable" or "very enjoyable." Six children (50%) rated the test as more enjoyable than a conventional psychophysical procedure (**Figure 4**, green lines), although two somewhat preferred the conventional procedure, and four rated both equally highly.

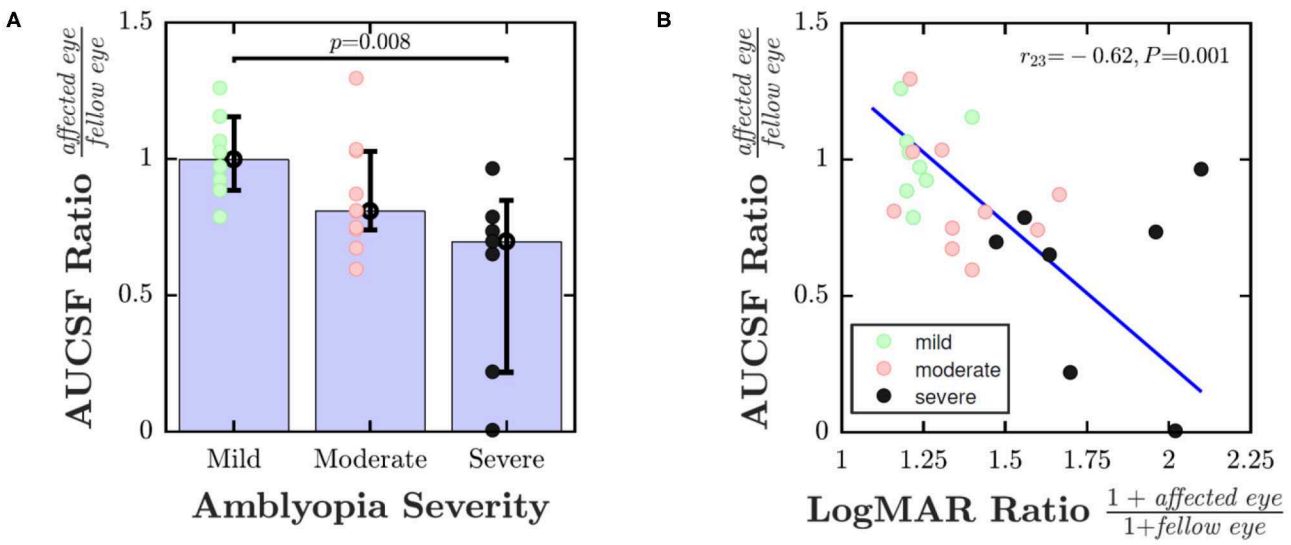


FIGURE 3 | Dose effect: pCSF performance (ratio of affected eye AUCSF to fellow eye AUCSF) as a function of amblyopic severity. In **(A)** severity of amblyopia was categorized by logMAR acuity in worse eye (Mild: ≤ 0.3; Moderate: 0.3–0.6; Severe: > 0.6). In **(B)** severity of amblyopia was computed as the ratio of 1+logMAR acuity in the affected vs. fellow eye (+1 to ensure all values non-negative). Error bars indicate medians ± bootstrapped 95% confidence intervals. The line in **(B)** is the least-square geometric mean regression slope. See main text for details regarding statistics.

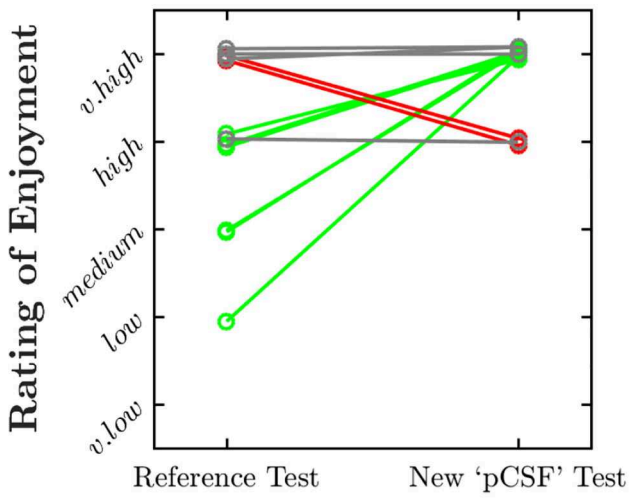


FIGURE 4 | Paired ratings of enjoyment for the present "pCSF" test, and a more conventional psychophysical procedure. Values have been jittered slightly along the y-axis for visibility. Green and red lines highlight instances in which the child rated the present test as more enjoyable or less enjoyable, respectively.

The Eye-YPAG were generally positive in their assessment of the test. Several members noted that the test was more comfortable than conventional psychophysical procedures such as microperimetry, and that it was more fun and engaging than a letter chart. Some children remarked that the pCSF was actually somewhat boring, but much less so than current eye tests. The group recommended the future use of more varied feedback, sound effects, and some form of narrative in any future iterations of the test. Several individuals with nystagmus appreciated the lack of fixation cross. It was also noted that the pCSF is harder to cheat on than a static letter chart, where it is sometimes possible to memorize letters across repeat assessments.

In general, it was evident that none of the Eye-YPAG members had any difficulty comprehending what to do, and several individuals, when left to perform the test unsupervised, proceeded to reach a score of 100 or more (i.e., ∼200+ trials): eventually having to be asked to stop.

Two potential issues of the current test were identified. Some individuals noted that, after extended use, smudges on the screen were liable to be mistaken for near-threshold stimuli. It was also noted that the current photopic background of the test made it unsuitable for individuals with photophobia (e.g., achromatopsia).

DISCUSSION

The results showed that a gamified, tablet-based test was able to produce plausible CSF estimates in a small cohort of young amblyopes (4.0–9.2 years). The test was able to separate moderately and severely amblyopic eyes from their fellow eye (case-control effect), and was able to distinguish between individuals with different degrees of visual impairment (dose effect). It was particularly encouraging that even the youngest children exhibited no difficulties completing the test or comprehending what to do, and in general, the children appeared to find the test genuinely enjoyable. This is particularly remarkable given that testing was performed after almost 2 h of clinical and psychophysical assessments, and given that the pCSF test itself was relatively rudimentary, with little in the way of sounds, graphics, or narrative. Conversely, we would hesitate to even attempt multiple CSF assessments in 4- or 5-year-olds using conventional psychophysical methods.

The test was far from perfect. It was insensitive to the effects of mild amblyopia. Furthermore, a minority of tests resulted in obviously spurious data, likely due to general non-compliance (e.g., low sensitivities in both eyes). These findings are to be expected given that the test was only a rough prototype, and also given the brevity of the test (which could have been allowed to run for longer).

Nevertheless, based on these preliminary findings, it appears that "gamified" vision assessments — such as the pCSF test described here — exhibit early promise as a potential means of estimating CSFs in young children: estimates which could in turn be used to identify, monitor, or stratify the severity of amblyopia.

At present, such functionality is provided by letter charts. However, digital tests could have substantial practical benefits by allowing vision to be measured outside of conventional eye-care facilities. Thus, for chronic conditions such as amblyopia, there is considerable interest in the idea of home monitoring, which has the potential to make treatment cheaper and more convenient by minimizing the number of in-person monitoring appointments (31). Home monitoring could also improve treatment outcomes by allowing for more frequent vision assessments (e.g., every few days or weeks, rather than every few months at present). This could be particularly beneficial for amblyopia given its low rates of treatment compliance (32, 33), and high rates of recurrence (34). Conventional tests such as letter charts are inappropriate for home monitoring, since a technician must be present to explain what to do, ensure the correct lighting and viewing distance, keep the child motivated and on-task, and record the results. In contrast, if, as in the present work, we can find tasks that children actually enjoy performing, and combine these with "smart" digital technologies (e.g., capable of monitoring viewing distance and ambient lighting autonomously), then home monitoring starts to become a realistic prospect. It is possible, for instance, to imagine a fully automated, cloud-based system in which the results of a digital test are transmitted securely to clinicians, who can then dynamically titrate or reinitiate patching remotely, or flag up high-risk cases for immediate, in-person review.

It is further possible that digital assessments of the whole CSF may be able to provide a more detailed and comprehensive characterization of visual impairment than conventional measures of acuity alone. For instance, previous studies have indicated that some amblyopes exhibit selective deficits at low spatial frequencies, independent of acuity (2–4). The present work is consistent with these previous findings, in that G_{max} (peak sensitivity) did not correlate with a conventional measure of amblyopic severity (based purely on acuity), and may therefore provide additional information not captured by acuity

alone. In short, by measuring contrast sensitivity across a broad range of spatial frequencies, it is conceivable that pathologies such as amblyopia may be detected earlier, or monitored more robustly. Such benefits remain unproven, however, and at this stage we consider them secondary to the more practical benefits of digital assessment (detailed *above*).

Limitations and Future Work

The present work represents a first step toward more engaging, child-friendly vision tests. However, it is only a first step: a preliminary assessment of feasibility. Further studies are required to formally assess the performance (sensitivity and specificity) of gamified measures, and to answer outstanding questions, such as whether gamified assessments are effective at sustaining a child's interest across repeated use, how robust they are when applied to a more diverse population (e.g., children with developmental delay), and whether the resulting data would be sensitive enough to detect changes in vision over time (e.g., due to treatment, or disease progression). These are questions that can only be answered by larger, longitudinal trials.

From a practical perspective, there are also myriad practical challenges to address before a test such as the one described in the present work could be made widely clinically available. These challenges include technical considerations (e.g., how to obtain more accurate estimates of viewing distance, how to factor these measurements into the psychophysical algorithm in real-time, how to store and transmit test data securely, and how to integrate the results with existing medical record systems), legal requirements (e.g., medical device certification), and issues surrounding usability and acceptability (i.e., among the patients themselves, their families, and also clinicians). Furthermore, even with a maximally engaging test, some instances of distraction or loss of concentration are inevitable. To achieve truly robust, unsupervised measurements will require autonomous means of verifying the user's identity, and of monitoring if/when they are performing the test correctly. These are non-trivial challenges, but ones that we have made initial steps toward solving using various computer vision and machine learning techniques (18, 35).

It may also be helpful in future to give further consideration to between-eye differences in response criterion. Thus, False Alarms were greater in the amblyopic eye, and tended to increase with severity. Put simply, children appeared disinclined to not press the screen for long periods, even when nothing was visible (note that while an adaptive algorithm would, given infinite trials, be expected to present the same proportion of visible stimuli to all eyes, in practice the algorithm always started from the same baseline stimulus level, and at some spatial frequencies amblyopic eyes might be at floor). The predicted effect of a more liberal response criterion would be to cause amblyopic severity to be *underestimated*. This may, however, have been offset in practice by the fact that the chance of a "lucky guess" was — in contrast to conventional n-alternative-forced-choice designs — relatively small (e.g., the probability of a random pixel/screen-press falling within any single Gabor was ∼1%). I.e., and as the probability of a guess being correct tends toward zero, the deleterious effect of guessing becomes negligible. Nevertheless, given sufficient normative data it might in future be possible to "correct," *post hoc* for the likely effect of response bias on performance (36). Furthermore, it may be prudent to display False Alarm rates as part of any test output, for general consideration by the assessing clinician (i.e., like when assessing visual fields via standard automated perimetry). Finally, the present work should not be taken to indicate that tried-and-tested psychophysical methods can easily be modified or replaced. For example, some clinical trials may benefit from the sorts of highly precise outcome measures that only a conventional psychophysical procedure can provide. While in older children, or in situations where the child *can* be manually supervised, the benefits of gamification may be negligible. Furthermore, it is important to note that even in the present study we focused only on one relatively successful approach to gamification (the pCSF test). During piloting, however, we also explored a range of other methods, many of which were unmitigated failures. This included, for example, one test in which the user is asked to "draw" their CSF directly, by tracing around striped parts of the screen (see **Supplemental Text**). Informal piloting (in non-naïve adults) indicated that this method was promising, and such as approach has been suggested previously as a potential, ultra-fast measure of the CSF (20, 37, 38). Children, however, seemed to find the task confusing — giving hesitant and highly variable responses — and the test was unable to differentiate between the two eyes, even in severe cases of amblyopia (see **Supplemental Text**).

SUMMARY AND CONCLUDING REMARKS

The present work demonstrates the feasibility of using a truly gamified psychophysical procedure to measure spatial vision (the CSF) in amblyopic children. The pCSF test, which involved pressing equiluminant Gabor patches as they bounced around a tablet screen, and which used head tracking to control for changes in viewing distance, was appealing and intuitive to children, and exhibited promising, though imperfect, sensitivity. These preliminary findings suggest that there may be merit in developing such gamified procedures further, and in performing larger-scale investigations regarding their reliability, accuracy, adherence, and clinical utility. Such measures could be particularly valuable for assessing children outside of conventional eye-care facilities (e.g., home-monitoring or school screening).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The research was carried out in accordance with the tenets of the Declaration of Helsinki, and was approved by the UK

Health Research Authority (REC ID #18/SC/0700; IRAS ID #248985). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

PJ conceived the work, created the test materials, analyzed the data, and wrote the manuscript. PJ and DE designed the experiments. DE performed the experiment and collected the data. All the authors contributed toward finalizing the draft manuscript.

FUNDING

This work was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, and by a Newton-Mosharafa Fund PhD Scholarship, administered by the Egyptian Ministry of Higher Education and Scientific Research, in collaboration with the British Council. The Eye YPAG was funded by Moofields Eye Charity, the NIHR, and Santen. Any views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the UK Department of Health.

ACKNOWLEDGMENTS

TD was supported by the Moorfields Eye Hospital and Insitute of Ophthalmology Biomedical Research Centre, and Moorfields Eye Charity (R190044A, R190029A).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2020.00469/full#supplementary-material

REFERENCES

1. Pelli DG, Bex P. Measuring contrast sensitivity. *Vision Res.* (2013) 90:10–4. doi: 10.1016/j.visres.2013.04.015
2. Howell ER, Mitchell DE, Keith CG. Contrast thresholds for sine gratings of children with amblyopia. *Invest Ophthalmol Vis Sci.* (1983) 24:782–7.
3. Wang G, Zhao C, Ding Q, Wang P. An assessment of the contrast sensitivity in patients with ametropic and anisometropic amblyopia in achieving the corrected visual acuity of 1.0. *Sci Rep.* (2017) 7:42043. doi: 10.1038/srep42043
4. Freedman RD, Thibos LN. Contrast sensitivity in humans with abnormal visual experience. *J Physiol.* (1975) 247:687–710. doi: 10.1113/jphysiol.1975.sp010952
5. Pelli D, Robson JG, Wilkins AJ. The design of a new letter chart for measuring contrast sensitivity. *Clin Vis Sci.* (1988) 2:187–99.
6. Farahbakhsh M, Dekker TM, Jones PR. Psychophysics with children: evaluating the use of maximum likelihood estimators in children aged 4 – 15 years (QUEST+). *J Vis.* (2019) 19:22. doi: 10.1167/19.6.22
7. Dorr M, Lesmes LA, Elze T, Wang H, Lu Z-L, Bex PJ. Evaluation of the precision of contrast sensitivity function assessment on a tablet device. *Sci Rep.* (2017) 7:46706. doi: 10.1038/srep46706
8. Hou F, Lesmes LA, Kim W, Gu H, Pitt MA, Myung I, et al. Evaluating the performance of the quick CSF method in detecting contrast sensitivity function changes. *J Vis.* (2016) 16:18. doi: 10.1167/16.6.18
9. Hou F, Lesmes L, Bex P, Dorr M, Lu Z-L. Using 10AFC to further improve the efficiency of the quick CSF method. *J Vis.* (2015) 15:2. doi: 10.1167/15.9.2
10. Hou F, Huang C-B, Lesmes L, Feng L-X, Tao L, Lu Z-L, et al. qCSF in clinical application: Efficient characterization and classification of contrast sensitivity functions in amblyopia. *Invest Ophthalmol Vis Sci.* (2010) 51:5365–77. doi: 10.1167/iovs.10-5468
11. Lesmes LA, Lu Z-L, Baek J, Albright TD. Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J Vis.* (2010) 10:17. doi: 10.1167/10.3.17
12. Dorr M, Lesmes LA, Lu Z-L, Bex PJ. Rapid and reliable assessment of the contrast sensitivity function on an iPad. *Invest Ophthalmol Vis Sci.* (2013) 54:7266–73. doi: 10.1167/iovs.13-11743
13. Watson AB. QUEST+: a general multidimensional Bayesian adaptive psychometric method. *J Vis.* (2017) 17:10. doi: 10.1167/17.3.10
14. Jones PR. QuestPlus: a matlab implementation of the QUEST+ adaptive psychometric method. *J Open Res Softw.* (2018) 6:195. doi: 10.5334/jors.195
15. Levitt H. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am.* (1971) 49:467–77. doi: 10.1121/1.1912375
16. Kaernbach C. Simple adaptive testing with the weighted up-down method. *Percept Psychophys.* (1991) 49:227–9. doi: 10.3758/BF03214307
17. Manning C, Jones PR, Dekker TM, Pellicano E. Psychophysics with children: Investigating the effects of attentional lapses on threshold estimates. *Percept Psychophys.* (2018) 80:1311–24. doi: 10.3758/s13414-018-1510-2
18. Jones PR. Sit still and pay attention: using the wii balance-board to detect lapses in concentration in children during psychophysical testing. *Behav Res Methods.* (2018) 51:28–39. doi: 10.3758/s13428-018-1045-4
19. Bosten J, Álvaro L, Alvarez J, Meyer B, Tang T, Maule J, et al. Tablet-based app for screening for CVD in young children. *J Vis.* (2019) 19:71. doi: 10.1167/19.8.71
20. Hosokawa K, Maruya K, Nishida S, Takahashi M, Nakadomari S. Gamified vision test system for daily self-check. In *2019 IEEE Games, Entertainment Media Conference*. New Haven, CT (2019). doi: 10.1109/GEM.2019.8811563
21. Greenwood JA, Tailor VK, Sloper JJ, Simmers AJ, Bex PJ, Dakin SC. Visual acuity, crowding, and stereo-vision are linked in children with and without amblyopia. *Invest Ophthalmol Vis Sci.* (2012) 53:7655–65. doi: 10.1167/iovs.12-10313
22. Levi DM, Harwerth RS, Smith EL. Binocular interactions in normal and anomalous binocular vision. *Doc Ophthalmol.* (1980) 49:303–24. doi: 10.1007/BF01886623
23. Tyler CW, Chan H, Liu L, McBride B, Kontsevich LL. Bit stealing: how to get 1786 or more gray levels from an 8-bit color monitor. *Int Soc Opt Eng.* (1992) 1666:351–65. doi: 10.1117/12.135981
24. Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C. What's new in Psychtoolbox-3. *Perception.* (2007) 36:1–16. doi: 10.1068/v070821
25. Baltrusaitis T, Zadeh A, Lim YC, Morency L-P. Openface 2.0: facial behavior analysis toolkit. In: *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. Xi'an (2018). doi: 10.1109/FG.2018.00019
26. Baltrušaitis T, Robinson P, Morency L-P. Openface: an open source facial behavior analysis toolkit. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Lake Placid, NY (2016). doi: 10.1109/WACV.2016.7477553
27. Pulli K, Baksheev A, Kornyakov K, Eruhimov V. Real-time computer vision with OpenCV. *Commun ACM.* (2012) 55:61–9. doi: 10.1145/2184319.2184337
28. Watson AB, Ahumada AJ. A standard model for foveal detection of spatial contrast. *J Vis.* (2005) 5:6. doi: 10.1167/5.9.6
29. Hamm LM, Mistry K, Black JM, Grant CC, Dakin SC. Impact of children's postural variation on viewing distance and estimated visual acuity. *Transl Vis Sci Technol.* (2019) 8:16. doi: 10.1167/tvst.8.1.16
30. Jones PR, Lindfield D, Crabb DP. Using an open-source tablet perimeter (Eyecatcher) as a rapid triage measure in a glaucoma clinic waiting area. *Br J Ophthalmol.* (2020). doi: 10.1136/bjophthalmol-2020-316018. [Epub ahead of print].

31. de Mul M, de Bont AA, Reus NJ, Lemij HG, Berg M. Improving the quality of eye care with tele-ophthalmology: shared-care glaucoma screening. *J Telemed Telecare.* (2004) 10:331–6. doi: 10.1258/1357633042602107
32. Awan M, Proudlock FA, Grosvenor D, Choudhuri I, Sarvananthan N, Gottlob I. An audit of the outcome of amblyopia treatment: a retrospective analysis of 322 children. *Br J Ophthalmol.* (2010) 94:1007–11. doi: 10.1136/bjo.2008.154674
33. Wallace MP, Stewart CE, Moseley MJ, Stephens DA, Fielder AR. Compliance with occlusion therapy for childhood amblyopia. *Invest Ophthalmol Vis Sci.* (2013) 54:6158–66. doi: 10.1167/iovs.13-11861
34. Bhola R, Keech RV, Kutschke P, Pfeifer W, Scott WE. Recurrence of amblyopia after occlusion therapy. *Ophthalmology.* (2006) 113:2097–100. doi: 10.1016/j.ophtha.2006.04.034
35. Jones PR, Demaria G, Tigchelaar I, Asfaw DA, Edgar DF, Campbell P, et al. The human touch: Using a webcam to autonomously monitor compliance during visual field assessments. *Transl Vis Sci Technol.* (2020) 9:31. doi: 10.1167/tvst.9.8.31
36. Jones PR, Moore DR, Shub DE, Amitay S. The role of response bias in perceptual learning. *J Exp Psychol Learn Mem Cogn.* (2015) 41:1456–70. doi: 10.1037/xlm0000111
37. Mulligan JB. A method for rapid measurement of contrast sensitivity on mobile touch-screens. *Electron Imaging.* (2016) 2016:1–6. doi: 10.2352/ISSN.2470-1173.2016.16.HVEI-104
38. Tardif J, Watson M, Giaschi D, Gosselin F. Measuring the contrast sensitivity function in just three clicks. *J Vis.* (2016) 16:966. doi: 10.1167/16.12.966

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 21 December 2020
doi: 10.3389/fnins.2020.555701

Multi-Stage Cortical Plasticity Induced by Visual Contrast Learning

Jie Xi[1,2], Pan Zhang[1,2,3], Wu-Li Jia[1,2,4], Nihong Chen[5,6], Jia Yang[1,2], Ge-Tong Wang[1,2], Yun Dai[7,8], Yudong Zhang[7,8] and Chang-Bing Huang[1,2]**

[1] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China, [2] Department of Psychology, University of Chinese Academy of Sciences, Beijing, China, [3] Center for Neural Science, New York University, New York, NY, United States, [4] School of Education Science, Huaiyin Normal University, Huaian, China, [5] Department of Psychology, School of Social Sciences, Tsinghua University, Beijing, China, [6] THU-IDG/McGovern Institute for Brain Research, Beijing, China, [7] Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu, China, [8] The Key Laboratory on Adaptive Optics, Chinese Academy of Sciences, Chengdu, China

Perceptual learning, the improved sensitivity via repetitive practice, is a universal phenomenon in vision and its neural mechanisms remain controversial. A central question is which stage of processing is changed after training. To answer this question, we measured the contrast response functions and electroencephalography (EEG) before and after ten daily sessions of contrast detection training. Behavioral results showed that training substantially improved visual acuity and contrast sensitivity. The learning effect was significant at the trained condition and partially transferred to control conditions. Event-related potential (ERP) results showed that training reduced the latency in both early and late ERPs at the trained condition. Specifically, contrast-gain-related changes were observed in the latency of P1, N1-P2 complex, and N2, which reflects neural changes across the early, middle, and high-level sensory stages. Meanwhile, response-gain-related changes were found in the latency of N2, which indicates stimulus-independent effect in higher-level stages. In sum, our findings indicate that learning leads to changes across different processing stages and the extent of learning and transfer may depend on the specific stage of information processing.

Keywords: contrast gain, ERP, latency, perceptual learning, response gain

OPEN ACCESS

Edited by:
Peter J. Bex,
Northeastern University, United States

Reviewed by:
Michael A. Silver,
University of California, Berkeley, United States
Luca Battaglini,
University of Padua, Italy

****Correspondence:***
Chang-Bing Huang
huangcb@psych.ac.cn
Yudong Zhang
ydzhang@ioe.ac.cn

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *26 April 2020*
Accepted: *25 November 2020*
Published: *21 December 2020*

Citation:
Xi J, Zhang P, Jia W-L, Chen N, Yang J, Wang G-T, Dai Y, Zhang Y and Huang C-B (2020) Multi-Stage Cortical Plasticity Induced by Visual Contrast Learning. Front. Neurosci. 14:555701. doi: 10.3389/fnins.2020.555701

INTRODUCTION

Visual perceptual learning (VPL) is a long-term performance improvement in visual tasks as a result of training or experience (Petrov et al., 2005; Sagi, 2011; Deveau et al., 2013; Dosher et al., 2013; Watanabe and Sasaki, 2015). The observed specificity to the trained stimulus, task, or retinal location in psychophysical studies has been generally taken as evidence for neural plasticity in early visual cortex (Karni and Sagi, 1991; Gilbert, 1994; Schoups et al., 1995; Watanabe et al., 2002; Chen and Fang, 2011; Crist et al., 2014). Alternatively, Mollon and Danilova (1996) hypothesized that learning occurs at a more central site but still predicts orientation and location specificity of learning. Models like improved readout or reweighting of representation neurons (e.g., V1) (Poggio et al., 1992; Dorsher and Lu, 1998) and the involvement of high-level processes beyond the visual cortex (Li W. et al., 2008) have been proposed in the last decades and received support from psychophysical (Liu, 1999; Liu and Weinshall, 2000; Xiao et al., 2008; Zhang et al., 2010), neurophysiological (Law and Gold, 2008), and brain imaging studies (Chen et al., 2015, 2017).

However, there is a growing consensus that perceptual learning involves neural processing in multiple brain regions. The reverse hierarchy theory proposed that learning back-propagate from higher to lower visual areas, providing predictive signals to lower-levels and learning site(s) depending on the task difficulty (Friston, 2003; Ahissar and Hochstein, 2004). Indeed, learning a simple task may involve a broad set of brain systems undergoing changes in sensory representations, read-out weights, decision rules, attention and feedback processes as well as sensorimotor changes (Maniglia and Seitz, 2018). The distribution of changes across the neural system may depend upon the physical stimuli as well as the training task. A similar two-stage model suggests that feature-based plasticity occurs in the early sensory processing stages, while task-based plasticity occurs in higher-level processing stages (Sasaki et al., 2013; Shibata et al., 2014, 2016).

Human electrophysiological studies can provide unique contributions to the question regarding learning stages, given different components of ERP reflected processing in different stages along the visual hierarchy (Voorhis and Hillyard, 1977; Luck et al., 2000; Fabiani et al., 2007). Modulations in both the early and late ERP components have been found in different perceptual training studies, ranging from early C1/P1 (Pourtois et al., 2008; Bao et al., 2010; Zhang et al., 2015) to enhancement in N1, P2 (Song et al., 2005; Shoji and Skrandies, 2006; Qu et al., 2010; Wang et al., 2010; Zhang et al., 2013), and later N2 and P3 components (Skrandies and Fahle, 1994; Wang et al., 2010; Hamamé et al., 2011). However, few studies have compared the contribution of early and late ERP components to perceptual learning within a unified theoretical framework.

In this study, we tested the multi-stage hypothesis of perceptual learning. Importantly, we measured ERP with quantitative modeling based on contrast response function (CRF) measurements. In this model, the facilitation of perceptual sensitivity induced by perceptual learning could be accounted for by three possible mechanisms – increased contrast gain, increased response gain, or additive baseline shift (**Figure 1A**). The contrast-gain change model predicts that changes in the ERP components interact with contrast level and lead to a leftward shift in the CRF, i.e., shifting the most sensitive operating range of the system toward lower contrast while the saturation points of the CRF remain fixed. The response-gain model predicts that learning leads to a constant multiplicative change in the ERP components irrespective of the contrast level, signifying by both slope and asymptotic changes of the CRF. The baseline shift model predicts that learning leads to an overall upward, additive gain of the ERP response. We also tested the psychophysical and electrophysiological transfer effect of learning in four control conditions that varied in spatial frequency, retinal location, and eye of origin.

MATERIALS AND METHODS

Subjects

Twenty subjects (23.8 ± 3.8 years, 12 males) participated in the study. All subjects were right-handed and had no psychiatric or neurological disorders, naïve to the task, and of normal or corrected-to-normal vision. All subjects received basic subsidies for their participation and additional bonus if they complete the whole experiments seriously. The study was approved by the Ethical Review Committee of Institute of Psychology, Chinese Academy of Sciences, and informed consent was obtained from each subject.

Apparatus and Stimuli

The experiments were controlled by a desktop computer running Matlab programs (Mathworks, Natick, Massachusetts) and PsychToolBox3 (Brainard, 1997; Pelli, 1997). The stimuli were presented on a gamma-corrected SONY G220 CRT monitor with a spatial resolution of 1600 × 1200 pixels, a refresh rate of 85 Hz, and a mean luminance of 28.7 cd/m^2. A special circuit combined two 8-bit output channels of the graphics card to produce 14-bit gray-level resolution (Li et al., 2003). Subjects viewed the stimuli monocularly with head on a chin rest.

Stimuli were circular sinusoidal gratings, subtending 2° at a distance of 1.38 meters, whose edges were smoothed to the background with a half-Gaussian ramp ($\sigma = 0.31°$) to minimize edge effects. The stimulus centered at 5° away from the fixation point in the upper left (trained location), upper right, or lower left location, depending on the test conditions (**Figure 1B**). The stimulus orientation was 45 or 135° relative to horizontal. Stimulus position was jittered slightly (0 – 0.5°) from trial to trial.

Experimental Design

The experiment consisted of pre-training assessment, training, and post-training re-assessment (**Figure 1C**). Training consisted of 10 sessions; each session was composed of seven blocks of 80 trials and lasted about 30–40 min. In both pre- and post-training assessments, contrast sensitivity function (CSF), visual acuity, and ERP recordings were measured in both eyes. CSF and visual acuity were measured on the first and last day of assessment, taking up to a total of ∼ 40 min. The ERP recordings were performed in the second and the day before the last day of assessment, taking up to a total of 3.5 h (including preparation of ERP recording, data acquisition, and voluntary breaks).

Tasks

Subjects performed a peripheral orientation discrimination task during all the CSF measurements, training, and EEG sessions (**Figure 1B**). Each trial started with a 500–800 ms blank (randomly jittered in time to minimize anticipation and was signaled by a brief tone) and was followed by a grating of 141 ms. Subjects indicated the orientation of the grating by a keypress within 1,500 ms. During training, a brief tone followed each correct response; during pre- and post-tests, a brief tone followed each response regardless of its accuracy. The next trial started after a 600 ms blank. Subjects were instructed to maintain fixation on a black dot at the center of the display. To ensure central fixation, the dot was randomly changed to letter "x" or "o" at a probability of 0.1, and subjects were asked to indicate the change with keypress, i.e., central task.

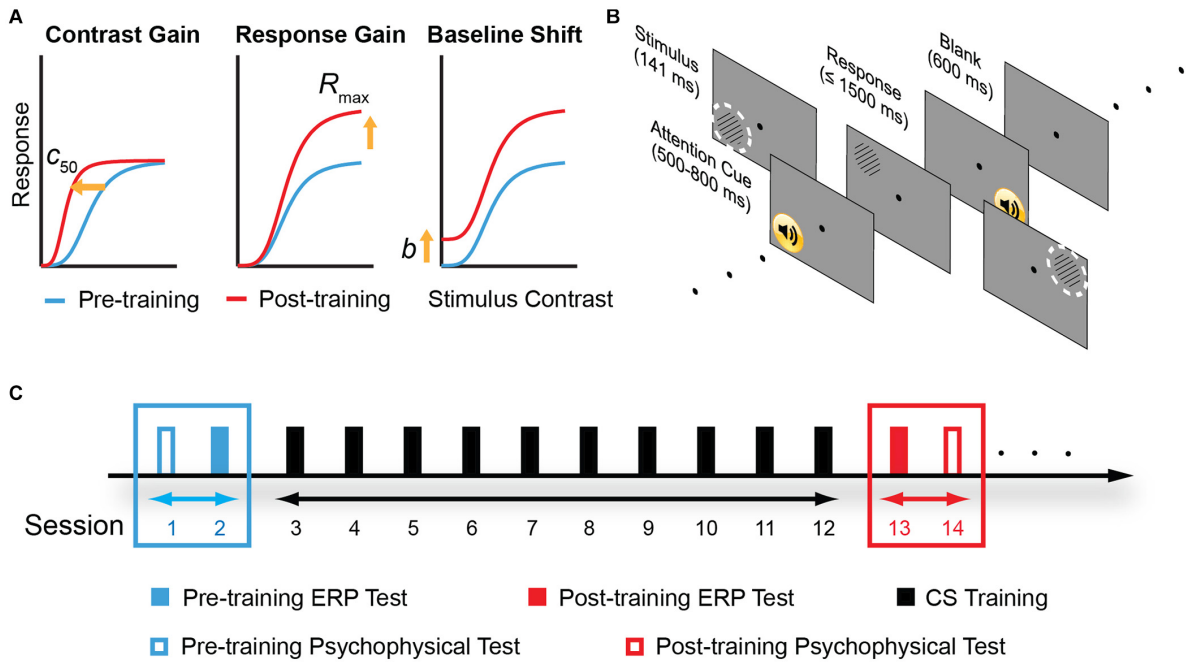


FIGURE 1 | Model predictions, experimental stimuli and protocol. **(A)** Three different mechanisms in the sensory gain model that predict different pattern of contrast response function (CRF) changes following perceptual learning. From left to right: change in contrast gain, multiplicative response gain, or baseline shift. c_{50}: the stimulus contrast that evokes half of the neuron's maximal response. R_{max}: maximal response to stimuli. b: baseline activity. **(B)** A typical trial procedure. Each trial started with an attention cue (500–800 ms). Stimulus was presented for 114 ms, and subjects were asked to report grating orientation within 1,500 ms. After response or 1,500 ms, a blank screen was presented for 600 ms and next trial started afterward. Training was performed in the upper left visual field location relative to the fixation dot. The dashed, white circles indicate two control locations: the upper right and the lower left visual field location. To ensure task compliance, subjects were asked to focus on the central fixation dot and press corresponding key when the black dot changed to "x" or "o" (with 5% probability each). **(C)** Experimental design. Participants were instructed to practice contrast detection tasks for ten consecutive days. Pre- and post-training psychophysical measurements covered contrast sensitivity function and visual acuity. ERP tests were conducted before and after contrast sensitivity training in different days to examine learning-induced changes in neural processing.

Pre- and Post-training Psychophysical Assessments

Visual acuity was measured with the Chinese Tumbling E Chart (Mou, 1966; Huang et al., 2008; Xi et al., 2014) and defined as the logMAR (log minimum angle of resolution) acuity associated with 75% correct identification (Xu et al., 2006; Zhou et al., 2006; Huang et al., 2009).

Contrast sensitivity (CS) was defined as the reciprocal of contrast threshold for detecting a grating with 79.4% accuracy. We measured CS using the quick CSF method (qCSF), which was recently developed by Lesmes et al. (2010) to accurately estimate CSF with greatly reduced testing times by sampling from pre-defined parameter space and updating the probability of CSF parameters based on subject's performance. The stimulus space consisted of gratings contrasts ranging from 0.1% to 99% in steps of 1.5 dBs and spatial frequencies from 0.5 to 8 cycles per degree (cpd) in steps of 3 dBs. The qCSF's parameter space is a four-dimensional grid of the four parameters that defined CSF, i.e., peak gain, peak frequency, bandwidth, and truncation level (Lesmes et al., 2010). The CSF curve was obtained after 100 qCSF trials. The area under contrast sensitivity function (AUCSF), a comprehensive measure of spatial vision over a wide range of spatial frequencies (van Gaalen et al., 2009; Lesmes et al., 2010), was calculated by integrating contrast sensitivity over spatial frequencies varying from 0.5 to 8 cpd. CSF in the upper right, upper left (trained location), and lower left visual field location of left eye (LE, trained eye), and the upper left of right eye (RE, untrained eye) was measured in four separate blocks and counterbalanced across subjects but held constant between pre- and post-training test sessions for a particular subject. Before pre-training CSF measurement, subjects practiced 20 trials to get familiar with the task.

Training

Training was performed in the upper left visual field location of left eye and training spatial frequency was fixed at 5 cpd. A 3-down-1-up adaptive staircase procedure in which three consecutive correct responses resulted in a reduction of signal contrast ($C_{n+1} = 0.90C_n$), and one wrong response resulted in an increase in contrast ($C_{n+1} = 1.10C_n$) was used to control grating contrast (Levitt, 1971).

EEG

The ten conditions conducted during pre- and post-training ERP measurements were summarized in **Table 1**. In the trained condition (spatial frequency: 5 cpd; retinal location: the upper left

visual field location; trained eye: left eye), six different contrast levels were employed to obtain full CRF: 0, 4.26, 8.90, 18.61, 38.90, and 81.13% Michelson contrasts. These six conditions were randomly intermixed in four blocks, each consists of 300 trials. In the control conditions, EEG signals were recorded for gratings of 10 cpd and 38.9% contrast at the trained location (i.e., the upper left visual field location in the left eye with higher spatial frequency and a fixed contrast, Frequency change condition), gratings of 8.9% contrast at the upper right (Location change-contralateral condition), and the lower left location in the left (trained) eye (Location change-ipsilateral condition); and the upper left location in the right (untrained) eye (Eye change condition). These four control conditions were separately presented in four blocks of 200 trials each. Training and control conditions were counterbalanced across subjects.

Scalp EEG data were recorded from 64 scalp electrodes (Neuroscan®) with an amplifier bandpass of DC to 100 Hz and a 60-Hz notch filter was digitized at 500 Hz. Vertical electro-oculogram (VEO) was recorded by electrodes placed above and below the left eye. Horizontal electro-oculogram (HEO) was recorded by electrodes placed at the outer canthus of the left and right eye. The reference electrode was placed on the top of the midline between electrodes C_Z and CP_Z. Electrode impedance was kept <5 kΩ throughout recording.

EEG data were analyzed using EEGLAB ([1]; Delorme and Makeig, 2004) and ERPLAB ([2]; Lopez-Calderon and Luck, 2014) with home-made scripts. Signals were first referenced offline to the average of all the electrodes and filtered with a bandpass filter of 0.1–30 Hz. The data were then epoched starting at 200 ms before stimulus onset and ending 1000 ms after stimulus onset. The data exceeding ± 50 μV at electrode VEO and ± 15 μV at electrode HEO, or other activities exceeding ± 100 μV at any electrodes were excluded from analysis. The overall rejection rate was 17.27%. Remaining epochs were averaged according to the stimulus condition.

The peak amplitude was calculated with a moving window technique: the peak(s) within a certain time window was first determined for each subject and each condition (trained condition: 90–140 ms for P1, 160–300 ms for N1-P2 complex, 400–800 ms for N2; control condition: 110–160 ms for P1, 160–300 ms for N1-P2 complex, 400–800 ms for N2); then the peak value within a certain time window surrounding the first peak was derived for each subject and each condition (30 ms for P1 and 50 ms for N1-P2 complex and N2). To quantify the peak amplitude and latency of each component, the largest three electrodes among six contralateral posterior-occipital electrodes (P4, P6, P8, PO4, PO6, and PO8 in the right hemisphere and P3, P5, P7, PO3, PO5, and PO7 in the left hemisphere) were chosen for further analysis. Electrode sites were selected in temporo-parietal-occipital positions based on previous ERP studies of VPL (Ding et al., 2003; Song et al., 2005; Qu et al., 2010; An et al., 2012; Zhang et al., 2013; Itthipuripat et al., 2014, 2017; Garner et al., 2015; Ahmadi et al., 2018). The amplitude of each component was defined as the height of the peak in this average signal, and the latency was defined as its time to the peak. Amplitudes were measured as peak-to-peak voltages for N1-P2 complex rather than the base-to-peak amplitude due to uncertainties in establishing a baseline voltage for N1 and P2. For statistical analysis, amplitudes and latency were averaged across trials for each condition.

For the trained condition, we subtracted the ERP evoked by 0%-contrast stimulus from the ERP response evoked by all other contrasts to minimize the potential effects of anticipatory ERPs (**Supplementary Figure 1**).

Statistical Analysis

The learning curve (i.e., $\log_{10}$ contrast sensitivity as a function of training session) was fitted with a linear function:

$$\log_{10} CS\,(\text{session}) = CS_0 + \alpha \times \log_{10}(\text{session})$$

where CS denotes contrast sensitivity, CS_0 is the intercept, and α is the slope of the learning curve (learning rate, or unit improvement at the trained condition).

To calculate the spatial frequency bandwidth of perceptual learning, we used the same methods as in our previous paper (Huang et al., 2008). Briefly, contrast sensitivity improvements of each observer were fit with a Gaussian function:

$$\log_{10}\left[CS_{\text{post-training}}(f)\right] - \log_{10}\left[CS_{\text{pre-training}}(f)\right] = a\exp\left[-\left(\frac{\log_2(f) - \log_2(f_0)}{\sigma}\right)^2\right] \quad (1)$$

where CS denotes contrast sensitivity, a is the amplitude of the improvement, f is the spatial frequency, f_o is the spatial frequency with the maximum improvement, and σ is the standard deviation of the Gaussian function. The bandwidth (BW) of perceptual learning was defined as:

$$BW = 2\sqrt{\ln 2}\,\sigma$$

Standard deviations of all the estimated parameters were computed with a resampling method (Maloney, 1990).

The improvement of AUCSF, CS, and the amplitude of each ERP component was defined as:

$$I = \frac{\text{Measure}_{\text{post-training}} - \text{Measure}_{\text{pre-training}}}{\text{Measure}_{\text{pre-training}}} \times 100\%$$

The improvement of visual acuity (in logMAR) and latency of each ERP component was calculated as:

$$I = \frac{\text{Measure}_{\text{pre-training}} - \text{Measure}_{\text{post-training}}}{\text{Measure}_{\text{pre-training}}} \times 100\%$$

Pre- and post-training visual acuity, CS, BW, and learning improvement were compared using paired t-tests and corrected for multiple comparison based on FDR. Pre- and post-training latency and amplitude of each ERP component of control conditions were also compared using paired t-tests and corrected for multiple comparison based on FDR. Evidence against the null

[1] http://sccn.ucsd.edu/eeglab/

[2] http://erpinfo.org/erplab/

TABLE 1 | Stimulus details for ERP measurements.

Conditions	Eye	Location	Spatial frequency (cpd)	Contrast %
Trained	Left	Upper left	5	0
	Left	Upper left	5	4.26
	Left	Upper left	5	8.9
	Left	Upper left	5	18.61
	Left	Upper left	5	38.9
	Left	Upper left	5	81.13
Frequency change	Left	Upper left	***10***	***38.9***
Location change-contralateral	Left	***Upper Right***	5	***8.9***
Location change-ipsilateral	Left	***Lower Left***	5	***8.9***
Eye change	***Right***	Upper left	5	***8.9***

Differences between the trained and each control condition are shown in italic bold.

hypothesis was quantified using Bayes factors (BF_{10}). Repeated ANOVA with Green house-Geisser correction was applied to the effects of training and contrast levels on the latency and amplitude of each ERP component of the trained condition.

ERP Model Fitting

The Naka-Rushton equation was fitted to the ERP amplitude CRFs, i.e., amplitude of P1, N1-P2 complex, and N2 as functions of contrasts (Tolhurst et al., 1981; Li X. et al., 2008).

$$R(c) = b + R_{\max}\frac{c^s}{c_{50}^s + c^s}$$

where c is the grating contrast, b is the baseline activity, c_{50} denotes the contrast at which the response reaches half of its maximum dynamic range, s is exponent controlling how quickly the CRF rises and reaches an asymptote, and R_{max} is the maximum response.

An inverted Naka-Rushton equation was fitted to the ERP latency CRFs, which was earlier shown to provide the best fit to the measured response latencies of neurons in striate cortex of cats and monkeys (Albrecht et al., 2002):

$$R(c) = L_{\max} - R_{\text{shift}}\frac{c^s}{c_{50}^s + c^s}$$

where c is the grating contrast, L_{max} is the max latency, c_{50} denotes the contrast at which the latency reaches half of its minimum dynamic range, s is exponent controlling how quickly the CRF decreases and reaches an asymptote, and R_{shift} is the maximum reduction in latency.

Pre- and post-training model fitting parameters were also compared using paired t-tests and corrected for multiple comparison based on FDR. Evidence against the null hypothesis was quantified using Bayes factors (BF_{10}). By systematically examining the best-fitting parameters of the Naka-Rushton equations to the amplitude and latency of different ERP components before and after training, we fulfilled the comparison between the contribution of early and late ERP components to perceptual learning within a unified theoretical framework.

RESULTS

Behavioral Outcomes

Central Task

Subjects performed the central letter identification task with high accuracy during all the CSF measurements, training, and EEG sessions. There was no significant difference among the central letter identification performances in the four CSF tests before and after training (94.06, 97.12, 96.97, and 96.70% correct in pre-tests at the upper right, upper left, and lower left visual field location in the left eye (LE) and the upper left location in the right eye (RE) vs. 95.90, 93.77, 97.60, and 96.92% in post-tests, respectively; all $p > 0.10$). There was also no significant change in the central task performance during EEG measurements [pre-test: 93.15%, post-test: 94.61%, $t(19) = 1.045$, $p = 0.31$]. We concluded that the learning effects were not compensated from performance decrements in the central task.

Visual Acuity

Training significantly improved visual acuity by 1.0 line in the left (trained) eye [from −0.13 to −0.23 logMAR, $t(19) = 8.025$, $p < 0.005$, $d = 1.617$] and 0.4 line in the right (untrained) eye [from −0.12 to −0.16 logMAR, $t(19) = 3.320$, $p < 0.01$, $d = 0.582$] after multiple comparison correction based on FDR (Benjamini and Yekutieli, 2001). The magnitude of improvement in the trained eye was significantly greater than that in the untrained eye [$t(19) = 3.113$, $p = 0.006$, $d = 0.828$]. In **Figure 2A**, we plotted visual acuity (logMAR) in the trained and untrained eyes after training versus that before training.

Learning Curve

Training at 5 cpd significantly improved CS by 38.17% [$t(19) = 6.108$, $p < 0.001$, $d = 1.404$]. Average learning curve was plotted in **Figure 2B**. The averaged best fitting curve has a slope of 0.306 $\log_{10}$ contrast sensitivity/$\log_{10}$ session ($r^2 = 0.887$).

Contrast Sensitivity Functions

Contrast Sensitivity Functions (CSFs) measured in the upper left (trained location), the upper right, and the lower left location in LE (trained eye) and the upper left location in RE (untrained eye) of all the subjects before and after training were shown in

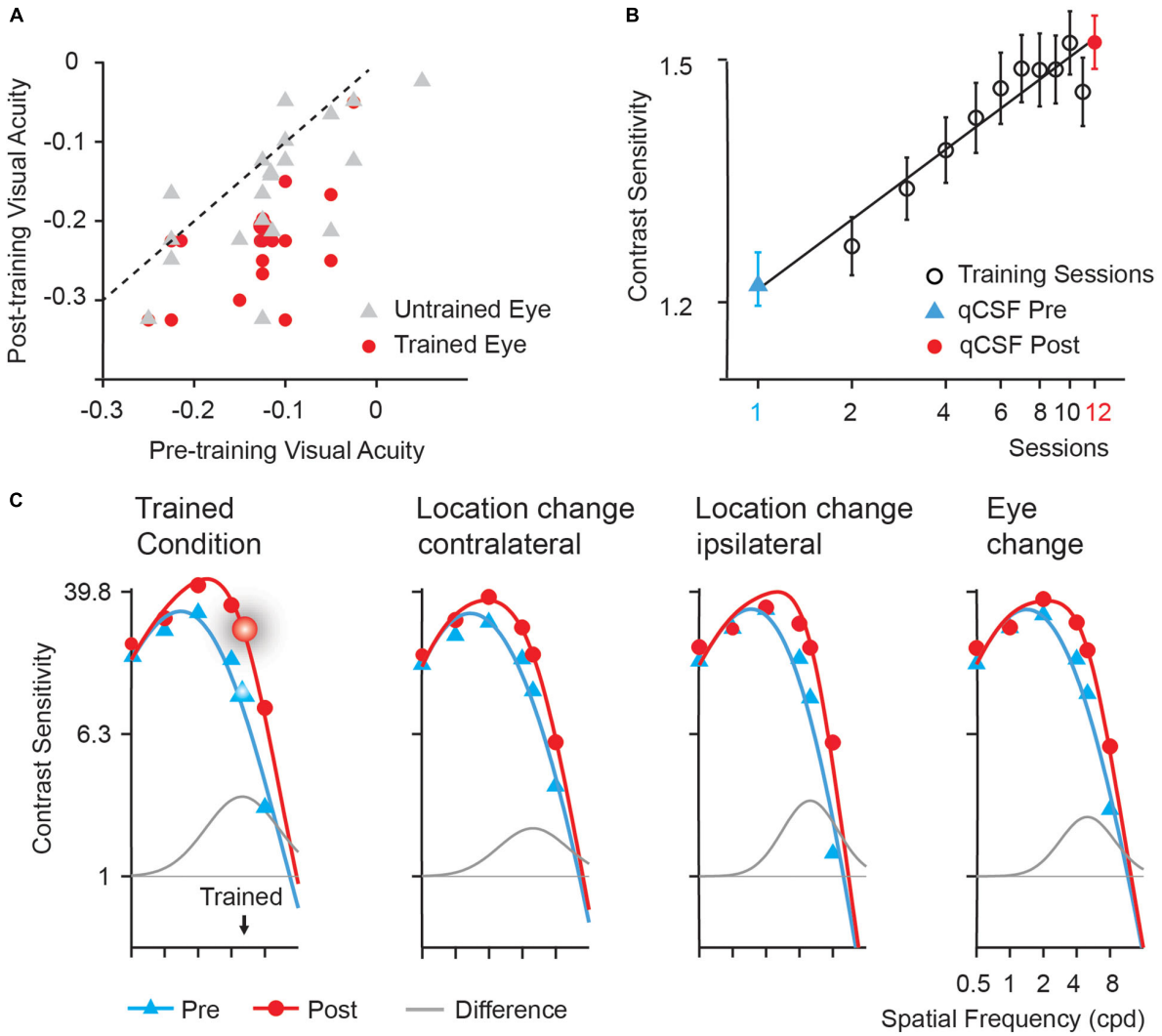


FIGURE 2 | Behavioral results. **(A)** Post- versus pre-training logMAR visual acuity of trained (red circles) and untrained (gray triangles) eyes. Each symbol represents the data of one subject. The dashed line is the identity line (slope = 1), indicating no improvement. **(B)** Learning curve. Error bars represent standard errors across subjects. The first (blue triangles) and last data points (red circles) were derived from pre- and post-training CSF measurements in the trained condition, respectively. Black open circle: data from training phase. **(C)** Pre- (blue curves) and post-training (red curves) CSFs and the difference between the best fitting post- and pre-training CSFs (gray curves) measured in the trained location (the upper left), the upper right (Location change-contralateral) and the lower left (Location change-ipsilateral) visual field location in LE, and the upper left location in RE (Eye change). The enlarged symbols indicate the trained condition (spatial frequency: 5 cpd; location: the upper left; eye: LE) before (blue triangles) and after training (red circles). BW: the bandwidth of perceptual learning.

Figure 2C. The AUCSF improved by 73.78, 53.35, 45.15, and 53.87% in the four conditions, respectively. The magnitude of AUCSF improvement in the trained location was significantly or marginally larger than that in the upper right [$t(19) = 1.957$, $p = 0.065$, $d = 0.462$], the lower left in LE [$t(19) = 3.127$, $p < 0.05$, $d = 0.667$], and the upper left in RE [$t(19) = 1.987$, $p = 0.093$, $d = 0.479$] after multiple comparison correction based on FDR. There was no significant difference among the magnitudes of improvement in the three control conditions (all $p > 0.10$, all $BF_{10} < 4.20$).

The spatial frequency bandwidth of perceptual learning indicates the generalization of training effect to other stimuli and tasks, were indexed by the full bandwidth at half height of the difference curve between the post- and pre-training CSFs, was 3.62 ± 1.96, 2.45 ± 1.62, 2.29 ± 1.19, and 3.55 ± 2.40 octaves (mean ± sd) for the upper left (trained location), the upper right, the lower left location in LE (trained eye), and the upper left in RE (untrained eye), respectively. The bandwidth of perceptual learning was significantly or marginally greater in the trained condition than in the upper right [$t(19) = 2.177$, $p = 0.063$, $d = 0.643$, paired t-test] and the lower left [$t(19) = 2.877$, $p = 0.030$, $d = 0.778$] in LE but not the upper left in RE [$t(19) = 0.102$, $p = 0.920$,

d = 0.033] after multiple comparison correction based on FDR.

ERP Outcomes

Overview

The grand average of stimulus-locked ERPs was shown in **Figures 3**, **4** for the trained and control conditions respectively. In the electrodes placed on the posterior-occipital cortex, we observed P1, N1, P2, and N2 components. The timing (**Figures 3A**, **4A**) and topography (**Figure 3C**) of each ERP component were largely consistent with previous reports (Voorhis and Hillyard, 1977; Johnson, 1989; Duncan et al., 1994; Gonzalez et al., 1994; Woldorff et al., 1997; Luck et al., 2000; Vogel and Luck, 2000; Pernet et al., 2003; Potts, 2004; Key et al., 2005).

Trained Condition

Training was performed at 5 cpd in the upper left visual field location of LE. We first conducted six three-way ANOVA for the latency and amplitude of each of the following ERP component: P1, N1-P2 complex and N2, with hemisphere (left hemisphere vs. right hemisphere), training (pre-training vs. post-training), and contrasts levels (4.26, 8.90, 18.61, 38.9, and 81.13%) as within-subject factors. We found shorter latencies of P1 and N2 in the right hemisphere (contralateral) compared to the ones in the left (ipsilateral) hemisphere [$F(1,19) = 5.290$, $p = 0.033$, $\eta_p^2 = 0.218$; $F(1,19) = 144.013$, $p < 0.001$, $\eta_p^2 = 0.883$]. The amplitudes of P1, N1-P2 complex and N2 from the right hemisphere were larger than the ones in the left hemisphere [$F(1,19) = 11.704$, $p = 0.003$, $\eta_p^2 = 0.381$; $F(1,19) = 16.108$, $p < 0.001$, $\eta_p^2 = 0.997$; $F(1,19) = 22.220$, $p < 0.001$, $\eta_p^2 = 0.539$].

Our further analyses focused on the contralateral (right) hemisphere. The latency and amplitude for each ERP component were then entered into a 2-way ANOVA with training (pre-training vs. post-training) and contrast levels (4.26, 8.9, 18.61, 38.9, and 81.13%) as two within-subject factors (**Figure 3B**). We found the latency of P1, N1-P2, and N2 components decreased significantly with contrast levels [$F(4,64) = 31.723$, 133.395, and 49.570, respectively, $\eta_p^2 = 0.625$, 0.875, and 0.723, all $p < 0.001$) and training [$F(1,19) = 6.128$, 20.062, and 13.611, $\eta_p^2 = 0.244$, 0.514, 0.417, respectively, all $p < 0.05$]. The interaction of the two factors was marginally significant for the latency of N2 component [$F(4,76) = 2.729$, $p = 0.060$, $\eta_p^2 = 0.126$]. A follow-up simple effect test indicated the three higher contrast conditions reached significance for the latency of N2 component [$F(1,19) = 0.30$, $p = 0.593$, $\eta_p^2 = 0.016$; $F(1,19) = 2.52$, $p = 0.129$, $\eta_p^2 = 0.117$; $F(1,19) = 6.23$, $p = 0.022$, $\eta_p^2 = 0.247$; $F(1,19) = 7.11$, $p = 0.015$, $\eta_p^2 = 0.272$; $F(1,19) = 9.08$, $p = 0.007$, $\eta_p^2 = 0.323$ for the five contrast levels separately].

The amplitudes increased significantly with contrast levels [$F(4,76) = 21.692$, 86.585, and 42.411, $\eta_p^2 = 0.533$, 0.820, 0.691, for P1, N1-P2, and N2, respectively, all $p < 0.001$). Training also significantly increased the amplitude of P1 [$F(1,19) = 6.085$, $p = 0.023$, $\eta_p^2 = 0.243$) and N1-P2 [$F(1,19) = 16.521$, $p = 0.001$, $\eta_p^2 = 0.465$] but not N2 [$F(1,19) = 0.463$, $p = 0.505$). The interaction of the two factors was only significant for the amplitude of P1 component [$F(4,76) = 3.607$, $p = 0.019$, $\eta_p^2 = 0.160$]. A follow-up simple effect test revealed that the amplitude of P1 component was only significantly increased when the stimulus contrast was 81.13% [$F(1,19) = 21.39$, $p < 0.001$, $\eta_p^2 = 0.530$]. In sum, we observed shorter latency and increased amplitude for ERP components in response to stimuli presented at the trained location.

Control Conditions

We conducted paired t-tests (with multiple comparison correction based on FDR) for the latency and amplitude for each ERP component of the right hemisphere for Frequency change, Location change-ipsilateral, Eye change condition, and left hemisphere for Location change-contralateral condition (**Figure 4B**). For latency, training decreased P1 latency only at Frequency change condition [$t(19) = 3.303$, $p < 0.01$, $d = 1.091$] after multiple-comparison correction based on FDR. In contrast, significant or marginally significant reduction of latency was found for N1-P2 complex at Frequency change, Location change-contralateral, and Eye change-condition [$t(19) = 3.629$, $p < 0.01$, $d = 0.803$; $t(19) = 3.138$, $p = 0.01$, $d = 0.675$; $t(19) = 2.144$, $p = 0.06$, $d = 0.504$, respectively]; and for N2 component at all the four control conditions [$t(19) = 2.125$, $p = 0.062$, $d = 0.630$, $t(19) = 3.668$, $p < 0.01$, $d = 0.944$, $t(19) = 2.991$, $p < 0.05$, $d = 0.719$, $t(19) = 1.893$, $p = 0.074$, $d = 0.497$ for Frequency change, Location change-contralateral, Location change-ipsilateral, and Eye change condition, respectively] after multiple-comparison correction based on FDR.

For amplitudes, we found significant or marginally significant changes in N1-P2 complex amplitude for Frequency change condition [$t(19) = 3.881$, $p < 0.005$, $d = 0.706$] and in N2 amplitude for both Frequency change and Eye change condition (untrained) eye [$t(19) = 2.990$, $p < 0.05$, $d = 0.549$ and $t(19) = 2.239$, $p = 0.074$, $d = 0.565$] after multiple-comparison correction based on FDR.

The magnitude of improvement in N2 amplitude and latency at the trained condition was marginally larger than that in the Higher SF condition [$t(19) = 1.782$, $p = 0.091$, $d = 0.25$] and the Location change-contralateral condition [$t(19) = 1.757$, $p = 0.095$, $d = 0.18$], respectively. No significant difference of ERP changes in other control and corresponding trained conditions was found [$t(19) = 1.582$–0.082, $p = 0.130$–0.936, $BF_{10} = 4.29$–1.48].

Model Analysis

We plotted the mean latency and amplitude of the P1, N1-P2 complex and N2 components of the right hemisphere at the trained condition as functions of stimulus contrasts (i.e., CRF) and fitted with the Naka-Rushton equation (**Figure 5**; Tolhurst et al., 1981; Albrecht et al., 2002; Li X. et al., 2008).

For the latency CRF (**Figures 5A–C**), training increased the effective contrast (c_{50}) by a factor of 0.72, or a decrease of 28% of its physical contrast, in the latency of P1 [$t(19) = 2.925$, $p < 0.05$, $d = 0.624$, multiple-comparison corrected based on FDR, **Figure 5A**] and by a factor of 0.70 in the N1-P2 complex [$t(19) = 2.765$, $p < 0.05$, $d = 0.637$, multiple comparison corrected based on FDR, **Figure 5B**]; and led to a shift of the contrast gain (c_{50}) by a factor of 0.59 [$t(19) = 4.179$, $p < 0.005$, multiple-comparison correction based on FDR, $d = 0.971$] and a multiplicative response increase by a factor of 1.61 [$t(19) = 2.090$,

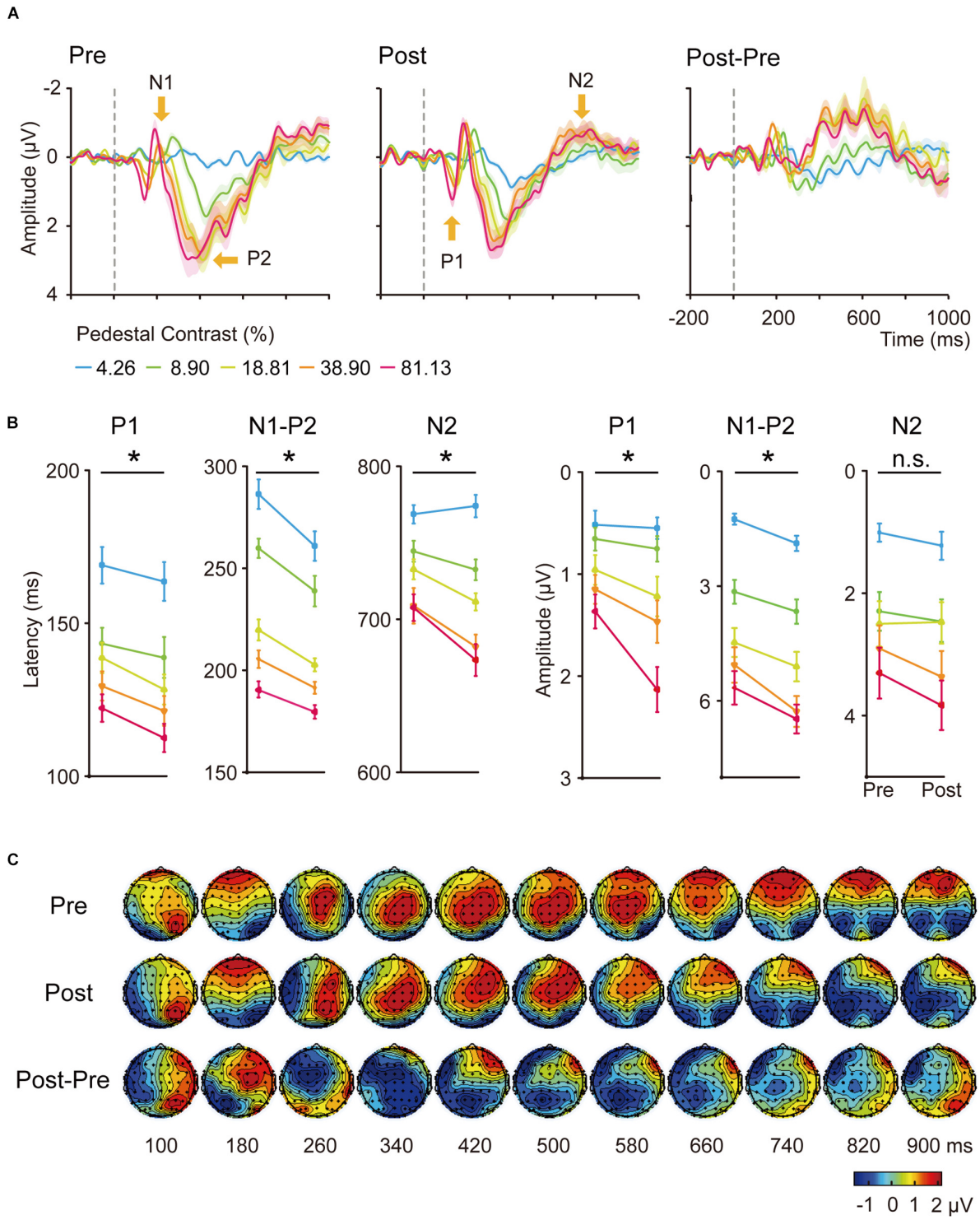


FIGURE 3 | (A) Averaged ERP waveforms of the trained condition. The ERPs evoked by contralateral stimuli of 4.26, 8.90, 18.61, 38.90, and 81.13% Michelson contrast levels were subtracted by that evoked by contralateral 0%-contrast stimuli. Significant sensory ERP components, e.g., P1, N1, P2, and N2, were identified. Shaded regions denote standard errors across subjects. **(B)** Latency and amplitudes from early to late ERP components at each contrast levels of the trained condition in pre-training and post-training sessions. Statistical analysis showed that the latency and amplitude from early to late ERP components at each contrast levels were modified differently by training. Error bars represent standard errors across subjects. *: significant main effects of training; n.s.: non-significant. **(C)** The grand-mean topographical map series from 100 to 900 ms in steps of 80 ms evoked by stimuli of 81.13% contrast level of the trained condition in pre-training (upper part) and post-training (middle part) sessions. The difference topographical maps were also displayed (lower part). Four components occurred at this time window, from P1, N1, P2, to N2.

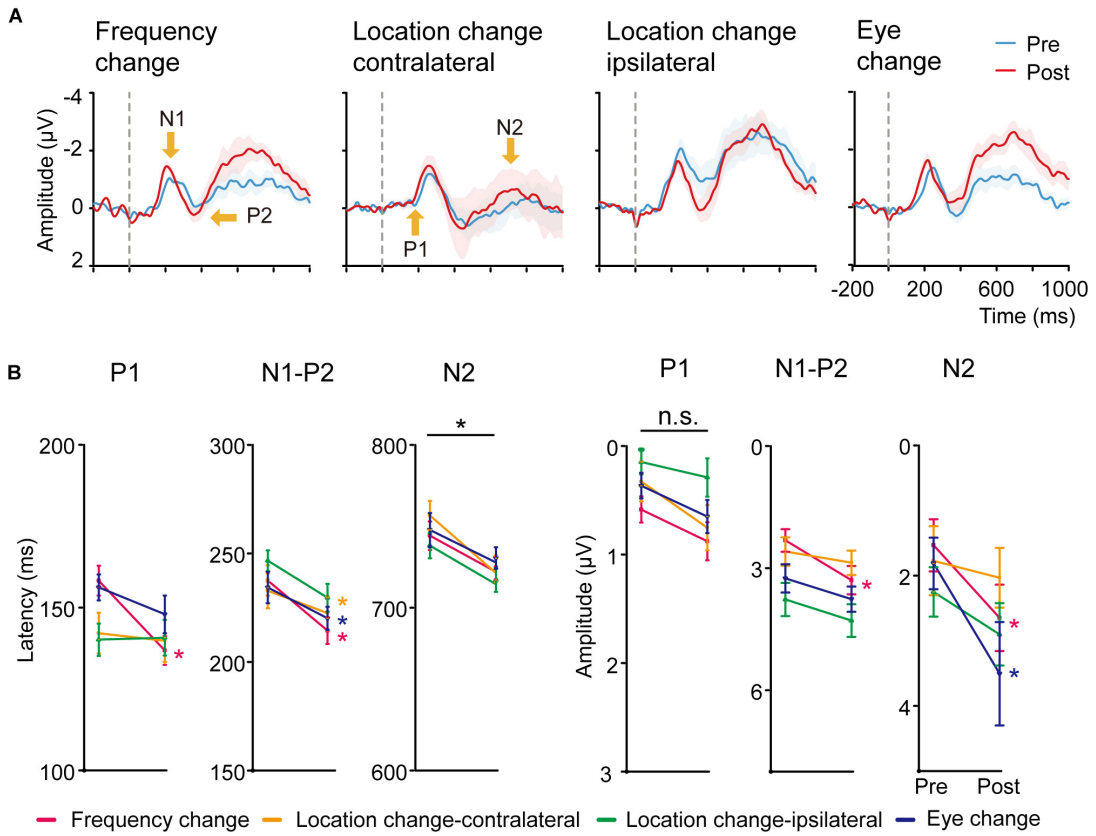


FIGURE 4 | (A) Averaged ERP responses of control conditions. From left to right: Frequency change, Location change-contralateral (the upper right visual field location in LE), Location change-ipsilateral (the lower left location in LE) and Eye change condition (the upper left location in RE). Shaded regions denote standard errors across subjects. **(B)** Latency and amplitudes from early to late ERP components of the four control conditions in pre-training and post-training sessions. Statistical analysis showed that the amplitude and latency from early to late ERP components of the four control conditions were also modified differently by training. Error bars represent standard errors across subjects.

$p = 0.076$, $d = 0.446$, multiple-comparison correction based on FDR] for N2 (**Figure 5C**).

For the amplitude CRF (**Figures 5D–F**), training led to a contrast gain (c_{50}) improvement by a factor of 0.7 [$t(19) = 2.673$, $p < 0.05$, $d = 0.472$, multiple-comparison correction based on FDR] and a multiplicative response increase by 1.72 [$t(19) = 3.713$, $p < 0.005$, $d = 0.490$, multiple-comparison correction based on FDR] in the amplitude of P1 (**Figure 5D**); and a multiplicative response increase by a factor of 2.29 [$t(19) = 2.160$, $p = 0.066$, $d = 0.381$, multiple-comparison correction based on FDR] and baseline shift by a factor of 1.26 [$t(19) = 2.576$, $p = 0.057$, $d = 0.517$, multiple-comparison correction based on FDR] for the N1-P2 complex (**Figure 5E**). These results further showed that perceptual learning impacted neural processing differently across neural events at the trained condition.

DISCUSSION

In the present study, we tested the multi-stage hypothesis of perceptual learning. Behavioral results showed that training substantially improved visual acuity and CSFs, with the learning effect being particularly pronounced at the trained condition and partially transferred to control conditions. ERP results showed that training reduced the latency and increased the amplitudes on both early and late components for the trained condition. Further modeling analysis revealed a contrast-gain-related change in the latency of P1, N1-P2 complex, and N2, as well as response-gain-related changes in the latency of N2. Finally, for the untrained conditions, P1 showed reduced latency only at the high spatial frequency condition while N2 showed decreased latency for all control conditions.

The specificity of VPL has been the hallmark of perceptual learning and is often regarded as the evidence of a singular low-level process. In support of this hypothesis, fMRI studies revealed increased responses in the early retinotopic visual areas (Schwartz et al., 2003; Furmanski et al., 2004; Jehee et al., 2012). These results were further substantiated by EEG recordings showing post-training improvements in early visually evoked components over occipital electrode sites (Pourtois et al., 2008; Censor et al., 2009; Bao et al., 2010) and electrophysiological recordings in non-human primates linking behavioral performance with improvements in neuronal

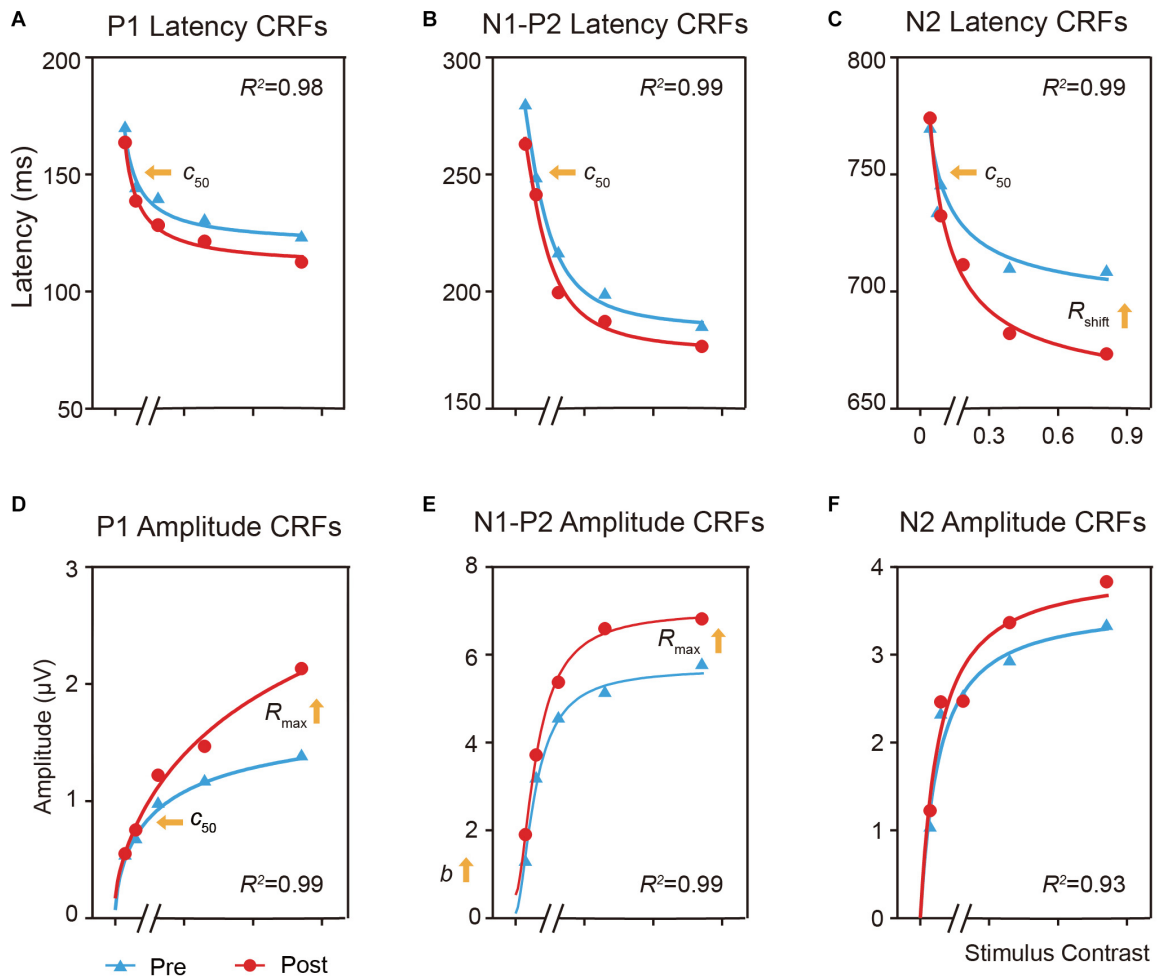


FIGURE 5 | Effects of perceptual learning on the mean latency and amplitude of the P1, N1-P2 complex, and N2 components in the trained condition as function of contrasts (i.e., ERP-dependent CRF). **(A,D)** CRFs for P1 latency and amplitude. **(B,E)** CRFs for N1-P2 complex latency and amplitude. **(C,F)** CRFs for N2 latency and amplitude. For the latency CRF **(A–C)**, training lead to c_{50} improvement for both the P1 and N1-P2 complex, and c_{50} and response increase for N2. For the amplitude CRFs **(D–F)**, training led to c_{50} improvement and multiplicative response increase for P1, and multiplicative response and baseline increase for N1-P2 complex.

sensitivity in primary sensory areas (Ghose et al., 2002; Hua et al., 2010; Yan et al., 2014). In the current study, we observed contrast-dependent gain change both in the latency and amplitude of early P1 component, which resembles a previous single-unit study that recorded the responses of V1 neurons in cats and found that training increased neuronal contrast gain (Hua et al., 2010). P1 is a visually evoked exogenous response that reflected the encoding of sensory information in visual cortex (Voorhis and Hillyard, 1977; Gonzalez et al., 1994; Woldorff et al., 1997; O'Shea et al., 2010; Souza et al., 2013). Moreover, we found there is little improvement in the latency and amplitude of P1 at the location and eye change control conditions, which is also indicative of learning specificity. These findings were confirmed by the behavioral result of CSF measurements, i.e. magnitude of AUCSF improvement in the training location was larger than untrained conditions.

In contrast, there are also studies proposed that learning could be explained by selective reweighting of sensory information readout (Dorsher and Lu, 1998, 1999; Petrov et al., 2005; Liu et al., 2010), changes in attention and/or decision-making areas (Xiao et al., 2008; Zhang et al., 2010, 2013; Wang et al., 2012), or changes in both the sensory coding and the communication between the visual and the decision-making related areas (Chen et al., 2015, 2017). Interestingly, in our study, model-based analysis also revealed mechanisms of both response and baseline improvements at later stages. N1 reflects selective attention to basic stimulus characteristics and intentional discrimination processing (Näätänen et al., 1982; Luck et al., 2000; Vogel and Luck, 2000); P2 may reflect stimulus classification, and its amplitude increases with the stimulus complexity (Näätänen et al., 1982; Pernet et al., 2003; Crowley and Colrain, 2004; Potts, 2004); Late negative N2 has been associated with high-level decision-related processing and task demands (Johnson, 1989;

Duncan et al., 1994; Mangun and Hillyard, 1995; Key et al., 2005). The response increment clearly indicates stimulus-independent effects of training on the latency of N1-P2 complex and N2 (Luck et al., 2000; Vogel and Luck, 2000; Pernet et al., 2003; Potts, 2004). The additive shift on the amplitude of N1-P2 complex ensures the effect that post-training responses will be higher than pre-training responses. These stimulus-independent improvements might reflect a top-down effect of training, such as decision and attention modulation on later stimulus processing stages. Also, we found significant improvement in the latency of N1-P2 complex and N2 at the spatial frequency, location, and eye transfer conditions. These results indicate unequal training effects across neuronal processing stages and the extent to which training transfers may depend on the specific stage of information processing. Using a motor training paradigm, Garner et al. (2015) also found transferability of training benefits was different across ERP components, i.e., N2 showed increased amplitudes and reduced latencies for both trained and untrained stimuli, while the onset of stimulus-locked lateralized readiness potential reduced only for the trained stimuli.

The multistage model could explain the existing divergent findings in perceptual learning (Sasaki et al., 2013; Shibata et al., 2016; Maniglia and Seitz, 2018). In accord with the multi-stage model, recent work in non-human primates found that V4 and the posterior inferior temporal (PIT) cortex both changed after training on an orientation discrimination task (Adab et al., 2014). Chen et al. (2015) has reported that training of a motion direction discrimination task is associated with changes in both V3A and connectivity between V3A and IPS. After training on a motion detection task, Shibata et al. (2016) found that the response changes in V3A were specific to the trained direction, independent of whether subjects performed the training task actively or only passively exposed to the stimuli, and significant response changes in V1 and the intraparietal sulcus (IPS) were found only when subjects performed the trained task on the trained motion stimulus, providing direct evidence for their two-plasticity model. Similarly, our results indicate different forms of experience-dependent plasticity: contrast-gain change in early P1 component and response gain/baseline increments in later N1-P2 complex and N2 component. Earlier ERP components might be more related to the physical properties (e.g., contrast) of stimuli which reflects the change of early sensory/feature processing stages, while later components were presumably modulated by top-down signals, which reflect the improvement in higher-level processing stages (Voorhis and Hillyard, 1977; Johnson, 1989; Duncan et al., 1994; Gonzalez et al., 1994; Woldorff et al., 1997; Luck et al., 2000; Vogel and Luck, 2000; Pernet et al., 2003; Potts, 2004; Key et al., 2005).

A recent ERP study also found significant changes in both early and late ERP components following training on a texture discrimination task (TDT) (Ahmadi et al., 2018). Specifically, they found a decrease in the C1 but not P1 amplitude, a decrease in both N1 amplitude and latency, and a significant increase in the P3 amplitude after training. In the current study, we found contrast detection training reduced the latency and increased the amplitude in both early and late ERP components, with different characteristics of contrast dependence and different underlying mechanisms explained within a quantitative modeling framework based on CRF measurements. The discrepancy was likely due to different training tasks and procedures used in the two studies. For example, Ahmadi et al. (2018) recorded ERPs while subjects performed the TDT at Session 1 and Session 2, with two full nights of sleep between sessions. Here we trained subjects on a contrast detection task for ten consecutive days and recorded ERPs during pre- and post-training tests.

We didn't observe any significant C1 in our subjects. C1 is the earliest visual ERP component and is thought to be generated by neurons in V1 (Foxe and Simpson, 2002; Russo et al., 2003). There are three possibilities: (1) C1 is more vulnerable and difficult to be identified because of the specific orientation and folding of calcarine sulci of individual subjects (Kelly et al., 2008). (2) In order to measure ERP-based CRF, a high proportion of low contrast stimuli was involved in the current study (e.g., 0, 4.26, 8.90, 18.61, and 38.90% Michelson contrasts), which might not be able to elicit the subtle C1 effects or were overlaid by large individual differences in the functional anatomy of early visual cortex (Dougherty et al., 2003; Pourtois et al., 2008). (3) It usually needs more trials to isolate C1. We averaged over 200 trials, less than the previous studies (Ludwig and Skrandies, 2002; Russo et al., 2003; Bao et al., 2010). For example, Zhang et al. (2015) trained subjects with a similar peripheral (5° retinal eccentricity) grating orientation discrimination task and successfully isolated stimulus-related C1 epochs with a total of 450 $\pm$ 65 trials for each condition.

Although a large number of studies failed to detect latency change (Song et al., 2002, 2005; Pourtois et al., 2008; Bao et al., 2010; Qu et al., 2010; Wang et al., 2010; Hamamé et al., 2011; An et al., 2012; Zhang et al., 2013, 2015) and claimed that ERP amplitude instead of latency was more sensitive to training (Qu et al., 2010; An et al., 2012), a few studies reported training-induced ERP latency change (Skrandies and Fahle, 1994; Ludwig and Skrandies, 2002; Shoji and Skrandies, 2006; Garner et al., 2015; Diaz et al., 2017). In the current study, we found that the latency of both early and late components was shortened after training while increased amplitudes were seen in P1 and N1-P2 complex but not N2 component. The decrements of ERP latency might reflect improved efficiency of visual transmission from the lateral geniculate to higher cortical areas. In line with this claim, Mukai et al. (2007) found BOLD responses in putative attention-control areas reduced but the functional connectivity between frontoparietal areas and early visual cortex increased after training, indicative of improved processing efficiency following training. Note that we didn't find changes in the amplitude of N2 components, which might be due to response saturation at this later stage.

In the current study, we found a mild improvement in visual acuity (e.g., 1.0 line in the trained eye and 0.4 lines in the untrained eye) following training on contrast detection. Visual acuity is usually thought to reflect the frequency limits of the visual system but the task, in fact, depends on a range of spatial frequencies, including low-to-medium spatial frequencies (Huang et al., 2007). Improvement of contrast sensitivity will likely benefit visual acuity, as evident in early studies with

normal subjects (e.g., Zhou et al., 2007) as well as suffered population (e.g., Polat et al., 2004; Huang et al., 2008; Yan et al., 2015). Previous psychophysical studies have found that perceptual learning of contrast detection might decrease internal noise and/or finely tune perceptual template (Huang et al., 2009), with related brain area possibly down to LGN (Yu et al., 2016). In the current study, we found significant ERP changes in both early and late ERP components, which may reflect neuronal changes in both the representation stage and attentional processing (Voorhis and Hillyard, 1977; Luck et al., 2000; Fabiani et al., 2007). Relations among different studies that involved varied technological measures remains to be elucidated. One limitation of the current study is the lack of a control group that took pre- and post-training assessments (without training), which might weaken the interpretation of visual acuity improvement following training, although our focus was the improvement in contrast sensitivity and associated early and late ERP changes at the trained location following training and within-subject comparison between relative changes in trained and untrained conditions.

We also observed significant improvement in contrast sensitivity at the upper right, the lower left location in LE (trained eye), and the upper left location in RE (untrained eye). Our results were in general consistent with previous findings (Sowden et al., 2002; Yu et al., 2004; Casco et al., 2014), although there were differences in experimental settings. For example, in order to elicit a more reliable ERP response, we used a training frequency of 5 cpd, which is much lower than that in earlier studies (e.g., Zhou et al., 2007; Huang et al., 2008; Wu et al., 2020). Some have indicated greater improvement magnitude and transfer of perceptual learning was related to higher spatial frequencies (Wu et al., 2020). Another interesting finding is that training based on lateral masking could be more effective than protocols based on isolated Gabor stimuli to compensate for myopic vision (Camilleri et al., 2014). Future studies are needed to investigate whether a paradigm with higher spatial frequency training or lateral masking would result in better learning effects.

Taken together, our findings indicate that visual perceptual training leads to changes across different visual processing stages and the extent of learning and transfers may depend on the specific stage of information processing. Perceptual learning has been considered to be effective in improving deficient vision in clinical populations, e.g., amblyopia (Polat et al., 2004; Zhou et al., 2006), myopia (Durrie and McMinn, 2007; Yan et al., 2015), and presbyopia (Polat, 2009; DeLoss et al., 2015). On the other hand, many visual diseases demonstrated decreased amplitude and/or increased delay in both early and late ERP components (Levi and Harwerth, 1978; Sokol, 1983; Hess et al., 1985; Sengpiel and Blakemore, 1996; Koertvelyes et al., 2012). The current study, together with others (Skrandies and Fahle, 1994; Song et al., 2005; Shoji and Skrandies, 2006; Pourtois et al., 2008; Bao et al., 2010; Qu et al., 2010; Wang et al., 2010; Hamamé et al., 2011; Zhang et al., 2013, 2015), provided a more integrated way to understand visual rehabilitation and a potential method to modulate the efficacy of visual training (e.g., neuro-feedback, Saxby and Peniston, 1995; Hanslmayr et al., 2005; Vernon, 2005; Shibata et al., 2011; Zoefel et al., 2011). Another interesting open question is whether changes in both early and late ERP components happen concurrently or sequentially with training. Future studies should track brain activities during the course of training to give a full theoretical framework for understanding visual perceptual learning.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Review Committee of Institute of Psychology, Chinese Academy of Sciences. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JX, YZ, and C-BH designed the experiment. JX, PZ, and W-LJ collected the data. JX, JY, NC, G-TW, and C-BH conducted the analyses. JX, G-TW, YD, YZ, and C-BH wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China (2020YFC2003800), the Scientific Instrument Developing Project of the Chinese Academy of Sciences (ZDKYYQ20200005), and the National Natural Science Foundation of China grants (NSFC 32071056 to C-BH, NSFC 31400877 to JX, and NSFC 31971031 and 31930053 to NC).

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (2020YFC2003800), the Scientific Instrument Developing Project of the Chinese Academy of Sciences (ZDKYYQ20200005), and the National Natural Science Foundation of China grants (NSFC 32071056 to C-BH, NSFC 31400877 to JX, and NSFC 31971031 and 31930053 to NC).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2020.555701/full#supplementary-material

REFERENCES

Adab, H. Z., Popivanov, I. D., and Vanduffel, W. (2014). Perceptual learning of simple stimuli modifies stimulus representations in posterior inferior temporal cortex. *J. Cogn. Neurosci.* 26, 2187–2200. doi: 10.1162/jocn_a_00641

Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci.* 8, 457–464. doi: 10.1016/j.tics.2004.08.011

Ahmadi, M., McDevitt, E. A., Silver, M. A., and Mednick, S. C. (2018). Perceptual learning induces changes in early and late visual evoked potentials. *Vis. Res.* 152, 101–109. doi: 10.1016/j.visres.2017.08.008

Albrecht, D. G., Geisler, W. S., Frazor, R. A., and Crane, A. M. (2002). Visual cortex neurons of monkeys and cats: temporal dynamics of the contrast response function. *J. Neurophysiol.* 88, 888–913. doi: 10.1152/jn.2002.88.2.888

An, A., Sun, M., Wang, Y., Wang, F., Ding, Y., and Song, Y. (2012). The N2pc is increased by perceptual learning but is unnecessary for the transfer of learning. *PLoS One* 7:e34826. doi: 10.1371/journal.pone.0034826

Bao, M., Yang, L., Rios, C., He, B., and Engel, S. A. (2010). Perceptual learning increases the strength of the earliest signals in visual cortex. *J. Neurosci.* 30, 15080–15084. doi: 10.1523/jneurosci.5703-09.2010

Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897x00357

Camilleri, R., Pavan, A., Ghin, F., and Campana, G. (2014). Improving myopia via perceptual learning: is training with lateral masking the only (or the most) efficacious technique? *Attent. Percept. Psychophys.* 76, 2485–2494. doi: 10.3758/s13414-014-0738-8

Casco, C., Guzzon, D., Moise, M., Vecchies, A., Testa, T., and Pavan, A. (2014). Specificity and generalization of perceptual learning in low myopia. *Restorat. Neurol. Neurosci.* 32, 639–653. doi: 10.3233/rnn-140389

Censor, N., Bonneh, Y., Arieli, A., and Sagi, D. (2009). Early-vision brain responses which predict human visual segmentation and learning. *J. Vis.* 9, 1–9.

Chen, N., Bi, T., Zhou, T., Li, S., Liu, Z., and Fang, F. (2015). Sharpened cortical tuning and enhanced cortico-cortical communication contribute to the long-term neural mechanisms of visual motion perceptual learning. *Neuroimage* 115, 17–29. doi: 10.1016/j.neuroimage.2015.04.041

Chen, N., and Fang, F. (2011). Tilt aftereffect from orientation discrimination learning. *Exp. Brain Res.* 215, 227–234. doi: 10.1007/s00221-011-2895-5

Chen, N., Lu, J., Shao, H., Weng, X., and Fang, F. (2017). Neural mechanisms of motion perceptual learning in noise. *Hum. Brain Mapp.* 38, 6029–6042. doi: 10.1002/hbm.23808

Crist, R. E., Kapadia, M. K., Westheimer, G., Gilbert, C. D., Mcgovern, D. P., Webb, B. S., et al. (2014). Perceptual learning of spatial localization: specificity for orientation, position, and context. *J. Neurophysiol.* 78, 2889–2894. doi: 10.1152/jn.1997.78.6.2889

Crowley, K. E., and Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin. Neurophysiol.* 115, 732–744. doi: 10.1016/j.clinph.2003.11.021

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

DeLoss, D. J., Watanabe, T., and Andersen, G. J. (2015). Improving vision among older adults: behavioral training to improve sight. *Psychol. Sci.* 26, 456–466. doi: 10.1177/0956797614567510

Deveau, J., Lovcik, G., and Seitz, A. R. (2013). The therapeutic benefits of perceptual learning. *Curr. Trends Neurol.* 7, 39–49.

Diaz, J. A., Queirazza, F., and Philiastides, M. G. (2017). Perceptual learning alters post-sensory processing in human decision-making. *Nat. Hum. Behav.* 1, 1–9.

Ding, Y. L., Song, Y., Fan, S., Qu, Z., and Chen, L. (2003). Specificity and generalization of visual perceptual learning in humans; an event-related potential study. *Neuroreport* 14, 587–590. doi: 10.1097/00001756-200303240-00012

Dorsher, B. A., and Lu, Z. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci. U.S.A.* 95, 13988–13993. doi: 10.1073/pnas.95.23.13988

Dorsher, B. A., and Lu, Z. L. (1999). Mechanisms of perceptual learning. *Vis. Res.* 39, 3197–3221.

Dosher, B. A., Jeter, P., Liu, J., and Lu, Z. L. (2013). An integrated reweighting theory of perceptual learning. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13678–13683.

Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., and Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *J. Vis.* 3, 586–598.

Duncan, C. C., Rumsey, J. M., Wilkniss, S. M., Denckla, M. B., Hamburger, S. D., and Odou-Potkin, M. (1994). Developmental dyslexia and attention dysfunction in adults: brain potential indices of information processing. *Psychophysiology* 31, 386–401. doi: 10.1111/j.1469-8986.1994.tb02447.x

Durrie, D., and McMinn, P. S. (2007). Computer-based primary visual cortex training for treatment of lowmyopia and early presbyopia. *Trans. Am. Ophthalmol. Soc.* 105, 132–140.

Fabiani, M., Gratton, G., and Federmeier, K. D. (2007). "Event-related brain potentials: methods, theory, and applications," in *Handbook of Psychophysiology*, eds J. T. Cacioppo, G. Berntson, and L. G. Tassinary (Cambridge: Cambridge University Press), 53–84.

Foxe, J. J., and Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans - A framework for defining "early" visual processing. *Exp. Brain Res.* 142, 139–150. doi: 10.1007/s00221-001-0906-7

Friston, K. (2003). Learning and inference in the brain. *Neural Netw.* 16, 1325–1352. doi: 10.1016/j.neunet.2003.06.005

Furmanski, C. S., Schluppeck, D., Engel, S. A., and Angeles, L. (2004). Learning strengthens the response of primary visual cortex to simple patterns. *Curr. Biol.* 14, 573–578. doi: 10.1016/j.cub.2004.03.032

Garner, K. G., Matthews, N., Remington, R. W., and Dux, P. E. (2015). Transferability of training benefits differs across neural events: evidence from ERPs. *J. Cogn. Neurosci.* 27, 2079–2094. doi: 10.1162/jocn_a_00833

Ghose, G. M., Yang, T., and Maunsell, J. H. R. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. *J. Neurophysiol.* 87, 1867–1888. doi: 10.1152/jn.00690.2001

Gilbert, C. D. (1994). Early perceptual learning. *Proc. Natl. Acad. Sci. U.S.A.* 91, 1195–1197.

Gonzalez, C. M. G., Clark, V. P., Fan, S., Luck, S. J., and Hillyard, S. A. (1994). Sources of attention-sensitive visual event-related potentials. *Brain Topogr.* 7, 41–51. doi: 10.1007/bf01184836

Hamamé, C. M., Cosmelli, D., Henriquez, R., and Aboitiz, F. (2011). Neural mechanisms of human perceptual learning: electrophysiological evidence for a two-stage process. *PLoS One* 6:e0019221. doi: 10.1371/journal.pone.0019221

Hanslmayr, S., Sauseng, P., Doppelmayr, M., Schabus, M., and Klimesch, W. (2005). Increasing individual upper alpha power by neurofeedback improves cognitive performance in human subjects. *Appl. Psychophysiol. Biofeedback* 30, 1–10. doi: 10.1007/s10484-005-2169-8

Hess, R. F., Baker, C. L., Verhoeve, J. N., Keesey, U. T., and France, T. D. (1985). The pattern evoked electroretinogram - its variability in normals and its relationship to amblyopia. *Investig. Ophthalmol. Vis. Sci.* 26, 1610–1623.

Hua, T., Bao, P., Huang, C. B., Wang, Z., Xu, J., Zhou, Y., et al. (2010). Perceptual learning improves contrast sensitivity of V1 neurons in cats. *Curr. Biol.* 20, 887–894. doi: 10.1016/j.cub.2010.03.066

Huang, C., Lu, Z., and Zhou, Y. (2009). Mechanisms underlying perceptual learning of contrast detection in adults with anisometropic amblyopia. *J. Vis.* 9, 1–14.

Huang, C., Tao, L., Zhou, Y., and Lu, Z. L. (2007). Treated amblyopes remain deficient in spatial vision: a contrast sensitivity and external noise study. *Vis. Res.* 47, 22–34. doi: 10.1016/j.visres.2006.09.015

Huang, C. B., Zhou, Y. F., and Lu, Z. L. (2008). Broad bandwidth of perceptual learning in the visual system of adults with anisometropic amblyopia. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4068–4073. doi: 10.1073/pnas.0800824105

Itthipuripat, S., Cha, K., Byers, A., and Serences, J. T. (2017). Two different mechanisms support selective attention at different phases of training. *PLoS Biol.* 15:e2001724. doi: 10.1371/journal.pbio.2001724

Itthipuripat, S., Ester, E. F., Deering, S., and Serences, J. T. (2014). Sensory gain outperforms efficient readout mechanisms in predicting attention-related improvements in behavior. *J. Neurosci.* 34, 13384–13398. doi: 10.1523/jneurosci.2277-14.2014

Jehee, J. F. M., Ling, S., Swisher, J. D., van Bergen, R. S., and Tong, F. (2012). Perceptual learning selectively refines orientation representations in early visual cortex. *J. Neurosci.* 32, 16747–16753. doi: 10.1523/jneurosci.6112-11.2012

Johnson, R. (1989). Developmental evidence for modality-dependent p300 generators: a normative study. *Psychophysiology* 26, 651–667. doi: 10.1111/j.1469-8986.1989.tb03167.x

Karni, A., and Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 88, 4966–4970. doi: 10.1073/pnas.88.11.4966

Kelly, S. P., Gomez-Ramirez, M., and Foxe, J. J. (2008). Spatial attention modulates initial afferent activity in human primary visual cortex. *Cereb. Cortex* 18, 2629–2636. doi: 10.1093/cercor/bhn022

Key, A. P., Dove, G. O., and Maguire, M. J. (2005). Linking brainwaves to the brain: an ERP primer. *Dev. Neuropsychol.* 27, 183–215. doi: 10.1207/s15326942dn2702_1

Koertvelyes, J., Banko, E. M., Andics, A., Rudas, G., Nemeth, J., Hermann, P., et al. (2012). Visual cortical responses to the input from the amblyopic eye are suppressed during binocular viewing. *Acta Biol. Hung.* 63, 65–79. doi: 10.1556/abiol.63.2012.suppl.1.7

Law, C., and Gold, J. I. (2008). Not a sensory cortical area. *Nat. Neurosci.* 11, 505–513. doi: 10.1038/nn2070

Lesmes, L. A., Lu, Z., and Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J. Vis.* 10, 1–21. doi: 10.1167/10.3.17

Levi, D. M., and Harwerth, R. S. (1978). Contrast evoked-potentials in strabismic and anisometropic amblyopia. *Investig. Ophthalmol. Vis. Sci.* 17, 571–575.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375

Li, W., Piëch, V., and Gilbert, C. D. (2008). Learning to link visual contours. *Neuron* 57, 442–451. doi: 10.1016/j.neuron.2007.12.011

Li, X., Lu, Z. L., Tjan, B. S., Dosher, B. A., and Chu, W. (2008). Blood oxygenation level-dependent contrast response functions identify mechanisms of covert attention in early visual areas. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6202–6207. doi: 10.1073/pnas.0801390105

Li, X., Lu, Z. L., Xu, P., Jin, J., and Zhou, Y. (2003). Generating high gray-level resolution monochrome displays with conventional computer graphics cards and color monitors. *J. Neurosci. Methods* 130, 9–18. doi: 10.1016/s0165-0270(03)00174-2

Liu, J., Lu, Z., and Dosher, B. A. (2010). Augmented Hebbian reweighting: interactions between feedback and training accuracy in perceptual learning. *J. Vis.* 10, 1–14.

Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14085–14087. doi: 10.1073/pnas.96.24.14085

Liu, Z., and Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vis. Res.* 40, 97–109. doi: 10.1016/s0042-6989(99)00140-6

Lopez-Calderon, J., and Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* 8:213. doi: 10.3389/fnhum.2014.00213

Luck, S. J., Woodman, G. F., and Vogel, E. K. (2000). Event-related potential studies of attention. *Trends Cogn. Sci.* 4, 432–440.

Ludwig, I., and Skrandies, W. (2002). Human perceptual learning in the peripheral visual field: sensory thresholds and neurophysiological correlates. *Biol. Psychol.* 59, 187–206. doi: 10.1016/s0301-0511(02)00009-1

Maloney, L. T. (1990). Confidence intervals for the parameters of psychometric functions. *Percept. Psychophys.* 47, 127–134. doi: 10.3758/bf03205977

Mangun, G., and Hillyard, S. (1995). "Mechanisms and models of selective attention," in *Oxford psychology series, No. 25. Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*, eds M. D. Rugg and M. G. H. Coles (Oxford: Oxford University Press), 40–86. doi: 10.1093/acprof:oso/9780198524168.003.0003

Maniglia, M., and Seitz, A. R. (2018). Towards a whole brain model of Perceptual Learning. *Curr. Opin. Behav. Sci.* 20, 47–55. doi: 10.1016/j.cobeha.2017.10.004

Mollon, J. D., and Danilova, M. V. (1996). Three remarks on perceptual learning. *Spat Vis.* 10, 51–58. doi: 10.1163/156856896x00051

Mou, T. (1966). Logarithmic visual acuity chart and five-score recording. *Chinese J. Ophthalmol.* 13, 96–106.

Mukai, I., Kim, D., Fukunaga, M., Japee, S., Marrett, S., and Ungerleider, L. G. (2007). Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. *J. Neurosci. Off. J. Soc. Neurosci.* 27, 11401–11411. doi: 10.1523/jneurosci.3002-07.2007

Näätänen, R., Simpson, M., and Loveless, N. (1982). Stimulus deviance and evoked potentials. *Biol. Psychol.* 14, 53–98. doi: 10.1016/0301-0511(82)90017-5

O'Shea, R. P., Roeber, U., and Bach, M. (2010). "Evoked potential: vision," in *Encyclopedia of Perception (pp. 399-400, xli [Color Plate])*, ed. E. B. Goldstein (Thousand Oaks, CA: Sage Publications, Inc).

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897x00366

Pernet, C., Basan, S., Doyon, B., Cardebat, D., Démonet, J. F., and Celsis, P. (2003). Neural timing of visual implicit categorization. *Cogn. Brain Res.* 17, 327–338. doi: 10.1016/s0926-6410(03)00134-4

Petrov, A. A., Dosher, B. A., and Lu, Z. L. (2005). The dynamics of perceptual learning: an incremental reweighting model. *Psychol. Rev.* 112, 715–743. doi: 10.1037/0033-295x.112.4.715

Poggio, T., Fahle, M., and Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science* 256, 1018–1021. doi: 10.1126/science.1589770

Polat, U. (2009). Making perceptual learning practical to improve visual functions. *Vis. Res.* 49, 2566–2573.

Polat, U., Ma-Naim, T., Belkin, M., and Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6692–6697.

Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain Cogn.* 56, 5–13. doi: 10.1016/j.bandc.2004.03.006

Pourtois, G., Rauss, K. S., Vuilleumier, P., and Schwartz, S. (2008). Effects of perceptual learning on primary visual cortex activity in humans. *Vis. Res.* 48, 55–62. doi: 10.1016/j.visres.2007.10.027

Qu, Z., Song, Y., and Ding, Y. (2010). ERP evidence for distinct mechanisms of fast and slow visual perceptual learning. *Neuropsychologia* 48, 1869–1874. doi: 10.1016/j.neuropsychologia.2010.01.008

Russo, F., Martinez, A., and Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cereb. Cortex* 13, 486–499. doi: 10.1093/cercor/13.5.486

Sagi, D. (2011). Perceptual learning in vision research. *Vis. Res.* 51, 1552–1566. doi: 10.1016/j.visres.2010.10.019

Sasaki, Y., Náñez, J. E., and Watanabe, T. (2013). Recent progress in perceptual learning research. *Wiley Interdiscip. Rev. Cogn. Sci.* 3, 34–37.

Saxby, E., and Peniston, E. G. (1995). Alpha-theta brainwave neurofeedback training: an effective treatment for male and female alcoholics with depressive symptoms. *J. Clin. Psychol.* 51, 685–693. doi: 10.1002/1097-4679(199509)51:5<685::aid-jclp2270510514>3.0.co;2-k

Schoups, A. A., Vogels, R., and Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *J. Physiol.* 483, 797–810. doi: 10.1113/jphysiol.1995.sp020623

Schwartz, S., Maquet, P., and Frith, C. (2003). Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. *Proc. Natl. Acad. Sci. U.S.A.* 100, 17137–17142. doi: 10.1073/pnas.242414599

Sengpiel, F., and Blakemore, C. (1996). The neural basis of suppression and amblyopia in strabismus. *Eye* 10, 250–258. doi: 10.1038/eye.1996.54

Shibata, K., Sagi, D., and Watanabe, T. (2014). Two-stage model in perceptual learning: toward a unified theory. *Ann. N.Y. Acad. Sci.* 1316, 18–28. doi: 10.1111/nyas.12419

Shibata, K., Sasaki, Y., Kawato, M., and Watanabe, T. (2016). Neuroimaging evidence for 2 types of plasticity in association with visual perceptual learning. *Cereb. Cortex* 26, 3681–3689. doi: 10.1093/cercor/bhw176

Shibata, K., Watanabe, T., Sasaki, Y., and Kawato, M. (2011). Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *Science* 334, 1413–1415. doi: 10.1126/science.1212003

Shoji, H., and Skrandies, W. (2006). ERP topography and human perceptual learning in the peripheral visual field. *Int. J. Psychophysiol.* 61, 179–187. doi: 10.1016/j.ijpsycho.2005.09.007

Skrandies, W., and Fahle, M. (1994). Neurophysiological correlates of perceptual learning in the human brain. *Brain Topogr.* 7, 163–168. doi: 10.1007/bf01186774

Sokol, S. (1983). Abnormal evoked-potential latencies in amblyopia. *Br. J. Ophthalmol.* 67, 310–314. doi: 10.1136/bjo.67.5.310

Song, Y., Ding, Y., Fan, S., and Chen, L. (2002). An event-related potential study on visual perceptual learning under short-term and long-term training conditions. *Neuroreport* 13, 2053–2057. doi: 10.1097/00001756-200211150-00013

Song, Y., Ding, Y., Fan, S., Qu, Z., Xu, L., Lu, C., et al. (2005). Neural substrates of visual perceptual learning of simple and complex stimuli. *Clin. Neurophysiol.* 116, 632–639. doi: 10.1016/j.clinph.2004.09.019

Souza, G. S., Gomes, B. D., Lacerda, E. M. C. B., Saito, C. A., Silva Filho, M. D., and Silveira, L. C. L. (2013). Contrast sensitivity of pattern transient VEP components: contribution from M and P pathways. *Psychol. Neurosci.* 6, 191–198. doi: 10.3922/j.psns.2013.2.07

Sowden, P. T., Rose, D., and Davies, I. R. L. (2002). Perceptual learning of luminance contrast detection: specific for spatial frequency and retinal location but not orientation. *Vis. Res.* 42, 1249–1258. doi: 10.1016/s0042-6989(02)00019-6

Tolhurst, D. J., Movshon, J. A., and Thompson, I. D. (1981). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Exp. Brain Res.* 41, 414–419.

van Gaalen, K. W., Jansonius, N. M., Koopmans, S. A., Terwee, T., and Kooijman, A. C. (2009). Relationship between contrast sensitivity and spherical aberration: comparison of 7 contrast sensitivity tests with natural and artificial pupils in healthy eyes. *J. Cataract. Refract. Surg.* 35, 47–56. doi: 10.1016/j.jcrs.2008.09.016

Vernon, D. J. (2005). Can neurofeedback training enhance performance? An evaluation of the evidence with implications for future research. *Appl. Psychophysiol. Biofeedback* 30, 347–364. doi: 10.1007/s10484-005-8421-4

Vogel, E. K., and Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology* 37, 190–203. doi: 10.1111/1469-8986.3720190

Voorhis, S., and Hillyard, S. A. (1977). Visual evoked potentials and selective attention to points in space. *Attent. Percept. Psychophys.* 22, 54–62. doi: 10.3758/bf03206080

Wang, R., Zhang, J. Y., Klein, S. A., Levi, D. M., and Yu, C. (2012). Task relevancy and demand modulate double-training enabled transfer of perceptual learning. *Vis. Res.* 61, 33–38. doi: 10.1016/j.visres.2011.07.019

Wang, Y., Song, Y., Qu, Z., and Ding, Y. (2010). Task difficulty modulates electrophysiological correlates of perceptual learning. *Int. J. Psychophysiol.* 75, 234–240. doi: 10.1016/j.ijpsycho.2009.11.006

Watanabe, T., Náñez, J. E., Koyama, S., Mukai, I., Liederman, J., and Sasaki, Y. (2002). Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nat. Neurosci.* 5, 1003–1009. doi: 10.1038/nn915

Watanabe, T., and Sasaki, Y. (2015). Perceptual learning: toward a comprehensive theory. *Annu. Rev. Psychol.* 66, 197–221. doi: 10.1146/annurev-psych-010814-015214

Woldorff, M. G., Fox, P. T., Matzke, M., Lancaster, J. L., Veeraswamy, S., Zamarripa, F., et al. (1997). Retinotopic organization of early visual spatial attention effects as revealed by PET and ERPs. *Hum. Brain Map.* 5, 280–286. doi: 10.1002/(sici)1097-0193(1997)5:4<280::aid-hbm13>3.0.co;2-i

Wu, D., Zhang, P., Li, C., Liu, N., Jia, W., Chen, G., et al. (2020). Perceptual learning at higher trained cutoff spatial frequencies induces larger visual improvements. *Front. Psychol.* 11:265. doi: 10.3389/fpsyg.2020.00265

Xi, J., Jia, W., Feng, L., Lu, Z. L., and Huang, C. B. (2014). Perceptual learning improves stereoacuity in amblyopia. *Invest. Ophthalmol. Vis. Sci.* 55, 2384–2391. doi: 10.1167/iovs.13-12627

Xiao, L. Q., Zhang, J. Y., Wang, R., Klein, S. A., Levi, D. M., and Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Curr. Biol.* 18, 1922–1926. doi: 10.1016/j.cub.2008.10.030

Xu, P., Lu, Z., Qiu, Z., and Zhou, Y. (2006). Identify mechanisms of amblyopia in Gabor orientation identification with external noise. *Vis. Res.* 46, 3748–3760. doi: 10.1016/j.visres.2006.06.013

Yan, F.-F., Zhou, J., Zhao, W., Li, M., Xi, J., Lu, Z.-L., et al. (2015). Perceptual learning improves neural processing in myopic vision. *J. Vis.* 15, 1–14.

Yan, Y., Rasch, M. J., Chen, M., Xiang, X., Huang, M., Wu, S., et al. (2014). Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nat. Neurosci.* 17, 1380–1387. doi: 10.1038/nn.3805

Yu, C., Klein, S. A., and Levi, D. M. (2004). Perceptual learning in contrast discrimination and the (minimal) role of context. *J. Vis.* 4, 169–182.

Yu, Q., Zhang, P., Qiu, J., and Fang, F. (2016). Perceptual learning of contrast detection in the human lateral geniculate nucleus report. *Curr. Biol.* 26, 3176–3182. doi: 10.1016/j.cub.2016.09.034

Zhang, G., Cong, L. J., Song, Y., and Yu, C. (2013). ERP P1-N1 changes associated with Vernier perceptual learning and its location specificity and transfer. *J. Vis.* 13, 1–13.

Zhang, G., Li, H., Song, Y., and Yu, C. (2015). ERP C1 is top-down modulated by orientation perceptual learning. *J. Vis.* 15, 1–11.

Zhang, J. Y., Zhang, G. L., Xiao, L. Q., Klein, S. A., Levi, D. M., and Yu, C. (2010). Rule-based learning explains visual perceptual learning and its specificity and transfer. *J. Neurosci.* 30, 12323–12328.

Zhou, Y., Huang, C., Xu, P., Tao, L., Qiu, Z., Li, X., et al. (2006). Perceptual learning improves contrast sensitivity and visual acuity in adults with anisometropic amblyopia. *Vis. Res.* 46, 739–750. doi: 10.1016/j.visres.2005.07.031

Zhou, Y., Huang, C., Xu, P., Tao, L., Qiu, Z., Li, X., et al. (2007). Perceptual learning improves contrast sensitivity and visual acuity in adults with anisometropic amblyopia (vol 46, pg 739, 2006). *Vis. Res.* 47, 2113–2113.

Zoefel, B., Huster, R. J., and Herrmann, C. S. (2011). Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. *Neuroimage* 54, 1427–1431. doi: 10.1016/j.neuroimage.2010.08.078

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Medicine

ORIGINAL RESEARCH
published: 28 January 2021
doi: 10.3389/fmed.2020.622877

Strongest Correlation Between Contrast Sensitivity and Morphological Characteristics in Bilateral nAMD

Laura Hoffmann[1†], *Petra Rossouw*[1,2†], *Maria-Magdalena Guichard*[1] *and Katja Hatz*[1,3*]

[1] *Vista Klinik, Binningen, Switzerland,* [2] *Department of Vision Science and Optometry, University of Aalen, Aalen, Germany,* [3] *Faculty of Medicine, University of Basel, Basel, Switzerland*

In patients with neovascular age-related macular degeneration (nAMD) there is often an inconsistency between their subjective visual impairment and a still relatively preserved standard Early Treatment of Diabetic Retinopathy Study (ETDRS) best corrected visual acuity. Therefore, in order to better capture the specific functional defects in nAMD, other tests need to be evaluated. In a previous study, we reported contrast sensitivity of the better eye to best correlate with near distance and distance vision related quality of life in patients with bilateral nAMD. Here, we evaluated Pelli-Robson contrast sensitivity, ETDRS visual acuity, low luminance visual acuity and Radner maximum reading speed and correlated them with several morphologic parameters as measured on fundus autofluorescence imaging, optical coherence tomography and optical tomography angiography in 54 patients. A multiple regression analysis was performed which correlated each visual function parameter with the anatomic features. The results showed the strongest correlations between the total area of macular geographic atrophy as well as the percentage of geographic atrophy in the central 1 mm and contrast sensitivity. Further, the regression model selected the total area of macular geographic atrophy, the photoreceptor inner and outer segments interface disruption score, the presence of subretinal fibrosis in the central 1 mm and the central retinal thickness as the variables that explained 71% of the variation in contrast sensitivity when including all eyes. Hence, our results suggest that among the evaluated measures of vision, contrast sensitivity is best correlated with the morphologic impairment in bilateral nAMD. Thus, contrast sensitivity may complement ETDRS visual acuity in clinical trials and serve as a standard diagnostic tool in clinical practice.

Keywords: contrast sensitivity, age-reated macular degeneration, AMD, visual function, OCT, OCTA, geographic atrophy, fundus autofluorecence

OPEN ACCESS

Edited by:
Fang Hou,
Wenzhou Medical University, China

Reviewed by:
William Ridder,
Marshall B. Ketchum University, United States
Carlo Gesualdo,
University of Campania Luigi Vanvitelli, Italy

***Correspondence:**
Katja Hatz
katja.hatz@vista.ch

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: *29 October 2020*
Accepted: *30 December 2020*
Published: *28 January 2021*

Citation:
Hoffmann L, Rossouw P, Guichard M-M and Hatz K (2021) Strongest Correlation Between Contrast Sensitivity and Morphological Characteristics in Bilateral nAMD. Front. Med. 7:622877. doi: 10.3389/fmed.2020.622877

INTRODUCTION

Age-related macular degeneration (AMD) represents the leading cause of blindness in the elderly in the industrialized world and a major public health concern (1). As a result of the current demographic development, a considerable rise in the affected subjects is estimated in the future (2). Despite typically mild visual impairment in the early stages, the progression of the disease results

in a severe central visual loss, which causes disability especially when both eyes are involved (3, 4). Progressive AMD is either represented by macular neovascularization (MNV) or an advanced stage of dry AMD mainly characterized by geographic atrophy (GA) of the outer retina and the retinal pigment epithelium (RPE) which might initially spare the fovea (5, 6). Eyes with neovascular AMD (nAMD) sometimes show GA zones at diagnosis or develop GA as well as subretinal scarring of MNV over the long-term (7). Subretinal fibrosis and GA, both leading to defects in the outer retina, account for the principal causes of vision loss in nAMD (8).

Typically, while patches of atrophy primarily present in the parafovea leading to difficulties in near-distance vision-related tasks such as reading, the central distance visual acuity [standard Early Treatment of Diabetic Retinopathy Study (ETDRS) protocol measurement at 4 m] is still relatively preserved (9). The loss of the RPE results in the decline of the photoreceptors so that the areas of GA correspond to an absolute scotoma (10). Within those lesions and, to a lesser extent in the still functional retina, a reduction in retinal sensitivity can be measured by the use of microperimetry (11, 12). In addition, with the enlargement of GA, a decrease in retinal sensitivity occurs (6, 13). The areas of reduced retinal sensitivity correlate morphologically with an impairment of the outer retina as observed by Optical Coherence Tomography (OCT), namely a disruption of the photoreceptors' inner and outer segments (IS/OS) interface and loss of the external limiting membrane (ELM) (14–16).

However, in AMD an inconsistency exists between the severity of the morphologic impairment and visual function since a wide range of visual acuity is observed in clinical practice (17). Furthermore, conventional high-contrast best-corrected visual acuity (BCVA) testing (ETDRS at 4 m), which is widely used in clinical practice and trials, does not adequately correlate with the patient's ability to execute vision-related tasks under conditions of everyday life. Even in the early stages, patients experience difficulties under low lightning and low contrast conditions, while central distance visual acuity is still reasonably intact (18). Hence, other functional tests have been evaluated to better correlate with the subjective visual impairment. Contrast sensitivity (CS) and low-luminance visual acuity (LLVA) have been reported to best capture the specific functional impairment in AMD when compared to healthy controls (19, 20). For our study population, contrast sensitivity is also documented to strongest correlate with patient's vision-related quality of life as measured by the National Eye Institute Visual Function Questionnaire (NEI-VFQ25) (21).

These findings suggest to evaluate whether alternative functional visual measures might be more reflective of macular morphology than conventional BCVA testing. Several quantitative and qualitative morphologic parameters, namely the retinal thickness, subretinal fibrosis and location of GA have been shown to correlate with standard distance visual acuity (15, 22–24). Furthermore, pre-existing RPE atrophy and disruption in the ELM are described to serve as predictors for an early poor visual response to anti-vascular endothelial growth factor (anti-VEGF) treatment in nAMD (7). Due to the heterogeneity of the morphologic impairment in AMD, data suggest that a combination of factors including location, size and component of the lesions explain the variety of visual function in AMD (23, 25). Currently, there are few studies describing the influence of foveal involvement of fibrosis and GA and size and location of MNV on contrast sensitivity which, in addition, did not use the latest image acquisition devices (23, 26, 27).

Amendments in OCT technology reveal detailed insight in outer retinal layer changes in eyes with AMD (28). Nevertheless, there are currently limited data correlating systemically alternative visual testing parameters and the plenitude of morphologic parameters as measured by fundus autofluorescence (FAF), spectral domain OCT (SD-OCT) and swept source OCT Angiography (SS-OCT-A). Therefore, the aim of our study was to evaluate the relations between different measures of visual function (distance acuity, contrast sensitivity, reading speed and LLVA) and morphologic characteristics in patients with bilateral nAMD to identify validated morphologic parameters which may serve as endpoints in clinical practice and research.

PATIENTS AND METHODS

This was a cross-sectional, non-interventional and single-visit study of patients treated at the Vista Klinik, Binningen, Switzerland with a clinically confirmed diagnosis of bilateral nAMD by a retina specialist. The study received approval from the local ethics approval board [Ethikkommission Nordwestschweiz (EKNZ No. 2016-002216)] and was performed in accordance with the tenets of the Declaration of Helsinki and Good Clinical Practice (ICH-GCP). The study is registered at ClinicalTrial.gov (NCT 03438669).

Data were obtained between February 2017 and October 2017. Patients had to meet the following inclusion criteria to be eligible: age $\geq$ 55 years, bilateral nAMD with AMD-related lesions such as MNV or GA or subretinal fibrosis or pigment epithelial detachment (PED) within the central 1 mm ETDRS grid subfield confirmed by Spectral Domain Optical Coherence Tomography (SD-OCT; Spectralis, Heidelberg Eng., Heidelberg, Germany) and Swept-Source-OCT-Angiography (SS-OCTA; Plex Elite 9000, Carl-Zeiss Meditec, Dublin, USA), best corrected visual acuity (BCVA) of at least 49 letters (Snellen equivalent 20/100 or better) in the better eye using ETDRS charts at a distance of 4 m, sufficiently clear ocular media and adequate pupillary dilation and fixation permitting quality fundus imaging. All eyes were currently treated with intravitreal anti-vascular endothelial growth factor (VEGF) agents. Exclusion criteria were significant ocular disease other than neovascular AMD or a history of neurologic disease or cognitive impairment.

Visual Function Assessment

Study examinations took place no earlier than 4 days after the last intravitreal treatment and were always performed in the same order. All patients underwent a standardized refraction protocol with ETDRS BCVA at 4 m in designated rooms with an ambient illumination of 97–109 lux evaluated with a calibrated illuminometer. LLVA testing was realized immediately afterwards following the LLVA testing protocol of our clinic. First light

switches were turned off and then the block out shutters were closed in order to have the ETDRS light box as the only illumination left. LLVA testing took place once the patients were ready to continue and no difficulties adapting to the lower lighting conditions were reported. Since the testing was not always performed in the same rooms because of differing vacancies and time of appointment, the illumination varied slightly between 02 and 10l ux. The low luminance deficit (LLD) was defined as the number of ETDRS letters read in standard BCVA testing minus LLVA. Maximal reading speed was tested monocularly and binocularly with Radner reading charts (Precision Vision, Inc.). Contrast sensitivity was measured using Pelli Robson charts (Precision Vision, Inc.) at 1 m. Further details of the visual function assessment in this study are described elsewhere (21).

Retinal Imaging Analysis

Morphological impairment was evaluated with SD-OCT (Heidelberg Engineering, Germany), using the following scans: horizontal volume scan 19 sections, macular star 9 sections and horizontal 6 mm scan. Grading was performed by two masked physicians according to a standardized protocol (LH, MG). The foveola was localized in the volume scans to overlay an ETDRS grid centered on the fovea. All lesions seen on OCT, FAF and OCT-A scans were classified dichotomously depending on presence or absence in each of the nine subfields of the disposed ETDRS grid. Lesions were then summarized as located within the central 1, 3, or 6 mm ring.

Photoreceptor inner segment and outer segments (IS/OS) interface and external limiting membrane (ELM) disruption were evaluated by calculating the mean of the score in the horizontal and vertical scans of the macular star scan (29). Disruption was graded as 0 (no disruption in 1 mm center), 1 (mild disruption <1/4 in 1 mm center), 2 (1/4 to 3/4 disruption in 1 mm center) and 3 (>3/4 disruption in 1 mm center) (29). Central subfield thickness (CRT) analysis was performed after centering of the scan and manual alignment of the automated lines. All scans were graded with regard to the presence of subretinal fibrosis and pigment epithelial detachment (PED). Subretinal fibrosis was considered as subretinal hyper-reflective material at the level above the retinal pigment epithelium. PED was defined as a focal elevation of the retinal pigment epithelium presenting with a height $\geq$200 µm or a width $\geq$400 µm (30). Maximum height of pigment epithelial detachment in the central 1 mm was measured.

Geographic atrophy (GA) was assessed by fundus autofluorescence imaging acquired with confocal scanning laser ophthalmoscope (HRA, Heidelberg Engineering, Germany) with an excitation wavelength of 488 nm and an emission spectrum of 500–700 nm. The GA pattern was classified as focal or multifocal and the surface area was measured with the built-in free-hand draw tool. Percentage of the spared surface within the central 1 mm ring was calculated.

All OCT-A analyses were performed with PLEX Elite 9000 Swept-Source OCT Angiography (Carl Zeiss Meditec Inc, Dublin, USA) using a 6 mm scan centered on the fovea. Manual correction was carried out to ensure an accurate segmentation. The CNV membrane was located in the en face OCT-A slabs of the outer retinal layer and choriocapillaris (ORCC) and surface was measured using the Image J tool (NIH Image, Bethesda, MD).

Statistical Analysis

Statistical analysis was performed using SPSS statistical package version 21 (SPSS, Inc., Chicago, IL). Data are presented as mean ± standard deviation (SD) or percentages, as appropriate. Lesions were subdivided depending on their presence or absence in the central 1, 3, or 6 mm ring and an unpaired *t*-test for comparison of means was used to test for differences in visual function between these groups. *P*-values for multiple comparisons were adjusted according to the Bonferroni-Holm adjustment (31). To evaluate correlations between morphologic and visual function parameters, Pearson correlation coefficients were calculated. Afterwards a multiple regression analysis was performed using the variables which were previously statistically significantly correlated with visual function as explanatory variables. To reduce potential collinearity, highly correlated variables ($r > 0.90$) were not included in the same model. Backward stepwise elimination was performed to select independent explanatory parameters. Each time a visual function parameter was used as the dependant variable and the model with the highest coefficient of determination (R^2) was chosen. Despite the known dependence of both eyes of a patient (31), here data are reported for the better, the worse and for both eyes of each patient nevertheless in order to evaluate a possible early biomarker reflective of retinal morphology when comparing the correlations with retinal morphology according to the status of the eye. *P*-values <0.05 were considered statistically significant.

RESULTS

Patient Characteristics

This study included 108 eyes from 54 patients with a mean age of 79.7 ± 7.8 years. 53.7% of patients were female, 46.3% were male.

Visual Function

Detailed analysis of the assessment of visual function of the study population has been described elsewhere (21). Mean BCVA for all eyes was 66.75 ± 24.33 ETDRS letters (0.37 ± 1.21 logMAR) with a statistically significant difference between the better and the worse eye (78.91 ± 7.93 letters (0.12 ± 1.54 logMAR) and 47.43 ± 35.96 letters (0.75 ± 0.98 logMAR), respectively, $p < 0.001$), see **Table 1**. Mean LLVA showed comparable results with a mean of 66.06 ± 24.17 letters (0.38 ± 1.22 logMAR). Mean contrast sensitivity for all eyes was 1.21 ± 0.42 log units. Mean binocular maximum reading speed did not differ significantly from mean maximum reading speed of the better eye ($p = 0.7287$). Mean monocular maximum reading speed for all eyes was significantly lower than binocular maximum reading speed (96.87 ± 48.16 wpm, $p < 0.001$).

Morphologic Parameters

Morphological characteristics are summarized in **Table 2**. There was a statistically significant difference for all of these parameters regarding the status of the eyes (better vs. worse eyes, $p < 0.01$)

TABLE 1 | Visual assessment.

	All eyes ($n = 108$)	Better eyes ($n = 54$)	Worse eyes ($n = 54$)
Mean BCVA (ETDRS letters) (logMAR) ±SD	66.75 ± 24.33 (0.37 ± 1.21)	78.91 ± 7.93 (0.12 ± 1.54)	47.43 ± 35.96 (0.75 ± 0.98)
Mean LLVA (ETDRS letters) (logMAR) ±SD	66.06 ± 24.17 (0.38 ± 1.22)	78.13 ± 8.04 (0.14 ± 1.54)	47.37 ± 35.39 (0.75 ± 0.99)
Low luminance deficit (LLD in letters)	0.69 ± 3.41	0.98 ± 3.81	0.41 ± 2.94
Mean CS (log units) ±SD	1.21 ± 0.42	1.33 ± 0.27	1.09 ± 0.50
Mean monocular MRS (wpm) ±SD	96.87 ± 48.16	118.16 ± 27.88	75.59 ± 54.61
Mean binocular MRS (wpm) ±SD	117.33 ± 28.42		
NEI-VFQ25 near distance subscale	74.69 ± 18.74		
NEI-VFQ25 distance subscale	74.15 ± 21.90		

TABLE 2 | Morphologic characteristics.

	All eyes ($n = 108$)	Better eyes ($n = 54$)	Worse eyes ($n = 54$)
FAF			
Mean total area of GA (mm^2) ±SD	3.02 ± 6.67	1.41 ± 3.90	4.60 ± 8.30
Mean percentage of GA in central 1 mm ±SD	19.40 ± 32.99	9.76 ± 23.95	29.05 ± 37.86
OCT			
Mean CRT (μm) ±SD	332.60 ± 187.45	294.39 ± 76.56	353.09 ± 201.94
Disruption IS/OS interface central 1 mm (%)	91.7	90.7	92.6
Disruption ELM central 1 mm (%)	83.3	81.5	85.2
Subretinal fibrosis central 1 mm (%)	17.6	3.7	31.5
PED in central 1mm (%)	59.3	66.7	51.9
OCT-A			
Mean total area of CNV (mm^2) ±SD	2.10 ± 3.23	1.53 ± 1.77	2.69 ± 4.17

except for mean total area of CNV, presence of PED in the central 1 mm and disruption of the IS/OS interface in the central 1 mm. 26.9% of all eyes ($n = 29$) presented without geographic atrophy. For all other eyes, mean area of total macular GA was 3.02 ± 6.67 mm^2, mean percentage of GA in the central 1 mm was 19.40 ± 32.99%. Considering only the eyes with GA ($n = 79$), 49.4% presented with multifocal GA. 63.9% of all eyes ($n = 69$) had sparing in the central 1 mm.

Mean central retinal thickness (CRT) was 332.60 ± 187.45 μm, mean height of pigment epithelium detachment (PED) in the central 1 mm was 156.02 ± 103.41 μm. 8.3% of all eyes ($n = 9$) showed an intact junction of the photoreceptors inner and outer segments (IS/OS junction line) and 16.7% ($n = 18$) an intact external limiting membrane (ELM). 17.6% of all eyes ($n = 19$) presented with subretinal fibrosis and 59.3% ($n = 64$) with a PED in the central 1 mm. 25 eyes (23.1%) presented with reticular pseudodrusen (RPD). 82.4% of all eyes ($n = 89$) showed a clearly identifiable choroidal neovascularization membrane in the OCT-A scans. In the remaining eyes the identification of a clear CNV membrane was limited at the time of the study due to progressed fibrosis and/or GA following a long-term anti-vascular endothelial growth factor (anti-VEGF) treatment. Seventy eyes (64.8%) were initially fluorescein angiography diagnosed as occult lesions, 27 (25%) as minimally classic, 8 (7.4%) as predominantly classic, and three as retinal angiomatous proliferation. Mean area of CNV was 2.10 ± 3.23 mm^2 (range 0–24.32 mm^2). Thirty-five eyes had sparing in the central 1 mm.

Correlations Between Visual Acuity Parameters and FAF, OCT, and OCT-A Parameters

The correlations are presented for all eyes as well as separately for the better and the worse eyes, respectively, in **Table 3**. All visual acuity parameters were statistically significantly correlated with the GA parameters measured by FAF when considering all eyes (**Table 3**).

Strongest correlations were observed between contrast sensitivity and total area of macular GA ($r = -0.766$, $p < 0.001$) as well as percentage of GA in the central 1 mm ($r = -0.693$, $p < 0.001$) for all eyes. Interestingly, in the better eye group BCVA and LLVA did not correlate with the FAF parameters whereas CS and maximum reading speed still exhibited strong correlations. **Figure 1** illustrates an eye with GA involving the central 1 mm but leaving an area of anatomic foveal preservation. The patient still had a well-preserved standard ETDRS distance BCVA of 80 letters but already showed a diminished CS and monocular MRS (1.20 log units and 81.28 wpm, respectively). For the correlations of maximum binocular reading speed and morphologic parameters, only the better eyes ($n = 54$) were included since there was no statistically significant difference between maximal binocular reading speed and maximal monocular reading speed of the better eyes. The binocular MRS was moderately correlated with the FAF parameters ($r = -0.395$, $p = 0.001$). Furthermore, significant associations were found between the NEI-VFQ25 near distance subscale and the total area of macular GA in the better eye group ($r = -0.493$, $p = 0.001$).

Considering the OCT parameters strong negative correlations were seen between IS/OS interface and ELM disruption score in the central 1 mm and all of visual acuity parameters except binocular MRS for which correlations were moderate ($p < 0.001$). Compared to the other functional measures, CS showed the strongest correlations with the mentioned OCT parameters with the highest correlation coefficient with the IS/OS interface disruption score when considering all eyes ($r = -0.619$, $p < 0.001$). The correlation coefficients were comparable between the better and the worse eye group although they were superior for CS than BCVA and LLVA in the better eyes. CRT exhibited moderate statistically significant negative correlations with all

TABLE 3 | Correlations between visual acuity measures, OCT, FAF and OCT-A parameters.

	BCVA (ETDRS letters)			LLVA (ETDRS letters)			CS (log units)			MRS monocular (wpm)		
	All eyes	Better eyes	Worse eyes	All eyes	Better eyes	Worse eyes	All eyes	Better eyes	Worse eyes	All eyes	Better eyes	Worse eyes
FAF												
Total area of macular GA	− 0.634	−0.024	− 0.694	− 0.643	−0.010	− 0.710	− 0.766	− 0.706	− 0.763	− 0.661	− 0.602	− 0.648
Percentage of GA in central 1 mm	− 0.651	−0.107	− 0.720	− 0.644	−0.062	− 0.725	− 0.693	− 0.549	− 0.709	− 0.688	− 0.532	− 0.690
OCT												
IS/OS interface disruption score in central 1 mm	− 0.537	− 0.451	− 0.548	− 0.525	− 0.390	− 0.584	− 0.619	− 0.523	− 0.548	− 0.596	− 0.470	− 0.632
ELM disruption score in central 1 mm	− 0.590	− 0.469	− 0.572	− 0.586	− 0.393	− 0.555	− 0.591	− 0.495	− 0.572	− 0.601	− 0.451	− 0.697
CRT	− 0.423	0.102	− 0.436	− 0.450	−0.046	− 0.457	− 0.461	0.054	− 0.539	− 0.324	0.250	− 0.400
PED height in central 1 mm	0.064	−0.119	0.106	0.050	−0.163	0.078	0.026	0.018	0.017	0.142	0.241	0.094
OCT-A												
Total area of CNV	− 0.558	0.197	− 0.640	− 0.572	0.078	− 0.647	− 0.588	0.135	− 0.698	− 0.416	0.106	− 0.502

Correlations between multimodal imaging parameters and visual acuity parameters for all eyes (n = 108), better and worse eyes (n = 54, respectively): p < 0.001—dark gray filled cells, p ≥ 0.001 to <0.01—medium gray filled cells.

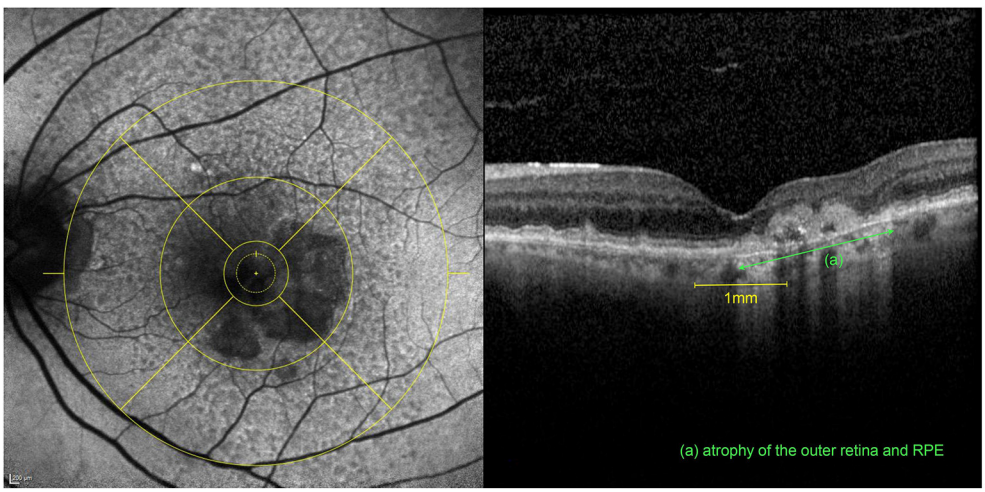


FIGURE 1 | Multimodal imaging of the left eye of an 81-year-old patient after 6 intravitreal aflibercept and 19 ranibizumab injections for nAMD. The patient had a preserved BCVA of 80 letters ETDRS with an already diminished CS (1.20 log units) and monocular MRS (81.28 wpm) (compared to the mean values of all included eyes of 66.7 letters ETDRS, 1.21 log units CS and 96.9 wpm monocular MRS, respectively). **Left:** FAF image with an ETDRS grid overlay centered on the fovea with a horseshoe-shaped area of atrophy involving the central 1mm. **Right:** Corresponding volume OCT scan revealing at the same time, an area of intact anatomic preservation on the left, associated with an atrophy of the outer retina (ELM and IS/OS interface disruption) and the retinal pigment epithelium (RPE) (green) on the right within the fovea.

measures of visual acuity measured monocularly in the worse eyes. The maximal height of pigment epithelial detachment in the central 1 mm was not significantly correlated with visual function. Eyes with reticular pseudodrusen ($n = 25$) did not show a significant difference in BCVA ($p = 0.156$), LLVA ($p = 0.297$), CS ($p = 0.276$), or MRS ($p = 0.875$) when compared to eyes without them ($n = 83$).

The total area of CNV exhibited moderate negative correlations with the visual acuity parameters with strongest correlation with contrast sensitivity for all eyes ($r = -0.588$, $p < 0.001$). However, the total area of CNV was not significantly correlated with visual function when including only the better eyes.

Relation Between Location of Lesions and Visual Acuity Parameters

Table 4 shows the influence of the presence or absence of each lesion component in the different ETDRS subfields on the different visual acuity parameters. The inner subfields of the ETDRS grid were summarized as the central 3 mm ring and the outer subfields as the central 6 mm ring. Comparison of means showed that all visual acuity parameters where significantly worse

TABLE 4 | Relations between location of lesions and visual function.

	p-values comparing presence of lesion and visual acuity parameters (unpaired *t*-test)											
	BCVA (ETDRS letters)			LLVA (ETDRS letters)			CS (log units)			MRS monocular (wpm)		
	All eyes	Better eyes	Worse eyes	All eyes	Better eyes	Worse eyes	All eyes	Better eyes	Worse eyes	All eyes	Better eyes	Worse eyes
FAF												
Presence of GA in central 1 mm	0.016	0.285	0.016	0.016	0.572	0.016	0.016	0.015	0.016	0.016	0.048	0.016
Presence of GA in central 3mm	0.016	1	0.016	0.016	0.068	0.016	0.016	0.156	0.016	0.016	0.650	0.016
Presence of GA in central 6mm	0.016	1	0.052	0.015	1	0.039	0.016	0.924	0.015	0.016	1	0.028
OCT												
IS/OS interface disruption												
Present in central 3 mm	1	0.285	1	1	0.252	1	1	0.042	1	1	1	1
Present in central 6 mm	1	0.742	0.770	1	1	0.616	0.126	1	0.015	1	1	0.015
ELM disruption												
Present in central 3 mm	1	0.285	1	1	0.252	1	1	0.042	1	1	1	1
Present in central 6 mm	1	1	0.228	1	1	0.192	0.126	1	0.042	1	1	0.015
Subretinal fibrosis												
Present in central 1 mm	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.182	0.016
Present in central 3 mm	0.016	0.016	0.016	0.016	0.015	0.016	0.016	0.015	0.016	0.016	0.150	0.016
Present in central 6 mm	0.015	1	0.016	0.015	1	0.016	0.015	1	0.016	0.015	1	0.016
PED												
Present in central 1 mm	0.280	1	0.450	0.392	1	0.550	0.364	1	0.672	0.015	1	0.120
Present in central 3 mm	1	1	1	1	1	1	1	1	1	1	1	1
Present in central 6 mm	1	1	1	1	1	1	1	1	1	1	1	1
OCT-A												
Presence of CNV in central 1 mm	0.015	1	0.028	0.015	1	0.028	0.444	1	0.130	1	1	1
Presence of CNV in central 3 mm	0.016	1	0.015	0.016	1	0.015	0.045	1	0.015	0.392	1	0.52
Presence of CNV in central 6 mm	0.943	1	1	1	1	1	1	1	1	1	1	1

Relations between presence or absence of lesion components in different locations and visual acuity parameters for all eyes (n = 108), better and worse eyes (n = 54, respectively), p-values adjusted for multiple comparisons with the Bonferroni-Holm adjustment. p ≥ 0.01 to 0.05—light gray filled cells.

in the presence of geographic atrophy and subretinal fibrosis in the central 1, 3, and 6 mm ring when considering all eyes. A disrupted IS/OS interface and ELM in the central 3 mm, respectively, 6 mm was associated with a significantly lower CS both in the better and the worse eye group whereas BCVA and LLVA were not significantly changed. Eyes with a CNV in the central 3 mm exhibited a significantly worse CS, LLVA and BCVA considering all eyes. The presence of a PED was not associated with significantly worse visual acuity parameters.

Multiple Regression Analysis

Multiple regression analysis run each time with a visual function parameter as the dependent variable, yielded values of adjusted R^2 of 0.707 for contrast sensitivity, 0.596 for BCVA, 0.588 for LLVA and 0.591 for monocular MRS when all eyes were included. The model with contrast sensitivity as the dependant variable with the highest adjusted R^2 is summarized in **Table 5**. This model selected the total area of macular GA, the IS/OS interface disruption score, the presence of subretinal fibrosis in the central 1 mm and CRT as significant independent parameters. Therefore, these factors explain ∼71% of the variation of contrast sensitivity.

When subdividing into two groups depending on the status of the eye (better vs. worse), results indicated a higher coefficient of determination for the worse eyes (0.772 vs. 0.586). Furthermore, only the total area of macular GA, the IS/OS junction line disruption score and CRT (only for the worse eyes) were selected as significant independent variables.

DISCUSSION

Several studies have reported that anatomic parameters such as CNV area and diameter do not adequately explain changes in distance visual acuity in nAMD (23, 26). Yet, the majority of those studies did not include a multitude of morphologic parameters graded on the basis of the ETDRS grid subfields using latest image acquisition devices such as SD-OCT and OCT-A and did not evaluate different measures of visual function.

Our results suggest that contrast sensitivity (CS) exhibits the most consistent correlations with many of the anatomic parameters compared to the other assessed measures of vision. Likewise, Ghoshal et al. correlated different functional measures and retinal morphology and found CS to be associated with a few OCT parameters including thickness and volume of RPE

TABLE 5 | Multiple regression analysis of contrast sensitivity and morphologic parameters.

Independent variables	Adjusted R^2	Beta coefficient of regression	*p*-value
All eyes ($n = 108$)	0.707		
Total area of macular GA		−0.035	<0.001
IS/OS interface disruption score		−0.088	0.004
Subretinal fibrosis in central 1mm		−0.191	0.030
CRT		−0.001	<0.001
Better eyes ($n = 54$)	0.586		
Total area of macular GA		−0.042	<0.001
IS/OS interface disruption score		−0.100	0.001
Worse eyes ($n = 54$)	0.772		
Total area of macular GA		−0.041	<0.001
IS/OS interface disruption score		−0.112	<0.001
CRT		−0.001	0.011

and Bruch membrane (32). In our study, although significant correlations were observed for all measures of visual function, CS exhibited the strongest correlation coefficients for all the evaluated morphologic parameters and yielded the highest coefficient of determination in multiple regression analysis.

With regards to the component of the lesions, in our study the area of atrophy and the presence of subretinal fibrosis in the central 1 mm were selected as independent variables explaining the variation in CS for all eyes. In addition, their presence in all of the ETDRS subfields was significantly correlated with CS (for the better eye group only in the central 1 mm). Their influence on visual function in general is not surprising as it is in accordance with several other studies (23, 33). Subretinal fibrosis as the end stage of nAMD leads to a destruction of the overlying photoreceptors; on the one hand by interfering with the metabolic exchange with the RPE and on the other hand by a direct toxic effect (34). Likewise, subretinal fibrosis and atrophy go along with a destruction of the outer retina and the RPE resulting in decreased visual function. Similarly, Keane et al. reported strongest correlations between a decreased CS and an increased subretinal tissue volume for newly diagnosed nAMD patients explaining 24% of the variation in CS at baseline (24).

In our study both the total area of geographic atrophy (GA) as defined by hypoautofluorescence in autofluorescence imaging as well as the percentage of GA in the central 1 mm were strongly associated with CS and other measures of visual function. Accordingly, Ooto et al. reported a negative correlation between CS and size of confluent hypoautofluorescence as well as its involvement of the fovea in dry AMD (35). Despite the additional presence of other components in nAMD our results suggest that GA is a major factor influencing visual function in nAMD, too. Furthermore, we identified the percentage of GA in the central 1 mm as a significant influencing factor which has been previously described to better correlate with distance visual acuity than total GA size (22). In addition, it can serve as a better tool to predict visual impairment over time than the simply binary grading of foveal-sparing status. Furthermore, in our study the presence of GA in the central 3 and 6 mm had a significant impact on visual function. Correspondingly, Sayegh et al. reported that progression of GA depends on the distance to the fovea and argue for a grading of sparing in the central 3 mm as it correlates with both atrophy progression and visual acuity (36).

With newer OCT devices, the visualization of the degeneration of the photoreceptors is thought to be visible by a disruption of ELM and the IS/OS interface which may be present not only in atrophic areas but also in the junctional zones (37). In our study, the ELM and IS/OS interface disruption score exhibited strongest correlations with CS and a disruption of ELM and IS/OS interface in the central 3 mm (better eye group) and the central 6 mm (worse eye group) was significantly associated only with CS but not the other measures of vision. Whereas, the size of the ellipsoid zone defect has been shown to correlate with microperimetry scores (35), there are currently no studies investigating the relation between ELM and IS/OS interface disruption and CS in anti-VEGF treated nAMD patients.

Although distance visual acuity is widely used in clinical trials, several studies documented that it correlated less with morphological features in nAMD than for example near visual acuity (23, 26), indicating that standard BCVA at 4 m does not adequately correlate with foveal impairment in nAMD. In the present study, correlation coefficients were also generally weaker compared to CS. We observed a similar significant but moderate correlation between central retinal thickness (CRT) and ETDRS BCVA as well as CS when considering all eyes. Retinal thickening associated with intraretinal fluid exudation is commonly observed in nAMD. Although a reduction in CRT after anti-VEGF treatment has been shown to moderately correlate with long-term visual outcomes, many studies failed to detect stronger correlations between CRT and visual function (24, 38, 39). Similarly, our results may be reflective of the heterogeneous morphology in AMD where reduced retinal thickness in advanced atrophic lesions is associated with severe vision loss. Further, in our study the total area of CNV as evaluated on OCT-A was moderately correlated with visual function in the worse eyes and its presence in the central 1 and 3 mm showed a highly significant association with ETDRS visual acuity. In other studies based on angiographic measurements, these correlations were weaker and inconsistent (23, 26), probably due to the fact that our study cohort had already experienced a long disease duration so that the CNV had often resulted in a considerable amount of atrophy or fibrosis within the macula. With growth of the CNV below the RPE, the formation of a fibrovascular PED may occur which is often accompanied by leakage or hemorrhage. Our results did not indicate a correlation between the presence or the height of PED and ETDRS BCVA which is consistent with findings of other studies (24, 40). This lack of association is not surprising as even large PED may go along with functioning overlying photoreceptors. Furthermore, we did not characterize the PED subtypes, thus a possible correlation dependent on the structural characteristics of the PED remains unclear.

Whereas, LLVA has been shown to capture visual impairment in nAMD better than standard distance acuity (41), there is currently limited data correlating systematically LLVA and

various morphologic characteristics. The Chroma and Spectri studies reported a weak negative correlation between LLVA and GA lesion size at baseline and week 48 in bilateral GA in dry AMD (42). Further, patients with foveal involvement already at baseline showed less decline in LLVA but not the other functional measures which reflects its relevance as a measure of foveal cone function in early disease stages in the absence of foveal lesions. In the present study LLVA showed globally similar correlations with retinal morphology than standard distance acuity, most likely due to their high intercorrelation level.

While reading speed is known to be significantly reduced in eyes with nAMD (9), in our study, correlations between the binocular MRS and the morphologic impairment of the better eye were globally weaker and inconsistent. This was probably due to the fact that the influence of the fellow eye was not taken into account. Since reading is performed best with central vision, monocular reading which prevents the compensation of a monocular central field loss by the fellow eye, is severely compromised in nAMD, possibly resulting in weaker correlations for the monocular MRS. Nevertheless, in our study monocular MRS exhibited stronger correlation coefficients with the autofluorescence parameters and ELM and IS/OS interface disruption score than standard BCVA. Reviewing the literature, reading speed has been described to correlate significantly with the size of GA (34, 42), central retinal thickness (24) as well as the foveal involvement of atrophy and subretinal fibrosis (23). Our results correspond to these findings with significant relations between monocular MRS and the presence of GA and subretinal fibrosis in the central 1, 3, and 6 mm considering all eyes. Especially the size of the foveal spared area seems to influence MRS as it has been shown that a minimal spared central area is required for fluent reading, otherwise the characters would not "fit" in the preserved space (43). Further, the Chroma and Spectri studies found MRS to be best correlated with progression of GA size compared to other functional measures in bilateral GA in dry AMD (42). Likewise, in our study strongest correlations for the monocular MRS were observed with the percentage of GA in the central 1 mm.

Analysis with two subgroups (better vs. worse eyes) revealed similar strong correlation coefficients between CS and the GA parameters as well as the disruption of the IS/OS interface and ELM independently of the status of the eye. However, standard BCVA was not reflective of the atrophic lesions in the better eyes. Hence, CS could serve as an early biomarker of impaired visual function especially in early disease stages and therefore be used even in the better eyes of patients with bilateral nAMD. Generally, the correlation coefficients were lower in the better eye than in the worse eye group which is consistent with findings in the literature (27). Possibly, the better eyes performed better than expected considering their morphologic lesions, since the subjects largely depended on it due to the impairment of the fellow eye.

To the best of our knowledge, there are no studies so far that systemically correlated a multitude of morphologic parameters graded on FAF, SD-OCT and SS-OCT-A with several measures of visual function of subjects with bilateral nAMD. The performance of a standardized visual assessment protocol under the same conditions and the homogenous study population contribute to the strengths of our study. Further standardized image evaluation protocols were used for FAF, SD-OCT and SS-OCT-A analyses. Limitations of this study included its cross-sectional design and relatively small study population. Unfortunately, due to a lengthy disease course a reliable statement about the activity status based on shape/branching pattern/anastomoses/morphology of vessel termini/perilesional halo was not always possible in the cross sectional OCTA exam. Furthermore, the LLVA testing protocol of our clinic did not include the use of a KODAK filter, thus potentially precluding a bigger low luminance deficit. Our results should be validated in larger cohorts evaluating how the progression of morphological impairment correlates with visual function over time in order to investigate its predictive value on future visual acuity change.

To summarize, in this study visual function did not depend on an unique anatomic parameter indicating that in nAMD which does not present as a single homogenous morphologic entity, the extent, composition and size of the lesions taken together account for the associated visual impairment. This study suggests that in bilateral nAMD contrast sensitivity is better correlated with anatomic characteristics than other functional visual measures including standard ETDRS BCVA with strongest correlations with total area of macular GA and IS/OS interface disruption score. These correlations were consistent even in the better eyes of a patient contrary to standard BCVA, thus CS may serve as a better biomarker in early disease stages. Furthermore, in comparison to the other assessed visual function measures, CS has been shown to best correlate with vision-related quality of life as reflected by the NEI-VFQ25 distance and near distance score (21). Thus, given the more consistent correlation with morphologic characteristics, the incorporation of CS as a surrogate endpoint in clinical trials may lead to improved accuracy. Furthermore, it may be used as a standard diagnostic tool in clinical practice to better reflect the patient's individual visual impairment.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, upon personal request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethikkommission Nordwest- und Zentralschweiz. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LH and PR: acquisition, interpretation of data, and statistics and writing of manuscript. M-MG: interpretation of data. KH: design of the study, acquisition, interpretation of data, supervision,

and writing of manuscript. All authors critically reviewed the manuscript and approved the final version.

FUNDING

This study was supported by a grant of Retina Suisse Foundation.

ACKNOWLEDGMENTS

The authors thank Mrs. Susanne Müller for her valuable support during document preparation for ethics approval, data collection and manuscript submission. They also thank Mr. Jaco Botha for providing extensive statistical assistance and support.

REFERENCES

1. Wong WL, Su X, Li X, Cheung CM, Klein R, Cheng CY, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health*. (2014) 2:e106–16. doi: 10.1016/S2214-109X(13)70145-1
2. Colijn JM, Buitendijk G, Prokofyeva E, Alves D, Cachulo ML, Khawaja AP, et al. Prevalence of age-related macular degeneration in Europe: the past and the future. *Ophthalmology*. (2017) 124:1753–63. doi: 10.1016/j.ophtha.2017.05.035
3. Cahill MT, Banks AD, Stinnett SS, Toth CA. Vision-related quality of life in patients with bilateral severe age-related macular degeneration. *Ophthalmology*. (2005) 112:152–8. doi: 10.1016/j.ophtha.2004.06.036
4. Elshout M, Webers CA, van der Reis MI, de Jong-Hesse Y, Schouten JS. Tracing the natural course of visual acuity and quality of life in neovascular age-related macular degeneration: a systematic review and quality of life study. *BMC Ophthalmol*. (2017) 17:120. doi: 10.1186/s12886-017-0514-3
5. Lindner M, Böker A, Mauschitz MM, Göbel AP, Fimmers R, Brinkmann CK, et al. Directional kinetics of geographic atrophy progression in age-related macular degeneration with foveal sparing. *Ophthalmology*. (2015) 122:1356–65. doi: 10.1016/j.ophtha.2015.03.027
6. Fleckenstein M, Mitchell P, Freund KB, Sadda S, Holz FG, Brittain C, et al. The progression of geographic atrophy secondary to age-related macular degeneration. *Ophthalmology*. (2018) 125:369–90. doi: 10.1016/j.ophtha.2017.08.038
7. Zandi S, Weisskopf F, Garweg JG, Pfister IB, Pruente C, Sutter F, et al. Pre-existing RPE atrophy and defects in the external limiting membrane predict early poor visual response to ranibizumab in neovascular age-related macular degeneration. *Ophthalmic Surg Lasers Imaging Retina*. (2017) 48:326–32. doi: 10.3928/23258160-20170329-07
8. Okeagu CU, Agrón E, Vitale S, Domalpally A, Chew EY, Keenan T, et al. Principal cause of poor visual acuity following neovascular age-related macular degeneration: AREDS2 Report Number 23. *Ophthalmol Retina*. (2020) S2468-6530(20)30408-5. Advance online publication. doi: 10.1016/j.oret.2020.09.025
9. Sunness JS, Rubin GS, Applegate CA, Bressler NM, Marsh MJ, Hawkins BS, et al. Visual function abnormalities and prognosis in eyes with age-related geographic atrophy of the macula and good visual acuity. *Ophthalmology*. (1997) 104:1677–91. doi: 10.1016/S0161-6420(97)30079-7
10. Sunness JS, Schuchard RA, Shen N, Rubin GS, Dagnelie G, Haselwood DM. Landmark-driven fundus perimetry using the scanning laser ophthalmoscope. *Invest Ophthalmol Vis Sci*. (1995) 36:1863–74.
11. Pilotto E, Guidolin F, Convento E, Spedicato L, Vujosevic S, Cavarzeran F, et al. Fundus autofluorescence and microperimetry in progressing geographic atrophy secondary to age-related macular degeneration. *Br J Ophthalmol*. (2013) 97:622–6. doi: 10.1136/bjophthalmol-2012-302633
12. Meleth AD, Mettu P, Agrón E, Chew EY, Sadda SR, Ferris FL, et al. Changes in retinal sensitivity in geographic atrophy progression as measured by microperimetry. *Invest Ophthalmol Vis Sci*. (2011) 52:1119–26. doi: 10.1167/iovs.10-6075
13. Wong EN, Chew AL, Morgan WH, Patel PJ, Chen FK. The use of microperimetry to detect functional progression in non-neovascular age-related macular degeneration: a systematic review. *Asia Pac J Ophthal (Phila)*. (2017) 6:70–9. doi: 10.22608/APO.201643
14. Wu Z, Ayton LN, Luu CD, Guymer RH. Relationship between retinal microstructures on optical coherence tomography and microperimetry in age-related macular degeneration. *Ophthalmology*. (2014) 121:1445–52. doi: 10.1016/j.ophtha.2014.01.025
15. Sayegh RG, Kiss CG, Simader C, Kroisamer J, Montuoro A, Mittermüller TJ, et al. A systematic correlation of morphology and function using spectral domain optical coherence tomography and microperimetry in patients with geographic atrophy. *Br J Ophthalmol*. (2014) 98:1050–5. doi: 10.1136/bjophthalmol-2014-305195
16. Pilotto E, Benetti E, Convento E, Guidolin F, Longhin E, Parrozzani R, et al. Microperimetry, fundus autofluorescence, and retinal layer changes in progressing geographic atrophy. *Can J Ophthalmol*. (2013) 48:386–93. doi: 10.1016/j.jcjo.2013.03.022
17. Bressler SB, Bressler NM, Fine SL, Hillis A, Murphy RP, Olk RJ, et al. Natural course of choroidal neovascular membranes within the foveal avascular zone in senile macular degeneration. *Am J Ophthalmol*. (1982) 93:157–63. doi: 10.1016/0002-9394(82)90410-X
18. Scilley K, Jackson GR, Cideciyan AV, Maguire MG, Jacobson SG, Owsley C. Early age-related maculopathy and self-reported visual difficulty in daily life. *Ophthalmology*. (2002) 109:1235–42. doi: 10.1016/S0161-6420(02)01060-6
19. Maynard ML, Zele AJ, Feigl B. Mesopic Pelli-Robson contrast sensitivity and MP-1 microperimetry in healthy ageing and age-related macular degeneration. *Acta Ophthalmol*. (2016) 94:e772–78. doi: 10.1111/aos.13112
20. Chandramohan A, Stinnett SS, Petrowski JT, Schuman SG, Toth CA, Cousins SW, et al. Visual function measures in early and intermediate age-related macular degeneration. *Retina*. (2016) 36:1021–31. doi: 10.1097/IAE.0000000000001002
21. Rossouw P, Guichard M, Hatz K. Contrast sensitivity and binocular reading speed best correlating with near distance vision-related quality of life in bilateral nAMD. *Ophthalmic Physiol Opt*. (2020) 40:760–9. doi: 10.1111/opo.12736
22. Bagheri S, Lains I, Silverman RF, Kim I, Eliott D, Silva R, et al. Percentage of foveal vs. total macular geographic atrophy as a predictor of visual acuity in age-related macular degeneration. *J Vitreoretin Dis*. (2019) 3:278–82. doi: 10.1177/2474126419859454
23. Hogg R, Curry E, Muldrew A, Winder J, Stevenson M, McClure M, et al. Identification of lesion components that influence visual function in age related macular degeneration. *Br J Ophthalmol*. (2003) 87:609–14. doi: 10.1136/bjo.87.5.609
24. Keane PA, Patel PJ, Ouyang Y, Chen FK, Ikeji F, Walsh AC, et al. Effects of retinal morphology on contrast sensitivity and reading ability in neovascular age-related macular degeneration. *Invest Ophthalmol Vis Sci*. (2010) 51:5431–7. doi: 10.1167/iovs.09-4846
25. Macular Photocoagulation Study Group. Visual outcome after laser photocoagulation for subfoveal choroidal neovascularization secondary to age-related macular degeneration. The influence of initial lesion size and initial visual acuity. Macular Photocoagulation Study Group. *Arch Ophthalmol*. (1994) 112:480–8. doi: 10.1001/archopht.1994.01090160056023
26. Moutray T, Alarbi M, Mahon G, Stevenson M, Chakravarthy U. Relationships between clinical measures of visual function, fluorescein angiographic and optical coherence tomography features in patients with subfoveal choroidal neovascularisation. *Br J Ophthalmol*. (2008) 92:361–4. doi: 10.1136/bjo.2007.123976
27. Doris N, Hart PM, Chakravarthy U, McCleland J, Stevenson M, Hudson C, et al. Relation between macular morphology and visual function in patients with choroidal neovascularisation of age related macular degeneration. *Br J Ophthalmol*. (2001) 85:184–8. doi: 10.1136/bjo.85.2.184
28. Fleckenstein M, Charbel Issa P, Helb HM, Schmitz-Valckenberg S, Finger RP, Scholl HP, et al. High-resolution spectral domain-OCT imaging in geographic atrophy associated with age-related macular degeneration. *Invest Ophthalmol Vis Sci*. (2008) 49:4137–44. doi: 10.1167/iovs.08-1967

29. Hatz K, Ebneter A, Tuerksever C, Pruente C, Zinkernagel M. Repeated dexamethasone intravitreal implant for the treatment of diabetic macular oedema unresponsive to anti-vegf therapy: outcome and predictive SD-OCT features. *Ophthalmologica*. (2018) 239:205–14. doi: 10.1159/000485852
30. Schmidt-Erfurth U, Waldstein SM, Deak GG, Kundi M, Simader C. Pigment epithelial detachment followed by retinal cystoid degeneration leads to vision loss in treatment of neovascular age-related macular degeneration. *Ophthalmology*. (2015) 122:822–32. doi: 10.1016/j.ophtha.2014.11.017
31. Holopigian K, Bach M. A primer on common statistical errors in clinical ophthalmology. *Doc Ophthalmol*. (2010) 121:215–22. doi: 10.1007/s10633-010-9249-7
32. Ghoshal R, Sharanjeet-Kaur S, Mohamad Fadzil N, Mutalib HA, Ghosh S, Ngah NF, et al. Correlation between visual functions and retinal morphology in eyes with early and intermediate age-related macular degeneration. *Int J Environ Res Public Health*. (2020) 17:6379. doi: 10.3390/ijerph17176379
33. Unver YB, Yavuz GA, Bekiroglu N, Presti P, Li W, Sinclair SH. Relationships between clinical measures of visual function and anatomic changes associated with bevacizumab treatment for choroidal neovascularization in age-related macular degeneration. *Eye (Lond)*. (2009) 23:453–60. doi: 10.1038/eye.2008.349
34. Willoughby AS, Ying GS, Toth CA, Maguire MG, Burns RE, Grunwald JE, et al. Subretinal hyperreflective material in the comparison of age-related macular degeneration treatments trials. *Ophthalmology*. (2015) 122:1846–53.e5. doi: 10.1016/j.ophtha.2015.05.042
35. Ooto S, Suzuki M, Vongkulsiri S, Sato T, Spaide RF. Multimodal visual function testing in eyes with nonexudative age-related macular degeneration. *Retina*. (2015) 35:1726–34. doi: 10.1097/IAE.0000000000000608
36. Sayegh RG, Sacu S, Dunavölgyi R, Kroh ME, Roberts P, Mitsch C, et al. Geographic atrophy and foveal-sparing changes related to visual acuity in patients with dry age-related macular degeneration over time. *Am J Ophthalmol*. (2017) 179:118–28. doi: 10.1016/j.ajo.2017.03.031
37. Wolf-Schnurrbusch UE, Enzmann V, Brinkmann CK, Wolf S. Morphologic changes in patients with geographic atrophy assessed with a novel spectral OCT-SLO combination. *Invest Ophthalmol Vis Sci*. (2008) 49:3095–9. doi: 10.1167/iovs.07-1460
38. Fung AE, Lalwani GA, Rosenfeld PJ, Dubovy SR, Michels S, Feuer WJ, et al. An optical coherence tomography-guided, variable dosing regimen with intravitreal ranibizumab (Lucentis) for neovascular age-related macular degeneration. *Am J Ophthalmol*. (2007) 143:566–83. doi: 10.1016/j.ajo.2007.01.028
39. Keane PA, Sadda SR. Predicting visual outcomes for macular disease using optical coherence tomography. *Saudi J Ophthalmol*. (2011) 25:145–58. doi: 10.1016/j.sjopt.2011.01.003
40. Pauleikhoff D, Löffert D, Spital G, Radermacher M, Dohrmann J, Lommatzsch A, et al. Pigment epithelial detachment in the elderly. Clinical differentiation, natural course and pathogenetic implications. *Graefes Arch Clin Exp Ophthalmol*. (2002) 240:533–8. doi: 10.1007/s00417-002-0505-8
41. Pondorfer SG, Heinemann M, Wintergerst M, Pfau M, Strömer AL, Holz FG, et al. Detecting vision loss in intermediate age-related macular degeneration: a comparison of visual function tests. *PLoS ONE*. (2020) 15:e0231748. doi: 10.1371/journal.pone.0231748
42. Tschosik EA, Gray S, Ferrara D, Guymer R, Chroma and Spectri Study Investigators. Visual function decline resulting from geographic atrophy: results from the chroma and spectri phase 3 trials. *Ophthalmol Retina*. (2020) 4:673–88. doi: 10.1016/j.oret.2020.01.019
43. Lindner M, Pfau M, Czauderna J, Goerdt L, Schmitz-Valckenberg S, Holz FG, et al. Determinants of reading performance in eyes with foveal-sparing geographic atrophy. *Ophthalmol Retina*. (2019) 3:201–10. doi: 10.1016/j.oret.2018.11.005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 19 February 2021
doi: 10.3389/fnins.2021.624472

The Effects of 0.01% Atropine on Adult Myopes' Contrast Sensitivity

Ziyun Cheng, Jianhui Mei, Suqi Cao, Ran Zhang, Jiawei Zhou and Yuwen Wang**

State Key Laboratory of Ophthalmology, Optometry and Vision Science, School of Ophthalmology and Optometry, Affiliated Eye Hospital, Wenzhou Medical University, Wenzhou, China

Purpose: Atropine at a low concentration is considered a safe and effective treatment to mitigate myopia progression. However, the potential unwanted side effects of administering atropine at a low dose on visual functions other than best corrected visual acuity has not been investigated. In this study, we investigate the short-term (12,16, and 20 h) and long-term (1, 2, and 4 weeks) effects of 0.01% atropine (i.e., 0.1 mg/ml) on contrast sensitivity (CS) in patients with myopia.

Methods: Thirty adults (23.33 ± 2.93 years old) with myopia between -1.00 and -6.00 diopters (D), astigmatism of -1.50 D or less, and anisometropia of 1.00 D or less, participated in this prospective, masked, placebo-controlled, randomized study. The participants were randomly assigned to receive 0.01% atropine or polyvinyl alcohol eye drops once nightly to both eyes for four weeks. CS was measured binocularly at baseline and 12, 16, 20 h, 1, 2, and 4 weeks after the first use of the eye drops.

Results: There was no statistically significant differences of CS found between atropine and placebo-controlled groups in both short-term and long-term. There was no statistically significant interaction effect found between the time and group.

Conclusion: We demonstrated no significant deleterious effect of 0.01% atropine on adult myopes' CS.

Keywords: atropine, contrast sensitivity, myopia control, visual perception, myopia

OPEN ACCESS

Edited by:
Peter J. Bex,
Northeastern University, United States

Reviewed by:
Janine Dale Mendola,
McGill University, Canada
Laura Restani,
Institute of Neuroscience, National Research Council (CNR), Italy

****Correspondence:***
Jiawei Zhou
zhoujw@mail.eye.ac.cn
Yuwen Wang
wyw-0721@163.com

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *31 October 2020*
Accepted: *29 January 2021*
Published: *19 February 2021*

Citation:
Cheng Z, Mei J, Cao S, Zhang R, Zhou J and Wang Y (2021) The Effects of 0.01% Atropine on Adult Myopes' Contrast Sensitivity. Front. Neurosci. 15:624472. doi: 10.3389/fnins.2021.624472

INTRODUCTION

The prevalence of myopia has increased sharply over the last several decades (Holden et al., 2016; Hopf and Pfeiffer, 2017; Grzybowski et al., 2020). Approximately 28.3% of the world's population was myopic in 2017 (Hopf and Pfeiffer, 2017), and the rate is expected to reach 40% by 2030 (Holden et al., 2016). At the same time, the prevalence of high myopia rose dramatically from 2.7 to 5.2% between 2000 and 2020, with associated concomitant surge of sight-threatening complications (Holden et al., 2016; Ziemssen et al., 2017). A variety of clinical interventions, including multifocal spectacles, contact lenses, and pharmaceutical agents, have been put into practice trying to slow down the rapid progression of myopia. Recent studies have suggested that topical atropine might be the most effective method (Huang et al., 2016; Walline et al., 2020a). Though efficacy of administration of atropine at a low concentration (e.g., 0.01%) is weaker than high concentration (e.g., 1, 0.5%), it has fewer side effects (e.g., photophobia, blurry near vision and allergic conjunctivitis) and minimum rebound after drop cessation (Chua et al., 2006;

Chia et al., 2012; Chia et al., 2014; Chia et al., 2016; Yam et al., 2019; Yam et al., 2020). Thus, there is a preference for 0.01% atropine among pediatric ophthalmologists in myopia control (Zloto et al., 2018; Wu et al., 2019).

One side effect is the effect of myopic control interventions on patient's visual perception. The application of low-concentration atropine (e.g., 0.01, 0.05%) has been shown to less likely impair myopes' best corrected visual acuity (VA) (Chia et al., 2012; Moon and Shin, 2018; Yam et al., 2019; Yam et al., 2020). For example, Chia et al. (2012) applied 0.5, 0.1, and 0.01% atropine respectively to three groups of subjects once nightly to both eyes for 2 year and found mean best-corrected distant VA was not significantly affected by atropine use; Yam et al. (2019) administered similar protocol as Chia et al. (2012) (i.e., eye drops once nightly to both eyes) with low-concentration atropine eye drops at 0.05, 0.025, and 0.01% and placebo for a year and demonstrated VA was not significantly influenced in each group. However, VA, as one of the most important visual functions, only describes patients' visual performance in recognizing high contrast letters. A minimal to no change of VA does not mean that patients' perception on low contrast targets is intact (Joltikov et al., 2017; Zhao et al., 2017; Alahmadi et al., 2018). Moreover, there is evidence that myopic control interventions, such as wearing multifocal spectacles or contact lenses, reduce myopes' low-contrast vision acuity compared with single vision lens (Lu et al., 2020; Walline et al., 2020b). Therefore, it is necessary to further assess the effect of low-concentration atropine on myopes' other visual functions before widespread application in clinical practice.

Contrast sensitivity function (CSF) provides a much more comprehensive assessment of spatial vision at different contrast conditions and a variety of spatial frequencies and has been used to evaluate and screen a variety of visual disorders (Chatzistefanou et al., 2005; Joltikov et al., 2017; Alahmadi et al., 2018). In animal models, it has been found that one drop of 1% atropine can actually increase contrast sensitivity (CS), at least, at low spatial frequency (e.g., 0.03 and 0.20 cycles per degree) in mice and chicks, measured by optomotor paradigm (Diether and Schaeffel, 1999; Schmucker and Schaeffel, 2006). In the mice, the contrast threshold of 0.03 cycles per degree was down to 16% from 24% (Schmucker and Schaeffel, 2006). Although there has been evidence that sustained penalization of the fellow eye with atropine is one of the standard treatments for amblyopia, improving their CS after the therapy (Menon et al., 2008; DeSantis, 2014), there has been, to our best knowledge, no study about the effects of atropine on myopes' CS. Previous experiments demonstrate that best corrected VA of myopia is not affected after the application of low-concentration atropine (Chia et al., 2016; Yam et al., 2019; Yam et al., 2020). In the current study, we are interested to know whether atropine could influence myopes' CS. To address this, we have applied 0.01% atropine or polyvinyl alcohol in two groups of myopes for four weeks and measured observers' CS before and at 12, 16, 20 h and 1, 2, 4 weeks after the first use of the eye drops. We have evaluated both short-term (e.g., within one day) and long-term (e.g., weeks) influences of low-concentration atropine application on myopes' visual perception.

MATERIALS AND METHODS

Participants

Thirty adults (23.33 ± 2.93 years old; 18 females) with myopic refraction between −1.0 D and −6.0 D, astigmatism of less than 1.5 diopters (D) in both eyes, and anisometropia of less than 1.0 D were enrolled. All participants have normal or corrected-to-normal vision (logMAR acuity ≤0.00). Excluded were those with ocular pathology (e.g., amblyopia, strabismus, glaucoma, conjunctivitis), ocular surgery history, abnormal binocular function, allergy to atropine, systemic ill health (e.g., diabetes or autoimmune diseases), or previous use of atropine or orthokeratology. All participants were naive to the purpose of the study. The study and protocol conformed to the tenets of the Declaration of Helsinki and was approved by the Ethics Committee of Eye hospital of Wenzhou Medical University. Written informed constant was obtained from each participant.

Design

Participants were allocated randomly to the atropine or the placebo-control group according to a computer-generated randomization list, and respectively received one drop of 0.01% atropine (0.04 mg/0.4 ml unit-concentration, preservative free, Shengyang Xinqi Eye Hospital Co., Ltd.) or polyvinyl alcohol (0.5 ml unit-concentration, preservative free, Xindongshengji Co., Ltd., Taiwan) once nightly (after 8:00 pm) in both eyes for 4 weeks. The eye drops were given by the authors (Z Cheng and J Mei), and the participants were unaware of which eye drop was given. CS was binocularly measured with individual's optimal spectacle correction in his/her each visit. Before the baseline test in the first visit, participants had been given a chance to be familiarized with the test. Follow-up visits were then scheduled after 12, 16, 20 h, 1, 2, and 4 weeks from their first use of the eye drops (**Figure 1**).

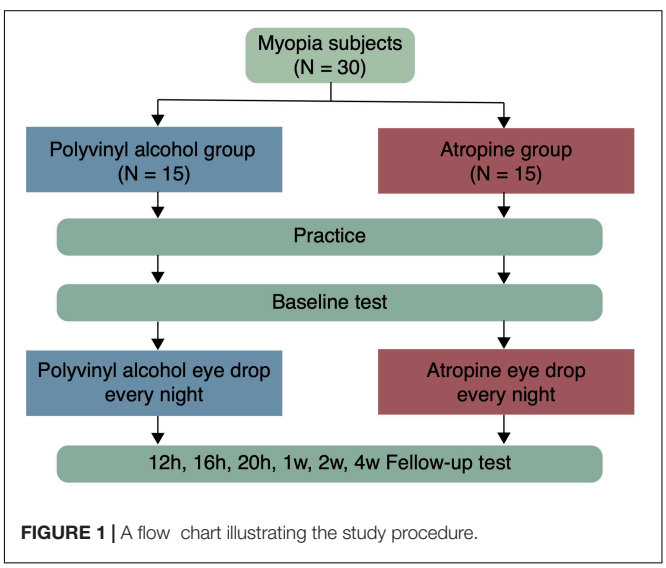


FIGURE 1 | A flow chart illustrating the study procedure.

Apparatus

All tests were performed on a visual function test workstation (Zhishiyuan, JH-P02, Model NO.102JST190828001, Jiangsu Juehua Medical Technology Co., Ltd.) in a dark room. The workstation consists of a PC and a display. Stimuli were generated and controlled by the PC, and presented on the display. The display was GAMMA-corrected, has a resolution of 2560 x 1440 pixels, a refresh rate of 60 Hz, and an average luminance of 74.5 cd/m^2.

Study Procedures

The stimulus was a sinusoidal grating with spatial frequency of 1.5, 3, 6, 12, 18, 24 cycles per degree (cpd) and subtended 3.0° × 3.0° at a viewing distance of 2 m. In order to reduce the edge effect, 0.5-degree Gaussian ramp was added around the stimulus. Before the start of the test, there was an introduction about the entire experimental process, stimuli and task. A brief beep prompted the start of the trial, together with presentation of a crosshair (3.0 × 3.0°) to indicate the location of the stimulus. After 150 ms, the cross disappeared and stimulus grating of vertical or horizontal orientation (with equal probability) were displayed for 167 ms. A blank background with mean luminance (74.5 cd/m^2) was then displayed and participants were asked to the orientation with corresponding arrow key in the keyboard. Inter-trial interval was 800 ms (**Figure 2**). A Psi method (Kontsevich and Tyler, 1999) controlled the grating contrast and estimated contrast threshold that corresponds to 80.3% correct for each spatial frequency, separately. CS was calculated as the reciprocal of contrast threshold. There were 270 trials in total, with 45 trials in each spatial frequency. The work station adopted the bit-stealing method (Tyler, 1997) to achieve high-precision gray-scale stimulation.

Data Analysis

A Shapiro–Wilks test was performed on each dataset to evaluate normality of our dataset. The demographic characteristics (e.g., age, gender, and refractive error) between the two treatment groups were evaluated by two-sample *t*-test for normality dataset, Wilcoxon rank sum test for non-normally distributed dataset, and Chi-square test for qualitative dataset. A one-way analysis of variance (ANOVA) was used to compare the CSF at baseline between groups. A repeated-measure ANOVA was used to evaluate the main effects of group, time and the interaction between them after treatment. The area under the log CSF curve (AULCSF), which is a widely used summary metric of the CSF function (Joltikov et al., 2017; Alahmadi et al., 2018; Zheng et al., 2018), was calculated by using the trapezoid method with CS in logarithmic values at 1.5–24 cpd (Rosén et al., 2014), and compared between groups and follow-up time sessions. *P* value <0.05 was considered statistically significant. Statistical analyses were performed using SPSS 26.0 software (IBM Crop., Armonk, NY. Released 2019).

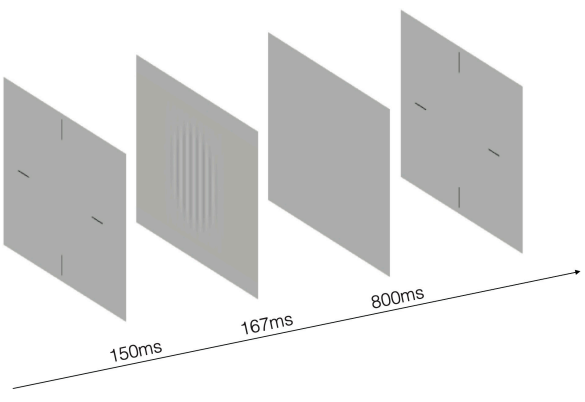


FIGURE 2 | An illustration of the test. The stimulus was a sinusoidal grating of vertical or horizontal orientation with spatial frequency of 1.5, 3, 6, 12, 18, 24 cycles per degree. A 0.5-degree Gaussian ramp was added around the stimulus to reduce the edge effect.

RESULTS

At the initial pretreatment visit, there were no significant differences between the groups in mean age, gender, and refractive error (mean age, $P = 0.407$; gender, $P = 1.000$; refractive error: OD, $P = 0.512$; OS, $P = 0.880$). The individual information of participants was exhibited in **Table 1**. One-way ANOVA, with spatial frequency (six levels) selected as within-subject factors, and treatment group (two levels) selected as between-subject factor, showed there was no significant differences on CS at baseline between the groups [$F(1,168) = 2.011$, $P = 0.158$] and the interaction between intervention group and spatial frequency was also not significant [$F(5,168) = 0.214$, $P = 0.956$]. Likewise, there were no significant differences between the groups in terms of AULCSF [$t(28) = -0.720$, $P = 0.477$] as demonstrated by independent-sample *t*-test.

To assess whether there was any effect of atropine on CS, we firstly conducted a three-way repeated-measure ANOVA, with spatial frequency (six levels) and time (seven levels) selected as within-subject factors, and treatment group (two levels) selected as between-subject factor. We found that there was no significant difference between two groups [$F(1,28) = 0.018$, $P = 0.895$], nor significant interaction effect of treatment group and time [$F(6,168) = 1.355$, $P = 0.254$].

We then conducted a two-way repeated-measure ANOVA (two within-subject: time of measurements, seven levels; spatial frequency, six levels) on each group to figure out whether there is any time cumulative effect of atropine compared with placebo-controlled group. The two-way repeated-measure ANOVA showed that there was no significant change on CS [0.01% atropine: $F(6,84) = 2.071$, $P = 0.114$; polyvinyl alcohol: $F(6,84) = 1.462$, $P = 0.201$].

We next evaluated the short-term effect (within one day) of low-concentration atropine (**Figure 3A**). Specifically, we conducted a two-way repeated-measure ANOVA for results of baseline and 12, 16, 20 h follow-up tests for the two groups. We found that there was no significant change on CS before and after using of the eye drops for each group: 0.01% atropine: $F(3,42) = 2.036$, $P = 0.123$; polyvinyl alcohol: $F(3,42) = 0.911$, $P = 0.444$. To assess whether there is any

TABLE 1 | Clinical details of the participants.

	Polyvinyl alcohol group				Atropine group		
	Age/Sex	**Refraction (OD/OS)**	**logMAR VA (OU)**		**Age/Sex**	**Refraction (OD/OS)**	**logMAR VA (OU)**
S1	35/M	−4.00/−0.50 × 176 −4.25/−0.50 × 178	0.00	*S16*	23/F	−2.00/−1.25 × 165 −1.75/-1.25 × 5	0.00
S2	25/F	−2.00/−0.50 × 105 −1.75/−0.25 × 40	−0.10	*S17*	23/F	−6.00 −5.25	0.00
S3	25/F	−5.25 −5.00	0.00	*S18*	24/F	−3.75/−0.75 × 100 −4.25/−1.00 × 25	0.00
S4	24/F	−4.75 −4.50	0.00	*S19*	22/F	−1.25/−0.25 × 140 −2.25/−0.50 × 165	0.00
S5	22/F	−3.50 −3.00	0.00	*S20*	23/F	−3.50 −4.00/−0.75 × 160	0.00
S6	23/F	−4.00/−1.50 × 180 −3.75/−1.25 × 180	0.00	*S21*	25/F	−2.25/−0.75 × 55 −1.75/−1.00 × 131	0.00
S7	19/M	−5.25 −5.00	0.00	*S22*	20/M	−5.00/−0.50 × 180 −4.50/−0.75 × 170	0.00
S8	23/M	−6.00 −6.00	0.00	*S23*	23/M	−4.25/−0.25 × 60 −4.25	−0.10
S9	22/F	−2.00/−1.00 × 40 −1.00/−1.00 × 140	−0.10	*S24*	19/M	−4.00/−1.00 × 170 −4.00/−1.00 × 175	0.00
S10	25/F	−3.75 −3.25	0.00	*S25*	21/M	−3.25/−0.75 ×33 −2.50/−1.25 × 147	0.00
S11	19/M	−4.50/−1.00 × 90 −5.00/−1.50 × 180	0.00	*S26*	23/F	−5.50/−0.75 × 180 −5.75/−1.00 × 165	−0.10
S12	25/F	−4.50 −4.25	0.00	*S27*	23/F	−4.25/−1.00 × 5 −4.50/−1.00 × 165	0.00
S13	22/M	−2.00/−1.00 × 100 −2.25/−1.00 × 140	0.00	*S28*	25/M	−1.25 −2.25	0.00
S14	27/F	−3.25 −3.25	0.00	*S29*	23/M	−2.75/−0.50 × 180 −1.75	0.00
S15	23/M	−1.75/−0.50 × 10 −1.00/−0.25 × 150	−0.10	*S30*	24/F	−2.25/−0.50 × 90 −1.25/−0.50 × 180	0.00

M, male; F, female; OD, oculus dexter (right eye); OS, oculus sinister (left eye); OU, oculus unati (binocular); VA, visual acuity.

difference between atropine group and placebo-controlled group, we then conducted a multi-factor repeated-measure ANOVA, and we found that there was no significant difference between two groups [$F(1,28) = 0.009$, $P = 0.924$], and the interaction effect of treatment group and time was also not significant [$F(3,84) = 2.401$, $P = 0.087$].

We conducted extra analysis for AULCSF dataset. The averaged and individual AULCSF as a function of follow-up time sessions are plotted in **Figure 4A**. A two-way repeated-measure ANOVA was used, which indicated no significant difference between two groups [$F(1,28) = 0.243$, $P = 0.626$], and no significant interaction effect of treatment group and time [$F(3,84) = 0.909$, $P = 0.441$]. In short, there was no significant effect of 0.01% atropine on CS in short-term within one day.

To assess whether there is long-term (e.g., a month) cumulative effect of time, we conducted a two-way repeated-measure ANOVA for results of baseline and 1, 2, and 4 weeks follow-up tests on each group (**Figure 3B**). The results showed there was no significant cumulative effect of 0.01% atropine on CS [atropine group: $F(3,42) = 0.389$, $P = 0.656$; polyvinyl alcohol: $F(3,42) = 2.273$, $P = 0.094$]. And multi-factor repeated-measure ANOVA suggested there was no significant difference between two groups [$F(1,28) = 0.127$, $P = 0.724$]. The interaction effect of treatment group and time was also not significant [$F(3,84) = 1.094$, $P = 0.347$]. The averaged and individual AULCSF as a function of time are plotted in **Figure 4B**. A multi-factor repeated-measure ANOVA was used, which revealed there was no significant difference between two groups [$F(1,28) = 0.127$, $P = 0.724$]. The interaction effect of treatment group and time was also not significant [$F(3,84) = 1.094$, $P = 0.347$]. In other word, there was no significant difference of CS found before and after using 0.01% atropine at a long-term (i.e., weeks up to one month).

DISCUSSION

In this study, we investigated whether the application of 0.01% atropine could influence CS in the short-term and long-term. Our results show that the effect of 0.01% atropine on adult myopes' visual perception is minimal.

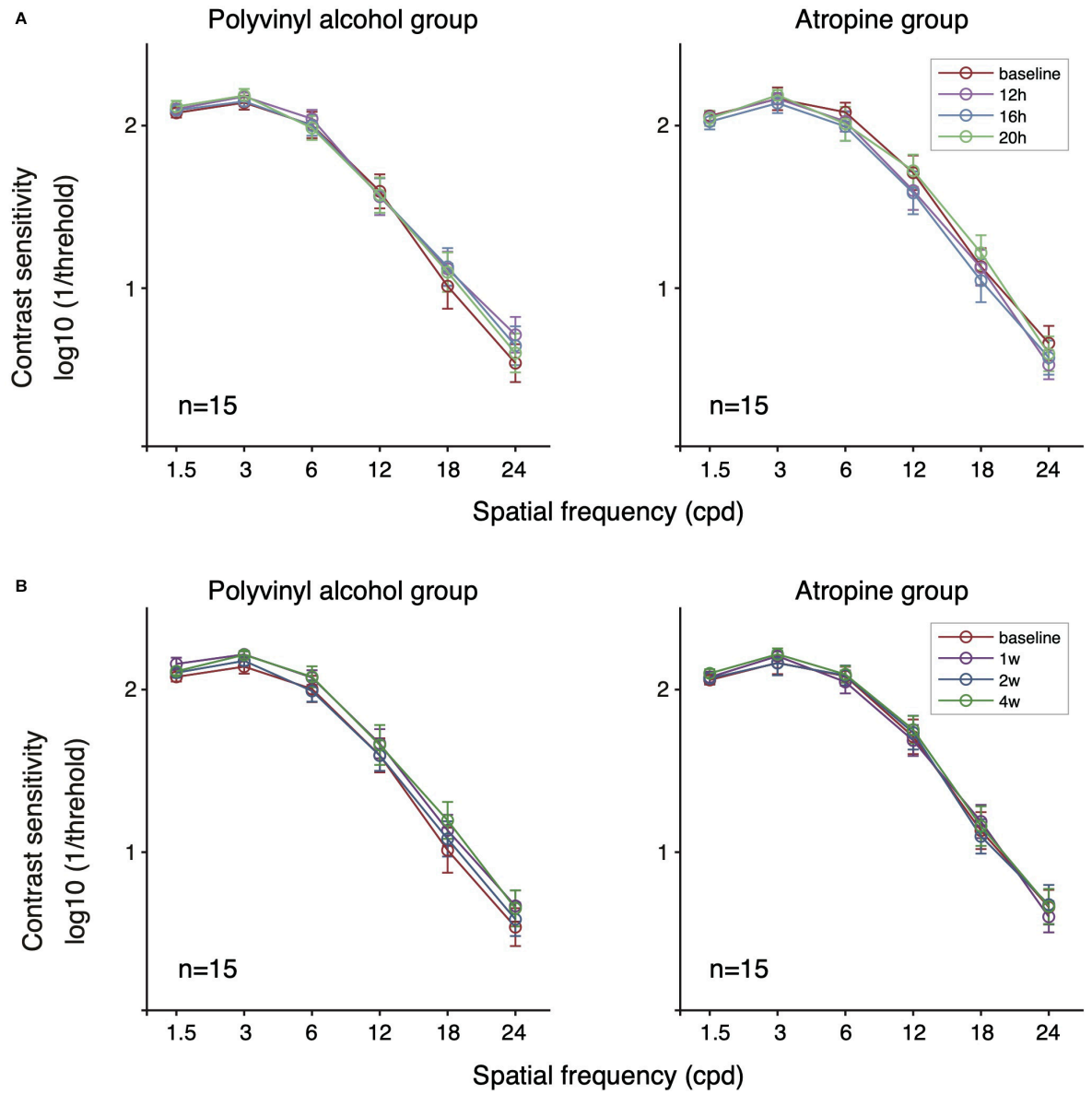


FIGURE 3 | Contrast sensitivity function of the two groups following short-term **(A)** and long-term **(B)** administration of 0.01% atropine or polyvinyl alcohol. There was no significant effect of time in the polyvinyl alcohol group and the atropine group (respectively, $P = 0.444$, $P = 0.123$ at short-term; $P = 0.094$, $P = 0.656$ at long-term). There was no significant difference between the polyvinyl alcohol group and the atropine group ($P = 0.924$ and $P = 0.724$ for short- and long-term, respectively).

Previous studies of atropine application in myopia control have used VA as an index of visual perception and found no significant effect of 0.01% atropine on distant VA compared to placebo-controlled group as well as no significant time cumulative effect over two years (Chia et al., 2012; Yam et al., 2019; Yam et al., 2020). Although VA is the most common functional endpoint, it has limitations. VA only represents the ability of distinguishing fine details at high contrast, although objects in the real world involve a wide range of luminance and contrast levels. CS was shown to better reflect the ability of detecting and identifying objects in day-to-day experience than that of VA (Owsley and Sloane, 1987; Roark and Stringham, 2019). Moreover, CS is a much more sensitive measure than VA when diseases, degeneration or other changes occur in the visual system (Joltikov et al., 2017; Zhao et al., 2017; Alahmadi et al., 2018; Xiong et al., 2020). To illustrate, there are cases with relatively normal VA but different extents of CS deficits across various ocular pathologies (Xiong et al., 2020). Recent studies demonstrate that myopia control measures, e.g., multifocal spectacles and contact lenses, could lower participants' low contrast VA, despite high contrast VA remaning intact (Przekoracka et al., 2020; Walline et al., 2020b). In the current study, we confirm that both short-term (e.g., within a day) and long-term (e.g., within a month) administration of

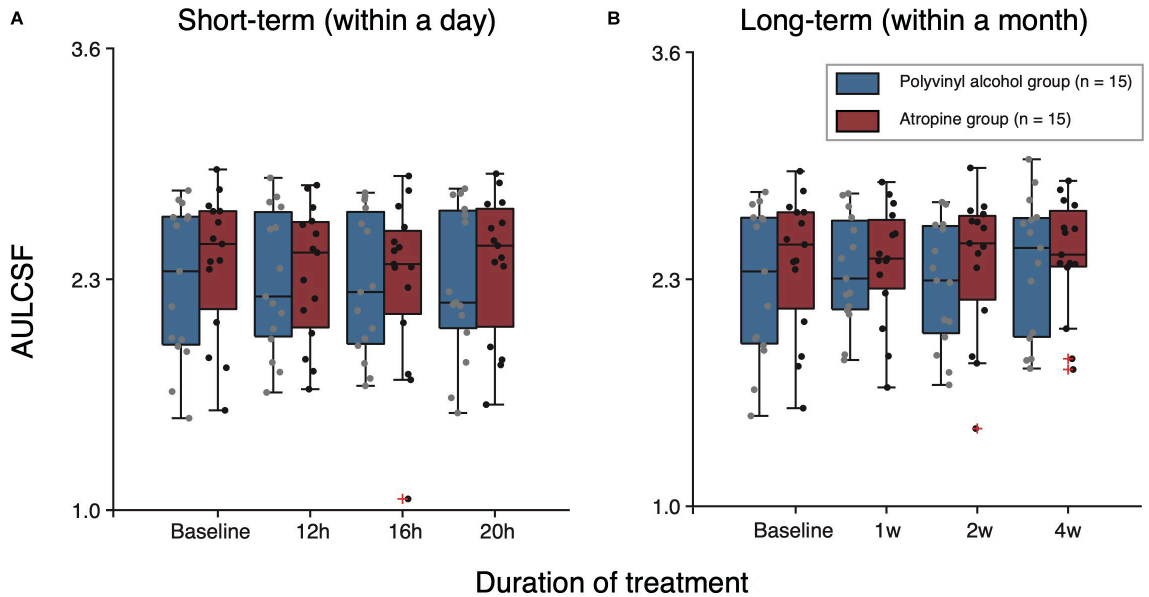


FIGURE 4 | Boxplots of AULCSF following short-term **(A)** and long-term **(B)** administration of polyvinyl alcohol (Blue) and atropine (red). The black solid line within each box represents the median. The box represents the interquartile range (IQR) of the data (25th to the 75th percentile). The whisker represents 1.5 × IQR either above the third quartile or below the first quartile. Crosses represent outliers.

0.01% atropine didn't have significant detrimental effect on CS in adult myopes.

Atropine is a non-selective antagonist of the muscarinic acetylcholine receptor, which is widely distributed in ocular tissues, including cornea, iris, ciliary body and ciliary muscles, epithelium of crystalline lens, retina, choroid, and sclera (Friedman et al., 1988; Gil et al., 1997; Collison et al., 2000; Qu et al., 2006; Barathi et al., 2009). Although the anti-myopia mechanism of atropine is not fully understood, recent experiments have shown that atropine may exert its myopia-protective effect mainly through muscarinic receptors on retina and sclera, stopping the remodeling and thinning of sclera and the consequent axial lengthening of the eye, even at a low concentration (Sánchez-González et al., 2020; Upadhyay and Beuerman, 2020). Moreover, topical atropine could cause pupil dilation, decrease of accommodation amplitude and change in corneal curvature, lens thickness, anterior chamber depth and vitreous chamber depth (Qu et al., 2006; Kumaran et al., 2015; Goldberg and Rucker, 2016; Yam et al., 2019; Upadhyay and Beuerman, 2020). It was proved that CS is affected by optical factors (e.g., aberration based on pupil size and intraocular forward scattering) as well as retinal and brain processing (Radhakrishnan et al., 2004; Yamaguchi et al., 2011; Kamiya et al., 2014; Karatepe et al., 2017). In animal experiments in myopia control, CS at low spatial frequency was enhanced after the administration of atropine (Diether and Schaeffel, 1999; Schmucker and Schaeffel, 2006). Schmucker and Schaeffel (2006) thought the increase of CS in mice was possibly due to the dilatation of the pupil with atropine and the brighter retinal image, despite the larger pupil also resulting in a decline in optical quality of retinal image. As a treatment of amblyopia, it has been found that the CS of amblyopic patients improves after the administration of atropine by suppressing fellow eye and meliorating VA of amblyopia eye (Menon et al., 2008; DeSantis, 2014). The anti-myopia mechanism of atropine is totally different from the rationale for amblyopia therapy. We were interested in whether low-concentration atropine could influence myopes' CS. Our results show that no significant differences existed in CSF before and after both short- and long-term administration of 0.01% atropine in adult myopes.

The results agrees with Anders et al.'s (2019) study, which showed there was no pronounced impact of 0.01% atropine on retinal processing, as reflected by the pattern electroretinogram (PERG). Moreover, Khanal et al. (2019) found that atropine may exert its anti-myopia effect mainly through affecting the responses to myopia defocus in peripheral retinal instead of central retinal as demonstrated by the global flash multifocal electroretinogram (gmfERG). And there was no significant change on VA after the application of 0.01% atropine (Chia et al., 2012; Yam et al., 2019; Yam et al., 2020). Our study, together with these previous reports, suggests that atropine might produce minimal effect on macular visual functions.

Another possible reason for the minimal effect is that the biochemical and structural changes in ocular system caused by one-month administration of 0.01% atropine was too slight to be detected. Yet there were significant dose-related effects of atropine in axial length, accommodation and pupil diameter (Chia et al., 2012; Yam et al., 2019; Fu et al., 2020; Li et al., 2020; Yam et al., 2020). In the LAMP study (Yam et al., 2019), the change of mesopic pupil size was only 0.23 mm in 0.01% atropine group compared to 0.43 and 0.58 mm in 0.025 and 0.05% groups after using eye drops for a year. Hence, a future study should investigate whether higher concentration (e.g., 0.025, 0.05%) of atropine affects CS.

In conclusion, our results indicate that 0.01% atropine has minimal deleterious effect on patient's CS. It should be noted that our research was conducted on adult myopia, while topical atropine is mainly used by preschool and school-age pediatric, who are at a stage of rapid myopia progression (Lin et al., 2018). A reduced visual perception (e.g., VA, CS) would be a disadvantage for kids, as it might produce amblyopia (McKee et al., 2003; Levi, 2020). It is an interesting question that whether a myopia control strategy that slightly decreases patients' visual perception would produce a worse or better myopia control effect than those strategies that didn't change patients' visual perception (Lau et al., 2020; Vincent et al., 2020). We have no clear answer to this question at this stage, since our knowledge of the change of visual functions following myopia control is limited. Further investigations on children myopia are needed to better show the effect of 0.01% atropine on their perception. Also, we only apply the atropine for one month. According to the results in **Figure 4**, there seem to be a trend, albeit non-significant, of a small reduction in sensitivity. It would be interesting to evaluate whether longer-term use (e.g., a year or longer) of atropine would affect the visual functions, since myopia control measures are typically prescribed until myopia progression is slowed down in late adolescent period (i.e., 15–18 years old) (Zadnik et al., 2015; Wu et al., 2019).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Affiliated eye hospital, Wenzhou Medical University. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ZC, JM, YW, and JZ conceived the experiments. ZC, JM, SC, and RZ performed the experiments. ZC, JM, and JZ analyzed the data and interpreted the data. ZC, YW, and JZ wrote the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

This work was supported by the National Key R&D Program of China (2016YFB0401203).

ACKNOWLEDGMENTS

We thank Dr. Chang-Bing Huang and Mr. Seung Hyun (Sam) Min for their help in manuscript editing.

REFERENCES

Alahmadi, B., Omari, A., Abalem, M., Andrews, C., Schlegel, D., Branham, K., et al. (2018). Contrast sensitivity deficits in patients with mutation-proven inherited retinal degenerations. *BMC Ophthalmol.* 18:313. doi: 10.1186/s12886-018-0982-0

Anders, L., Heinrich, S., Lagrèze, W., and Joachimsen, L. (2019). Little effect of 0.01% atropine eye drops as used in myopia prevention on the pattern electroretinogram. *Docum. Ophthalmol.* 138, 85–95. doi: 10.1007/s10633-019-09671-0

Barathi, V., Weon, S., and Beuerman, R. (2009). Expression of muscarinic receptors in human and mouse sclera and their role in the regulation of scleral fibroblasts proliferation. *Mol. Vis.* 15, 1277–1293.

Chatzistefanou, K. I., Theodossiadis, G. P., Damanakis, A. G., Ladas, I. D., Moschos, M. N., and Chimonidou, E. (2005). Contrast sensitivity in amblyopia: the fellow eye of untreated and successfully treated amblyopes. *J. Am. Assoc. Pediatr. Ophthalmol. Strab.* 9, 468–474. doi: 10.1016/j.jaapos.2005.05.002

Chia, A., Chua, W.-H., Cheung, Y.-B., Wong, W.-L., Lingham, A., Fong, A., et al. (2012). Atropine for the treatment of childhood myopia: safety and efficacy of 0.5%, 0.1%, and 0.01% doses (Atropine for the Treatment of Myopia 2). *Ophthalmology* 119, 347–354. doi: 10.1016/j.ophtha.2011.07.031

Chia, A., Chua, W. H., Wen, L., Fong, A., Goon, Y. Y., and Tan, D. (2014). Atropine for the treatment of childhood myopia: changes after stopping atropine 0.01%, 0.1% and 0.5%. *Am. J. Ophthalmol.* 157, 451–457.e1.

Chia, A., Lu, Q., and Tan, D. (2016). Five-year clinical trial on atropine for the treatment of myopia 2: myopia control with atropine 0.01% Eyedrops. *Ophthalmology* 123, 391–399. doi: 10.1016/j.ophtha.2015.07.004

Chua, W.-H., Balakrishnan, V., Chan, Y.-H., Tong, L., Ling, Y., Quah, B.-L., et al. (2006). Atropine for the treatment of childhood myopia. *Ophthalmology* 113, 2285–2291.

Collison, D., Coleman, R., James, R., Carey, J., and Duncan, G. (2000). Characterization of muscarinic receptors in human lens cells by pharmacologic and molecular techniques. *Invest. Ophthalmol. Vis. Sci.* 41, 2633–2641.

DeSantis, D. (2014). Amblyopia. *Pediatr. Clin. North Am.* 61, 505–518.

Diether, S., and Schaeffel, F. (1999). Long-term changes in retinal contrast sensitivity in chicks from frosted occluders and drugs: relations to myopia? *Vis. Res.* 39, 2499–2510. doi: 10.1016/s0042-6989(99)00005-x

Friedman, Z., Hackett, S., and Campochiaro, P. (1988). Human retinal pigment epithelial cells possess muscarinic receptors coupled to calcium mobilization. *Brain Res.* 446, 11–16. doi: 10.1016/0006-8993(88)91291-7

Fu, A., Stapleton, F., Wei, L., Wang, W., Zhao, B., Watt, K., et al. (2020). Effect of low-dose atropine on myopia progression, pupil diameter and accommodative amplitude: low-dose atropine and myopia progression. *Br. J. Ophthalmol.* 104, 1535–1541.

Gil, D., Krauss, H., Bogardus, A., and Woldemussie, E. (1997). Muscarinic receptor subtypes in human iris-ciliary body measured by immunoprecipitation. *Invest. Ophthalmol. Vis. Sci.* 38, 1434–1442.

Goldberg, L., and Rucker, F. (2016). Opposing effects of atropine and timolol on the color and luminance emmetropization mechanisms in chicks. *Vis. Res.* 122, 1–11. doi: 10.1016/j.visres.2016.03.001

Grzybowski, A., Kanclerz, P., Tsubota, K., Lanca, C., and Saw, S. (2020). A review on the epidemiology of myopia in school children worldwide. *BMC Ophthalmol.* 20:27. doi: 10.1186/s12886-019-1220-0

Holden, B., Fricke, T., Wilson, D., Jong, M., Naidoo, K., Sankaridurg, P., et al. (2016). Global prevalence of myopia and high myopia and temporal trends from

2000 through 2050. *Ophthalmology* 123, 1036–1042. doi: 10.1016/j.ophtha.2016.01.006

Hopf, S., and Pfeiffer, N. (2017). [Epidemiology of myopia]. *Ophthalmologe* 114, 20–23.

Huang, J., Wen, D., Wang, Q., Mcalinden, C., Flitcroft, I., Chen, H., et al. (2016). Efficacy comparison of 16 interventions for myopia control in children: a network meta-analysis. *Ophthalmology* 123, 697–708. doi: 10.1016/j.ophtha.2015.11.010

Joltikov, K., De Castro, V., Davila, J., Anand, R., Khan, S., Farbman, N., et al. (2017). Multidimensional functional and structural evaluation reveals neuroretinal impairment in early diabetic retinopathy. *Invest. Ophthalmol. Vis. Sci.* 58, BIO277–BIO290.

Kamiya, K., Shimizu, K., Iijima, A., and Kobashi, H. (2014). Factors influencing contrast sensitivity function in myopic eyes. *PLoS One* 9:e113562. doi: 10.1371/journal.pone.0113562

Karatepe, A., Köse, S., and Eğrilmez, S. (2017). Factors affecting contrast sensitivity in healthy individuals: a pilot study. *Turkish J. Ophthalmol.* 47, 80–84. doi: 10.4274/tjo.93763

Khanal, S., Turnbull, P., Lee, N., and Phillips, J. (2019). The effect of atropine on human global flash mfERG responses to retinal defocus. *Invest. Ophthalmol. Vis. Sci.* 60, 218–225. doi: 10.1167/iovs.18-24600

Kontsevich, L., and Tyler, C. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vis. Res.* 39, 2729–2737. doi: 10.1016/s0042-6989(98)00285-5

Kumaran, A., Htoon, H., Tan, D., and Chia, A. (2015). Analysis of changes in refraction and biometry of atropine- and placebo-treated eyes. *Invest. Ophthalmol. Vis. Sci.* 56, 5650–5655. doi: 10.1167/iovs.14-14716

Lau, J., Vincent, S., Cheung, S., and Cho, P. (2020). Higher-order aberrations and axial elongation in myopic children treated with orthokeratology. *Invest. Ophthalmol. Vis. Sci.* 61:22. doi: 10.1167/iovs.61.2.22

Levi, D. M. (2020). Rethinking amblyopia 2020. *Vision Res.* 176, 118–129. doi: 10.1016/j.visres.2020.07.014

Li, F., Kam, K., Zhang, Y., Tang, S., Young, A., Chen, L., et al. (2020). Differential effects on ocular biometrics by 0.05%, 0.025%, and 0.01% atropine: low-concentration atropine for myopia progression study. *Ophthalmology* 127, 1603–1611.

Lin, H., Long, E., Ding, X., Diao, H., Chen, Z., Liu, R., et al. (2018). Prediction of myopia development among Chinese school-aged children using refraction data from electronic medical records: a retrospective, multicentre machine learning study. *PLoS Med.* 15:e1002674. doi: 10.1371/journal.pmed.1002674

Lu, Y., Lin, Z., Wen, L., Gao, W., Pan, L., Li, X., et al. (2020). The adaptation and acceptance of defocus incorporated multiple segment lens for Chinese children. *Am. J. Ophthalmol.* 211, 207–216. doi: 10.1016/j.ajo.2019.12.002

McKee, S., Levi, D., and Movshon, J. (2003). The pattern of visual deficits in amblyopia. *J. Vision* 3, 380–405.

Menon, V., Shailesh, G., Sharma, P., and Saxena, R. (2008). Clinical trial of patching versus atropine penalization for the treatment of anisometropic amblyopia in older children. *J. AAPOS* 12, 493–497. doi: 10.1016/j.jaapos.2008.03.006

Moon, J., and Shin, S. (2018). The diluted atropine for inhibition of myopia progression in Korean children. *Int. J. Ophthalmol.* 11, 1657–1662.

Owsley, C., and Sloane, M. E. (1987). Contrast sensitivity, acuity, and the perception of 'real-world' targets. *Br. J. Ophthalmol.* 71, 791–796. doi: 10.1136/bjo.71.10.791

Przekoracka, K., Michalak, K., Olszewski, J., Zeri, F., Michalski, A., Paluch, J., et al. (2020). Contrast sensitivity and visual acuity in subjects wearing multifocal contact lenses with high additions designed for myopia progression control. *Cont. Lens. Anterior Eye* 43, 33–39. doi: 10.1016/j.clae.2019.12.002

Qu, J., Zhou, X., Xie, R., Zhang, L., Hu, D., Li, H., et al. (2006). The presence of m1 to m5 receptors in human sclera: evidence of the sclera as a potential site of action for muscarinic receptor antagonists. *Curr. Eye Res.* 31, 587–597. doi: 10.1080/02713680600770609

Radhakrishnan, H., Pardhan, S., Calver, R., and O'leary, D. (2004). Effect of positive and negative defocus on contrast sensitivity in myopes and non-myopes. *Vis. Res.* 44, 1869–1878. doi: 10.1016/j.visres.2004.03.007

Roark, M., and Stringham, J. (2019). Visual performance in the "Real World": contrast sensitivity, visual acuity, and effects of macular carotenoids. *Mol. Nutr. Food Res.* 63:e1801053.

Rosén, R., Lundström, L., Venkataraman, A., Winter, S., and Unsbo, P. (2014). Quick contrast sensitivity measurements in the periphery. *J. Vis.* 14:3. doi: 10.1167/14.8.3

Sánchez-González, J., De-Hita-Cantalejo, C., Baustita-Llamas, M., Sánchez-González, M., and Capote-Puente, R. (2020). The combined effect of low-dose atropine with orthokeratology in pediatric myopia control: review of the current treatment status for myopia. *J. Clin. Med.* 9:2371. doi: 10.3390/jcm9082371

Schmucker, C., and Schaeffel, F. (2006). Contrast sensitivity of wildtype mice wearing diffusers or spectacle lenses, and the effect of atropine. *Vis. Res.* 46, 678–687. doi: 10.1016/j.visres.2005.04.015

Tyler, C. W. (1997). Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spat. Vis.* 10, 369–377. doi: 10.1163/156856897x00294

Upadhyay, A., and Beuerman, R. (2020). Biological mechanisms of atropine control of myopia. *Eye Contact Lens* 46, 129–135. doi: 10.1097/icl.0000000000000677

Vincent, S., Tan, Q., Ng, A., Cheng, G., Woo, V., and Cho, P. (2020). Higher order aberrations and axial elongation in combined 0.01% atropine with orthokeratology for myopia control. *Ophthalmic Physiol. Optics* 40, 728–737. doi: 10.1111/opo.12730

Walline, J., Lindsley, K., Vedula, S., Cotter, S., Mutti, D., Ng, S., et al. (2020a). Interventions to slow progression of myopia in children. *Cochrane Datab. Syst. Rev.* 1:CD004916.

Walline, J., Walker, M., Mutti, D., Jones-Jordan, L., Sinnott, L., Giannoni, A., et al. (2020b). Effect of high add power, medium add power, or single-vision contact lenses on myopia progression in children: the BLINK randomized clinical trial. *JAMA* 324, 571–580. doi: 10.1001/jama.2020.10834

Wu, P., Chuang, M., Choi, J., Chen, H., Wu, G., Ohno-Matsui, K., et al. (2019). Update in myopia and treatment strategy of atropine use in myopia control. *Eye (London, England)* 33, 3–13. doi: 10.1038/s41433-018-0139-7

Xiong, Y.-Z., Kwon, M., Bittner, A. K., Virgili, G., Giacomelli, G., and Legge, G. E. (2020). Relationship between acuity and contrast sensitivity: differences due to eye disease. *Invest. Ophthalmol. Vis. Sci.* 61:40. doi: 10.1167/iovs.61.6.40

Yam, J., Li, F., Zhang, X., Tang, S., Yip, B., Kam, K., et al. (2020). Two-year clinical trial of the low-concentration atropine for myopia progression (LAMP) study: phase 2 report. *Ophthalmology* 127, 910–919. doi: 10.1016/j.ophtha.2019.12.011

Yam, J. C., Jiang, Y., Tang, S. M., Law, A. K. P., Chan, J. J., Wong, E., et al. (2019). Low-concentration atropine for myopia progression (LAMP) study: a randomized, double-blinded, placebo-controlled trial of 0.05%, 0.025%, and 0.01% atropine eye drops in myopia control. *Ophthalmology* 126, 113–124.

Yamaguchi, T., Negishi, K., Ohnuma, K., and Tsubota, K. (2011). Correlation between contrast sensitivity and higher-order aberration based on pupil diameter after cataract surgery. *Clin. Ophthalmol.* 5, 1701–1707. doi: 10.2147/opth.s21819

Zadnik, K., Sinnott, L., Cotter, S., Jones-Jordan, L., Kleinstein, R., Manny, R., et al. (2015). Prediction of juvenile-onset myopia. *JAMA Ophthalmol.* 133, 683–689. doi: 10.1001/jamaophthalmol.2015.0471

Zhao, W., Jia, W. L., Chen, G., Luo, Y., Lin, B., He, Q., et al. (2017). A complete investigation of monocular and binocular functions in clinically treated amblyopia. *Sci. Rep.* 7:10682.

Zheng, H., Wang, C., Cui, R., He, X., Shen, M., Lesmes, L. A., et al. (2018). Measuring the contrast sensitivity function using the qCSF method with 10 digits. *Transl. Vis. Sci. Technol.* 7:9. doi: 10.1167/tvst.7.6.9

Ziemssen, F., Lagrèze, W., and Voykov, B. (2017). [Secondary diseases in high myopia]. *Der. Ophthalmol.* 114, 30–43.

Zloto, O., Wygnanski-Jaffe, T., Farzavandi, S., Gomez-De-Liaño, R., Sprunger, D., and Mezer, E. (2018). Current trends among pediatric ophthalmologists to decrease myopia progression-an international perspective. *Graefes Arch. Clin. Exp. Ophthalmol.* 256, 2457–2466. doi: 10.1007/s00417-018-4078-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 23 February 2021
doi: 10.3389/fnins.2021.591302

Validation of Computer-Adaptive Contrast Sensitivity as a Tool to Assess Visual Impairment in Multiple Sclerosis Patients

*Sina C. Rosenkranz[1,2], Barbara Kaulen[1,2], Hanna G. Zimmermann[3,4], Ava K. Bittner[5,6], Michael Dorr[7] and Jan-Patrick Stellmann[1,2,8,9]**

[1] Institut für Neuroimmunologie und Multiple Sklerose, Zentrum für Molekulare Neurobiologie, Hamburg, Germany, [2] Klinik und Poliklinik für Neurologie, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany, [3] Experimental and Clinical Research Center, Max Delbrück Center for Molecular Medicine and Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany, [4] NeuroCure Clinical Research Center, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany, [5] College of Optometry, Nova Southeastern University, Fort Lauderdale, FL, United States, [6] Department of Ophthalmology, Stein Eye Institute, University of California, Los Angeles, Los Angeles, CA, United States, [7] Adaptive Sensory Technology, Lübeck, Germany, [8] APHM, Hôpital de la Timone, CEMEREM, Marseille, France, [9] Aix Marseille Université, CRMBM, CNRS UMR 7339, Marseille, France

OPEN ACCESS

Edited by:
Alexandre Reynaud, McGill University, Canada

Reviewed by:
Fang Hou, Wenzhou Medical University, China
Sara Collorone, University College London, United Kingdom

***Correspondence:**
Jan-Patrick Stellmann
jan-patrick.stellmann@univ-amu.fr

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *04 August 2020*
Accepted: *02 February 2021*
Published: *23 February 2021*

Citation:
Rosenkranz SC, Kaulen B, Zimmermann HG, Bittner AK, Dorr M and Stellmann J-P (2021) Validation of Computer-Adaptive Contrast Sensitivity as a Tool to Assess Visual Impairment in Multiple Sclerosis Patients. Front. Neurosci. 15:591302. doi: 10.3389/fnins.2021.591302

Background: Impairment of visual function is one of the major symptoms of people with multiple sclerosis (pwMS). A multitude of disease effects including inflammation and neurodegeneration lead to structural impairment in the visual system. However, the gold standard of disability quantification, the expanded disability status scale (EDSS), relies on visual assessment charts. A more comprehensive assessment of visual function is the full contrast sensitivity function (CSF), but most tools are time consuming and not feasible in clinical routine. The quantitative CSF (qCSF) test is a computerized test to assess the full CSF. We have already shown a better correlation with visual quality of life (QoL) than for classical high and low contrast charts in multiple sclerosis (MS).

Objective: To study the precision, test duration, and repeatability of the qCSF in pwMS. In order to evaluate the discrimination ability, we compared the data of pwMS to healthy controls.

Methods: We recruited two independent cohorts of MS patients. Within the precision cohort (n = 54), we analyzed the benefit of running 50 instead of 25 qCSF trials. The repeatability cohort (n = 44) was assessed by high contrast vision charts and qCSF assessments twice and we computed repeatability metrics. For the discrimination ability we used the data from all pwMS without any previous optic neuritis and compared the area under the log CSF (AULCSF) to an age-matched healthy control data set.

Results: We identified 25 trials of the qCSF algorithm as a sufficient amount for a precise estimate of the CSF. The median test duration for one eye was 185 s (range 129–373 s). The AULCSF had better test–retest repeatability (Mean Average Precision, MAP) than visual acuity measured by standard high contrast visual acuity charts or

CSF acuity measured with the qCSF (0.18 vs. 0.11 and 0.17, respectively). Even better repeatability (MAP = 0.19) was demonstrated by a CSF-derived feature that was inspired by low-contrast acuity charts, i.e., the highest spatial frequency at 25% contrast. When compared to healthy controls, the MS patients showed reduced CSF (average AULCSF 1.21 vs. 1.42, $p < 0.01$).

Conclusion: High precision, usability, repeatability, and discrimination support the qCSF as a tool to assess contrast vision in pwMS.

Keywords: qCSF, AULCSF, precision, repeatability, discrimination, multiple sclerosis, vision

INTRODUCTION

Visual impairment can be one of the major symptoms in multiple sclerosis (MS) patients. Because vision is rated as one of the three most important bodily functions by MS patients, its impairment has a high impact on quality of life (QoL; Balcer et al., 2015; Heesen et al., 2018). Structural pathological changes in the retina are caused not only by acute optic neuritis, but also as the consequence of chronic inflammation, demyelination, and progressive neurodegeneration (Talman et al., 2010). Key pathological features observed in the retinas of MS patients are decreased thickness of the retinal nerve fiber layer (RNFL), ganglion cell layer, and inner plexiform layer (Balcer et al., 2015; Martinez-Lapiscina et al., 2016; Petzold et al., 2017). However, vision outcomes are not yet routinely implemented as a disease monitoring tool in clinical care or clinical trials for MS. This is mainly due to the fact that the available tests either do not sufficiently represent the pathological changes in the central nervous system or are too time-consuming to incorporate into routine clinical practice.

Currently, visual impairment in clinical practice is usually assessed by the Snellen chart, which records high-contrast visual acuity (HCVA) (Gilbert and Hopkinson, 1949). For example, the gold standard for MS disability quantification, the expanded disability status scale (EDSS), relies on HCVA to estimate its visual functioning score. In MS patients, however, low-contrast visual acuity (LCVA) seems to better correlate with the alterations of retinal morphology (Schinzel et al., 2014) and cognitive function (Wieder et al., 2013). LCVA is usually assessed by the low-contrast Sloan letter charts, but the evidence supporting this method is controversial as association with vision-related QoL in MS patients is inconsistent (Mowry et al., 2009; Stellmann et al., 2015b; Sabadia et al., 2016). Sloan LCVA charts usually measure at selected contrast levels [for example 1.25% or 2.5% (Balcer et al., 2012, 2017)] whereas the affected contrast sensitivity changes on an individual basis for different letter sizes. Consequently, contrast sensitivity should be assessed across a range of different letter sizes or spatial frequencies (Balcer et al., 2017), but the most established tools to measure the full contrast sensitivity function (CSF) are time-consuming, unreliable, and not feasible in routine clinical practice (Kalia et al., 2014). The quantitative CSF (qCSF, previously also termed quick CSF) test is a computerized test that uses a Bayesian adaptive method to assess the full CSF, which implicitly includes LCVA and HCVA, quickly but precisely (Lesmes et al., 2010; Dorr et al., 2013).

We recently showed that qCSF could be a useful tool for the assessment of visual function in MS patients as it correlated best with vision-related QoL measured by the National Eye Institute Visual Functioning Questionnaire (NEI-VFQ) (Stellmann et al., 2015b), whereas VA with Sloan charts was not significantly associated with the NEI-VFQ administered to MS patients. However, to establish the qCSF as a diagnostic tool in routine clinical care and research, further validation is necessary in the intended patient population. Here, we studied two independent consecutive MS cohorts in order to (i) optimize the precision and usability and (ii) evaluate repeatability and discrimination of the qCSF in MS patients.

MATERIALS AND METHODS

Patient Cohorts

Between August 2014 until November 2016 we included consecutive patients with a clinically isolated syndrome highly suggestive for MS or patients with a definite MS diagnosis according to the revised McDonald criteria (Polman et al., 2011) for two independent cohorts: (i) $n = 54$ in the *precision* cohort; and (ii) $n = 44$ in the repeatability cohort. Within the precision cohort, we aimed to compare the precision of two different test settings for the qCSF device, namely 25 vs. 50 qCSF trials. The repeatability cohort performed two consecutive qCSF assessments at a single visit to determine within-session variability with the previously established number of trials, and an additional HCVA assessment, which was chosen as the most commonly used standard clinical visual outcome measure. We assessed the typical median test duration for the qCSF method in the repeatability cohort. The cohorts were recruited sequentially without any overlap in the recruitment periods; thus, no patient was included in both cohorts. For an evaluation of the ability of the qCSF to discriminate between normal vision and abnormal vision in MS, we pooled the patients with a RRMS disease course and without any current or previous optic neuritis of both cohorts ($n = 13$ patients, 40 measurements) and compared them to a published normative dataset of 61 age-matched healthy controls (186 measurements because some subjects were tested repeatedly at two visits about a week apart) (Lesmes et al., 2017). The diagnosis of an acute optic neuritis was assessed by an Ophthalmologist and Neurologist. Historical optic neuritis was assessed with clinical history. Previous optic neuritis was

excluded to focus on the neurodegenerative changes rather than previous inflammatory damage. All participants with MS were recruited at the MS outpatient clinic, University Medical Centre Hamburg-Eppendorf, and the healthy controls were recruited and tested at the Nova Southeastern University, College of Optometry. All subjects gave their written informed consent prior to any testing. The local ethics committees approved the studies (Ethics Committee of the Board of Physicians in the State of Hamburg, PV4455 for both MS cohorts and the Institutional Review Board at the Nova Southeastern University in Fort Lauderdale, Florida for the healthy controls).

Data Acquisition

All tests were performed in the same room, under the same daylight illumination, and in the same order for both MS cohorts; the same stipulations applied for the normal healthy controls tested at a different site. Each eye was assessed separately and measured with best habitual correction (e.g., glasses or contact lenses). For the precision analysis, the repeatability and the analysis of the time duration we treated within-subject eyes as independent measurements, for the discrimination analysis we averaged the available measurements per subject. To assess the full CSF, for each of the 25 trials, the qCSF device (Manifold Contrast Vision Meter, Adaptive Sensory Technology) presented three bandpass-filtered Sloan letters of varying size and contrast on a 46-inch computer screen at a viewing distance of 4.5 m (for details see Lesmes et al., 2010; Dorr et al., 2013). The letters were presented until the patient gave a response; hence the test duration depended on the individual being tested. Based on the participant's responses, the method chose the most informative combination of size and contrast for the next trial. Test duration for the qCSF was calculated by the file timestamps of subsequent tested eyes. This estimation includes changing the eye patch, any break (rest period) requested by the subject, and re-entering test details. For the precision cohort we ran the qCSF with 50 instead of 25 trials. For the repeatability cohort, the HCVA at 5 m (VA 500) was additionally determined with standard Snellen charts presenting nine lines with 1 to 10 letters. The smallest line with less than two mistakes was defined VA500 (possible values in LogMAR: −0.1, 0, 0.2, 0.3, 0.5, 0.6, 0.7, 0.85, and 1). For all MS subjects, the clinical neurological status was assessed by trained neurologists with the EDSS including its functional system subscore for vision (Kurtzke, 1983). [For correlations of area under the log CSF (AULCSF) and CSF acuity to VA500 see **Supplementary Figures 1A,B**, for correlations of AULCSF with EDSS and disease duration see **Supplementary Figures 2A,B**].

Contrast Sensitivity Function Outcomes

The shape of the CSF can be modeled with four parameters: (i) peak spatial frequency, (ii) peak contrast sensitivity, (iii) bandwidth, and (iv) a low-frequency truncation parameter (Lesmes et al., 2010). The qCSF provides not only scalar estimates of these parameters, but their full joint posterior distribution. However, to simplify statistics and facilitate comparison with the point estimates of existing charts, here we calculated and report several scalar features of the CSF: (i) the AULCSF in the range from 1.5 to 18 cycles per degree; (ii) the CSF acuity, the spatial frequency for which sensitivity reaches zero (i.e., 100% contrast is required for recognition); and (iii) a "low-contrast CSF acuity" as an approximate equivalent of low-contrast acuity charts. Because of the different stimuli (bandpass-filtered vs. unfiltered letters) and the different contrast definitions (Michelson vs. Weber contrast), contrast values (typically 1.25% or 2.5% for LCVA charts) are not directly comparable, so we varied the threshold to 2.5% and in 5% steps from 5 to 50%. In our repeatability cohort, best repeatability precision was obtained with the spatial frequency for which sensitivity reached 0.6 (=log10(4), i.e., 25% contrast required for recognition), and we will thus refer to this parameter as "CS4."

Data Analysis

Because the ground truth of the actual CSF of an observer is unknown, we estimated the accuracy by calculating the convergence in the precision cohort by running the qCSF for 50 trials. We also calculated the difference of the CSF estimates after 25 and 50 trials (mean bias and mean absolute bias). For calculating the within-visit test-retest repeatability we used the Bland-Altman coefficient of repeatability (COR) (Bland and Altman, 1986). However, standard vision charts have a limited number of values and low discriminant abilities for the continuum of visual functioning. From a conceptual point of view, this leads to a risk of artificially high repeatability at the cost of low sensitivity to detect subtle differences, and restricts the use of COR for comparison of quantized tests (e.g., vision charts) and continuous outcomes, such as AULCSF. In recent work, we therefore have developed a new metric that penalizes test quantization, namely, Mean Average Precision (MAP) (Dorr et al., 2018). This metric assesses how uniquely an individual is identified by their test and repeated test pair: Repeated tests from the same subject should yield the same result, whereas different subjects should typically yield different results. MAP ranks all retests by their similarity to the first subject's test and an average rank precision over all participants can be computed. A MAP score of 1.0 indicates high precision and resolution, whereas the MAP score approaches zero for a poor test. For the discrimination between healthy controls and MS patients, we calculate a between-group *t*-test on the AULCSF summary statistic.

RESULTS

Quantitative CSF Precision and Number of Trials

We have already shown that 25 trials of the qCSF algorithm are sufficient for a good estimate of the CSF in MS patients (Stellmann et al., 2015b). However, greater precision may be obtained by running the adaptive Bayesian procedure for more trials. In the precision cohort, we analyzed the difference of the estimated CSF between 25 and 50 trials by calculating the

TABLE 1 | Baseline characteristics of the precision cohort.

Baseline characteristics (*n*)	Patients (54)
Age, mean (SD), range, years	41 (10), 20–63
Gender, male/female	15/39
EDSS, mean (SD), range	2.39 (1.41), 0–6.5
Time since first symptoms, years (SD), range	8 (8), 0–29
CIS, *n* (%)	10 (18.5)
RRMS, *n* (%)	38 (70.4)
PPMS, *n* (%)	3 (5.5)
SPMS, *n* (%)	3 (5.5)

SD = Standard deviation; EDSS = Expanded Disability Status Scale; CIS = Clinically isolated syndrome; RRMS = Relapsing remitting MS; SPMS = Secondary progressive MS; PPMS = Primary progressive MS.

convergence. Descriptive statistics for the precision cohort are summarized in **Table 1**.

Two of the 54 included subjects with MS were measured monocularly only because they had no remaining vision in one of their eyes. One measurement in one subject had to be excluded due to technical problems. This resulted in a total of 105 qCSF measurements. **Figure 1A** shows the convergence of the AULCSF estimate over the time course of 50 trials (internal variability, not variability across measurements). For many patients, the first incorrect response leads to a drop in the AULCSF estimate, which then recovers after the first 5–10 trials. After 25–30 trials, the mean AULCSF estimate and the credible interval width have mostly converged with a small but statistically significant difference between the AULCSF estimate after 25 and after 50 trials (**Figure 1B**); (for 105 measurements mean = −0.02 log10 units, SD = 0.049; 95% CI (−0.03, −0.01), *one-sample t-test t* (104) = −4.34, $p < 0.001$). In line with previous work and the negligible benefit above 25 trials compared to the cost of additional test time, we performed subsequent testing and analyses with 25 trials.

Quantitative CSF Repeatability

Next, we assessed the individual within-visit reliability of two repeated tests in the repeatability cohort of MS patients. We included 44 subjects with one who was measured in one eye only. This resulted in 87 pairs of test and retest measurements; one eye was excluded from the CS4 calculation because its contrast sensitivity was too low (no stimulus size could be recognized at 25% contrast). **Table 2** shows the baseline descriptive statistics of this cohort. Numerically, AULCSF had the highest ("least repeatable") COR in comparison to VA500 based on standard charts, CSF acuity, and CS4 (0.23 vs. 0.08, 0.14, 0.13, respectively, **Figure 2A**; for Bland-Altman plots, see **Supplementary Figures 3A–D**). However, the test score ranges are not directly comparable and paper charts strongly quantize results (see **Supplementary Figure 3D**), i.e., lack resolution. Using the novel metric MAP that penalizes coarse quantization, the AULCSF, high-contrast CSF acuity and low-contrast CS4 had better repeatability and precision (higher MAP values) than VA500 (0.18, 0.17, 0.19. vs. 0.11, respectively, **Figure 2B**).

Quantitative qCSF Test Duration

Test duration is an important factor for clinical usability. We looked at the duration between subsequent tests of 25 trials each for alternating eyes in the repeatability cohort (43 patients, the patient with only one measured eye was excluded). As the data were not normally distributed, we fit a mixture of two Gaussian distributions (mean 172 and 219 s; standard deviation 17 and 10 s, respectively). Median test duration per eye including all preparation, such as changing the eye patch and a break between eyes, was 185 s (5th and 95th percentiles were 154 and 260 s, respectively, **Figure 3**).

Discriminative Power of the qCSF in MS

We further investigated whether the qCSF could serve as a diagnostic instrument in MS and could discriminate these patients from healthy controls. For this aim, we pooled all qCSF

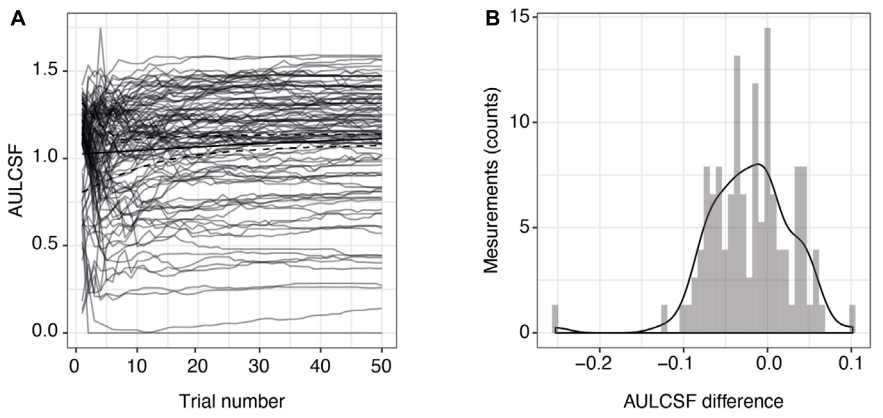


FIGURE 1 | Precision of AULCSF estimate over time. **(A)** AULCSF estimates for each eye (gray lines) (*n* = 105 measurements) and their mean over all measurements (solid black line). The dashed black lines show the mean lower and upper ends of the 68.3% credible interval estimates averaged across all measurements and are therefore indicative of the internal measurement variability (not of variability across measurements). **(B)** Distribution of differences between AULCSF estimates after 25 and 50 trials (*n* = 105 measurements). The *y*-axis shows the absolute number of measurements within each bin of differences. AULCSF = area under the log CSF.

TABLE 2 | Baseline characteristics of the repeatability cohort.

Baseline characteristics (*n*)	All patients (44)
Age, mean (SD), range, years	41 (9.7), 22–58
Gender, male/female	13/31
EDSS, mean (SD), range	2.49 (1.73), 0–6.5
Time since first symptoms, years (SD), range	9.2 (9.1), 0–36
Visual acuity,% (SD), range	83.2 (25.3), 10–125
CIS, *n* (%)	5
RRMS, *n* (%)	25
PPMS, *n* (%)	1
SPMS, *n* (%)	2

SD = Standard deviation; EDSS = Expanded Disability Status Scale; CIS = Clinically isolated syndrome; RRMS = Relapsing remitting MS; SPMS = Secondary progressive MS; PPMS = Primary progressive MS.

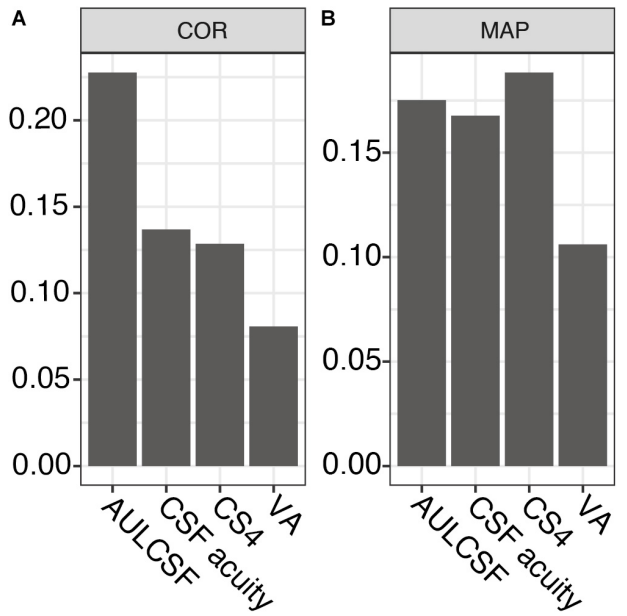


FIGURE 2 | Repeatability metrics for AULCSF, CSF acuity, CS4, and VA500. **(A)** Coefficient of repeatability (only for the same feature, lower values correspond to greater precision). **(B)** Mean Average Precision (higher values correspond to greater precision, independent of test feature). AULCSF = area under the log CSF; CSF acuity = visual acuity of contrast sensitivity function; CS4 = CSF acuity at 25% contrast; VA = visual acuity; COR = coefficient of repeatability; MAP = Mean Average Precision.

measurements from both MS cohorts (precision and repeatability cohort), excluded all patients with a previous optic neuritis or an unknown visual symptom history and compared them to a normative data set of age-matched healthy normal controls (n = 186 measurements from 61 subjects, age range 20–59 years). On average, MS subjects (mean EDSS 2.65; SD 1.62) had reduced CSF results when compared to healthy controls (mean AULCSF after averaging measurements per subject 1.21 vs. 1.42, 95% CI of group differences (0.12, 0.3), Welch's t (36.28) = 4.77, $p < 0.001$). The most pronounced difference is seen on the top left side of the curve indicating a larger gap of performance for large and medium-sized letters at low contrast (**Figure 4**).

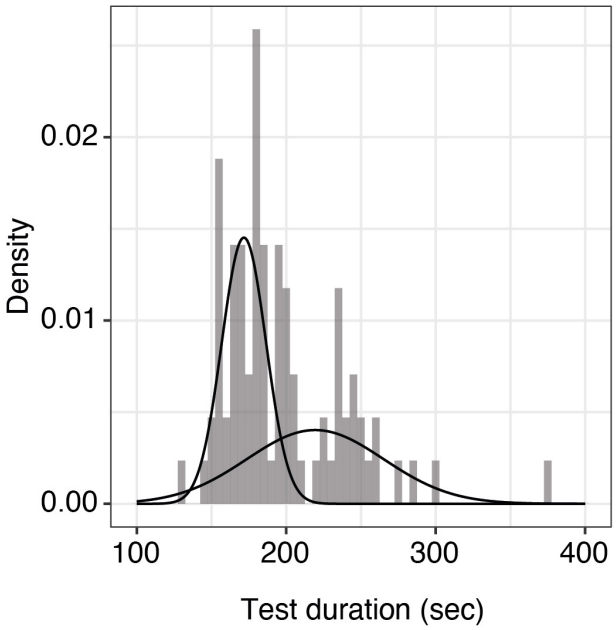


FIGURE 3 | Histogram of test durations per eye including all preparations. Shown here is the distribution of qCSF test duration per eye (n = 86 measurements) including all preparations. The raw data were best fit by a mixture of two Gaussian distributions with means of 172 and 219 s, respectively. The overall median test duration per eye is 185 s.

DISCUSSION

Precise, reliable assessment of visual function is an important part of disability quantification in MS (Balcer et al., 2015; Heesen et al., 2018). However, currently available tools have several limitations and their impact on clinical decisions is probably negligible. Here, we provide important feasibility and reliability information about a new outcome to assess visual impairment in MS. The qCSF method shows higher precision than the current standard of care (VA500) without major burden to the patient or clinician. The qCSF has previously been compared to standard vision outcomes in MS patients, namely HCVA and LCVA, and already has demonstrated higher correlation with visual quality of life from the MS patient's perspective than HCVA or LCVA (Stellmann et al., 2015b).

In a first consecutive cohort, we aimed to define a reliable trade-off between duration and precision of the qCSF computer-adaptive algorithm. Conceptually, precision of an estimate increases as a function of time as the impact of outliers and lapses decreases with the number of trials. However, long test duration directly reduces the feasibility and acceptance in clinical care and trials. For example, longer walking tests of 2 or 6 min for several years already have been considered to replace the MS standard of a timed 25 foot walk (which usually takes less than 1 min), but implementation of such outcomes is still rare (Gijbels et al., 2012; Stellmann et al., 2015a). We found that 25 trials were sufficient for a reliable convergence of the algorithm, and a longer test duration did not change the estimate of the

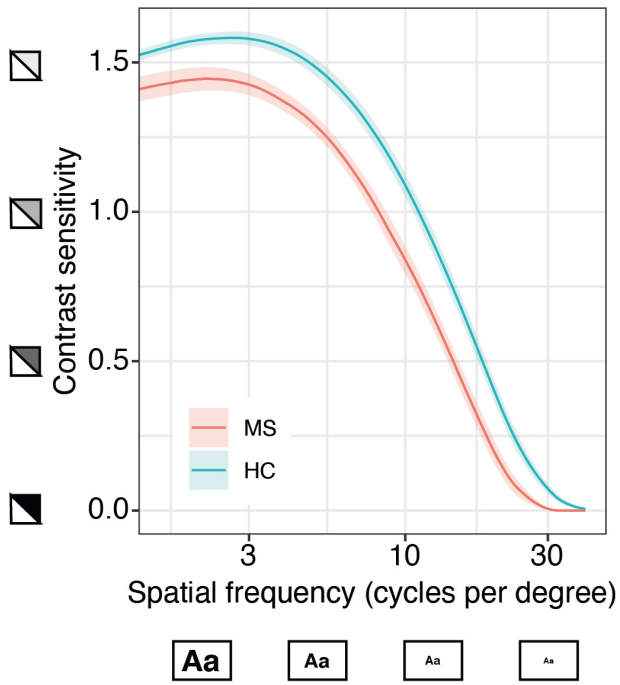


FIGURE 4 | Quantitative contrast sensitivity function (qCSF)-examination in MS patients compared to age-matched healthy controls (HC). Curves show mean contrast sensitivity function (CSF) in MS patients (n = 13, based on patient-specific averages of a total of 40 measurements) and HC (n = 61, 186 measurements) ± SEM, x-axis represents logarithmic spatial frequencies, i.e., decreasing size of the letters; y-axis represents logarithmic sensitivity to decreasing contrast.

AULCSF substantially. This is in line with a previous study that showed that 25 trials of the qCSF were enough to demonstrate loss of visual function in diabetic subjects with and without retinopathy compared to healthy controls (Joltikov et al., 2017). However, longer test durations might still be necessary if other, potentially noisier features than the summary statistic AULCSF were of interest.

The second aim of our study was to evaluate the repeatability performance of the qCSF method in MS patients. First, we found that the visual function estimate did not differ substantially between two separate assessments in each individual, i.e., we found high stability of the method or high intra-individual reliability. This held true for both the evaluated parameters of the qCSF and for visual acuity VA500. However, care must be taken not to confound the commonly reported COR as a tool to compare tests with different outputs. The smaller range and the strong quantization of VA500 scores led to a seemingly excellent COR. However, repeatability (or precision) is only a necessary but not sufficient condition for the more important quality of a biomarker, which is the sensitivity to detect changes in the underlying signal. In the absence of a ground truth of visual change, we hypothesize that different individuals differ in their visual function, and the MAP metric quantifies the discriminant ability of a test for each individual in comparison to the whole group. Here, in a reversal of the ranking based on the COR, the summary statistic AULCSF provided higher discriminant ability than VA500 and CSF acuity.

Similar to low-contrast Sloan VA charts, we further calculated the cut-off frequency for stimuli presented at 25% contrast (CS4), and found that this parameter was even more precise than the AULCSF. In principle, the combination of several parameters, or the distributions thereof, might provide even further precision gains. This also demonstrated the advantage of estimating the full CSF versus paper charts that are limited to a fixed contrast level (HCVA or LCVA) or fixed spatial frequency (e.g., Pelli-Robson CS chart) (Pelli and Bex, 2013). In addition to the summary statistic AULCSF, we were able to calculate cut-off frequencies for a large number of contrast levels, in order to select the parameter with the highest test-retest precision. In general, different ocular or neurological pathologies might affect vision differently at different stages, and the most informative part of the CSF may therefore differ across individuals or diseases. For example, low vision patients may experience floor effects at low contrast levels, which was the case for one MS patient at our 25% contrast criterion. Testing the entire CSF thus solves the conundrum of knowing where to test in advance.

Furthermore, we also assessed typical test duration. With an average of about 3 min, results are in agreement with previous qCSF data from healthy controls (Dorr et al., 2017). Taken together with the comparable COR (Lesmes et al., 2017), this implies that mildly to moderately impaired MS patients do not need specialized test protocols because they are as fast and precise, albeit at a lower performance level, than healthy controls. Notably, the test duration for the qCSF in our study seems comparable to Sloan charts in previous studies. The literature reports a test duration of 10–15 min for a complete monocular and binocular testing at two contrast levels for the Sloan charts (Balcer et al., 2017).

The descriptive analysis comparing average CSF estimates from MS and healthy controls indicates a worse performance over the complete frequency range with marked impairment of sensitivity to lower contrast levels. This finding in our study is consistent with reports for Sloan charts, that provide best discriminant abilities for MS with 1.25 and 2.5% low contrast VA charts. While 0.6% seems to be biased by a floor effect, 5% charts are already limited by a raising ceiling effect (Balcer et al., 2017). The qCSF approach avoids test restrictions at selected contrast levels, which may differ with disease progression, or spatial frequencies as the full CSF is estimated. The current main outcome, AULCSF, might thus be sensitive to differentiate patients from healthy controls as subtle differences over the complete CSF might sum up to a significant difference of the AULCSF. However, a sufficiently powered future head-to-head comparison of matched individuals tested under the same conditions might reveal even more discriminatory features of the CSF. Moreover, to establish qCSF as a diagnostic tool for MS, prospective studies are needed to determine sensitivity for change over time and define meaningful clinically relevant cut-offs.

The main limitation of the current study was the lack of comparison with Sloan VA charts in MS patients. However, we aimed to keep the burden for recruited patients low and decided to contrast our findings only with the clinical and EDSS

standard, which is HCVA with a Snellen chart. Interestingly, repeatability data for Sloan charts in MS patients have not been previously published, although a good intra-rater agreement has been reported (ICC 0.86–0.95) (Balcer et al., 2000). It should also be noted that the data from healthy controls were collected in a different clinic under slightly different conditions, at a viewing distance of 400 vs. 450 cm. However, these subtle factors likely did not substantially affect the rather large (qualitative) difference in qCSF on average that we observed between MS and healthy controls.

Taken together, qCSF provides excellent test characteristics and has already been linked to visual quality of life in MS. Our findings indicate that the further evaluation of the qCSF method in longitudinal studies as an outcome measure for MS seems promising. These studies should also include comparisons to the Sloan VA charts.

DATA AVAILABILITY STATEMENT

The raw data of all MS patients supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Board of Physicians in the State of Hamburg and the Institutional Review Board at the Nova Southeastern University in Fort Lauderdale, Florida. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

J-PS conducted and designed the experiments. AKB designed the study involving healthy controls. SCR, MD, and J-PS analyzed the data and wrote the manuscript. BK analyzed the data. HZ and AKB supervised the study and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SCR was supported by a Clinician-Scientist Fellowship from the Stifterverband für die Deutsche Wissenschaft and the Hertie Network of Excellence in Clinical Neuroscience of the Gemeinnützige Hertie-Stiftung. The study of the healthy controls was funded by an investigator-initiated award from Adaptive Sensory Technology to AKB at Nova Southeastern University. AKB was also supported by an unrestricted grant from Research to Prevent Blindness, Inc., to the Department of Ophthalmology at the University of California, Los Angeles, Los Angeles, CA, United States.

ACKNOWLEDGMENTS

We thank all study nurses from the MS outpatient clinic for their excellent assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.591302/full#supplementary-material

REFERENCES

Balcer, L. J., Baier, M. L., Kunkle, A. M., Rudick, R. A., Weinstock-Guttman, B., Simonian, N., et al. (2000). Self-reported visual dysfunction in multiple sclerosis: results from the 25-item National Eye Institute Visual Function Questionnaire (Vfq-25). *Mult. Scler.* 6, 382–385. doi: 10.1177/135245850000600604

Balcer, L. J., Galetta, S. L., Polman, C. H., Eggenberger, E., Calabresi, P. A., Zhang, A., et al. (2012). Low-contrast acuity measures visual improvement in phase 3 trial of natalizumab in relapsing MS. *J. Neurol. Sci.* 318, 119–124. doi: 10.1016/J.JNS.2012.03.009

Balcer, L. J., Miller, D. H., Reingold, S. C., and Cohen, J. a (2015). Vision and vision-related outcome measures in multiple sclerosis. *Brain* 138(Pt 1), 11–27. doi: 10.1093/brain/awu335

Balcer, L. J., Raynowska, J., Nolan, R., Galetta, S. L., Kapoor, R., Benedict, R., et al. (2017). Validity of low-contrast letter acuity as a visual performance outcome measure for multiple sclerosis. *Mult. Scler.* 23, 734–747. doi: 10.1177/1352458517690822

Dorr, M., Elze, T., Wang, H., Lu, Z. L., Bex, P. J., and Lesmes, L. A. (2018). New precision metrics for contrast sensitivity testing. *IEEE J. Biomed. Heal. Inform.* 22, 919–925. doi: 10.1109/JBHI.2017.2708745

Dorr, M., Lesmes, L. A., Elze, T., Wang, H., Lu, Z. L., and Bex, P. J. (2017). Evaluation of the precision of contrast sensitivity function assessment on a tablet device. *Sci. Rep.* 7:46706. doi: 10.1038/srep46706

Dorr, M., Lesmes, L. A., Lu, Z., and Bex, P. J. (2013). Rapid and reliable assessment of the contrast sensitivity function on an iPad. *Invest. Ophthalmol. Vis. Sci.* 54, 7266–7273. doi: 10.1167/iovs.13-11743

Gijbels, D., Dalgas, U., Romberg, A., de Groot, V., Bethoux, F., Vaney, C., et al. (2012). Which walking capacity tests to use in multiple sclerosis? A multicentre study providing the basis for a core set. *Mult. Scler.* 18, 364–371. doi: 10.1177/1352458511420598

Gilbert, M., and Hopkinson, R. G. (1949). The illumination of the Snellen chart. *Br. J. Ophthalmol.* 33, 305–310. doi: 10.1136/bjo.33.5.305

Heesen, C., Haase, R., Melzig, S., Poettgen, J., Berghoff, M., Paul, F., et al. (2018). Perceptions on the value of bodily functions in multiple sclerosis. *Acta Neurol. Scand.* 137, 356–362. doi: 10.1111/ane.12881

Joltikov, K. A., de Castro, V. M., Davila, J. R., Anand, R., Khan, S. M., Farbman, N., et al. (2017). Multidimensional functional and structural evaluation reveals neuroretinal impairment in early diabetic retinopathy. *Invest. Ophthalmol. Vis. Sci.* 58, BIO277–BIO290. doi: 10.1167/iovs.17-21863

Kalia, A., Lesmes, L. A., Dorr, M., Gandhi, T., Chatterjee, G., Ganesh, S., et al. (2014). Development of pattern vision following early and extended blindness. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2035–2039. doi: 10.1073/pnas.1311041111

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33, 1444–1452.

Lesmes, L. A., Bittner, A. K., Lu, Z.-L., Bex, P. J., and Dorr, M. (2017). Distinguishing the contribution of precision and repeatability to vision testing. *Invest. Ophthalmol. Vis. Sci.* 58:2204.

Lesmes, L. A., Lu, Z.-L., Baek, J., and Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J. Vis.* 10, 17.1–21. doi: 10.1167/10.3.17

Bland, J. M., and Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1, 307–310. doi: 10.1016/S0140-6736(86)90837-8

Martinez-Lapiscina, E. H., Arnow, S., Wilson, J. A., Saidha, S., Preiningerova, J. L., Oberwahrenbrock, T., et al. (2016). Retinal thickness measured with optical coherence tomography and risk of disability worsening in multiple sclerosis: a cohort study. *Lancet Neurol.* 15, 574–584. doi: 10.1016/S1474-4422(16)00068-5

Mowry, E. M., Loguidice, M. J., Daniels, a B, Jacobs, D. a, Markowitz, C. E., Galetta, S. L., et al. (2009). Vision related quality of life in multiple sclerosis: correlation with new measures of low and high contrast letter acuity. *J. Neurol. Neurosurg. Psychiatry* 80, 767–772. doi: 10.1136/jnnp.2008.165449

Pelli, D. G., and Bex, P. (2013). Measuring contrast sensitivity. *Vis. Res.* 90, 10–14. doi: 10.1016/j.visres.2013.04.015

Petzold, A., Balcer, L. J., Calabresi, P. A., Costello, F., Frohman, T. C., Frohman, E. M., et al. (2017). Retinal layer segmentation in multiple sclerosis: a systematic review and meta-analysis. *Lancet Neurol.* 16, 797–812. doi: 10.1016/S1474-4422(17)30278-8

Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., et al. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann. Neurol.* 69, 292–302. doi: 10.1002/ana.22366

Sabadia, S. B., Nolan, R. C., Galetta, K. M., Narayana, K. M., Wilson, J. A., Calabresi, P. A., et al. (2016). 20/40 or better visual acuity after optic neuritis: not as good as we once thought? *J. Neuro Ophthalmol.* 36, 369–376. doi: 10.1097/WNO.0000000000000421

Schinzel, J., Zimmermann, H., Paul, F., Ruprecht, K., Hahn, K., Brandt, A. U., et al. (2014). Relations of low contrast visual acuity, quality of life and multiple sclerosis functional composite: a cross-sectional analysis. *BMC Neurol.* 14:31. doi: 10.1186/1471-2377-14-31

Stellmann, J. P., Neuhaus, A., Götze, N., Briken, S., Lederer, C., Schimpl, M., et al. (2015a). Ecological validity of walking capacity tests in multiple sclerosis. *PLoS One* 10:e0123822. doi: 10.1371/journal.pone.0123822

Stellmann, J. P., Young, K., Pottgen, J., Dorr, M., and Heesen, C. (2015b). Introducing a new method to assess vision: computer-adaptive contrast-sensitivity testing predicts visual functioning better than charts in multiple sclerosis patients. *Mult. Scler. J. Exp. Transl. Clin.* 1:2055217315596184. doi: 10.1177/2055217315596184

Talman, L. S., Bisker, E. R., Sackel, D. J., Long, D. A. Jr., Galetta, K. M., Ratchford, J. N., et al. (2010). Longitudinal study of vision and retinal nerve fiber layer thickness in MS. *Ann Neurol.* 67, 749–760. doi: 10.1002/ana.22005

Wieder, L., Gäde, G., Pech, L. M., Zimmermann, H., Wernecke, K.-D., Dörr, J.-M., et al. (2013). Low contrast visual acuity testing is associated with cognitive performance in multiple sclerosis: a cross-sectional pilot study. *BMC Neurol.* 13:167. doi: 10.1186/1471-2377-13-167

Conflict of Interest: MD holds equity in and employment by Adaptive Sensory Technology and holds patents related to the qCSF. The qCSF device was provided to INIMS free of charge; no further compensation was granted. HZ received research grants from Novartis and speaking honoraria from Bayer Healthcare, unrelated to this study.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 09 March 2021
doi: 10.3389/fnins.2021.626466

The Curve Visible on the Campbell-Robson Chart Is Not the Contrast Sensitivity Function

Jessica Tardif[1], Marcus R. Watson[2], Deborah Giaschi[3] and Frédéric Gosselin[1]*

[1] Département de Psychologie, Université de Montréal, Montréal, QC, Canada, [2] Department of Biology, York University, Toronto, ON, Canada, [3] Department of Ophthalmology and Visual Sciences, The University of British Columbia, Vancouver, BC, Canada

The Campbell-Robson chart is a highly popular figure used in psychophysics and visual perception textbooks to illustrate the Contrast Sensitivity Function (CSF). The chart depicts a grating which varies logarithmically in spatial frequency (SF) from left to right and in contrast from bottom to top. Campbell and Robson's (1964) intuition was that the boundary between the grating and the homogeneous gray area (below threshold) would trace the shape of the observer's own CSF. In this paper, we tested this intuition. A total of 170 participants (96 adults and 74 children) adjusted the four parameters of a truncated log-parabola directly onto a Campbell-Robson chart rendition and completed a gold-standard CSF evaluation. We hoped that this procedure which requires a mere three clicks on the computer mouse, would speed up the measurement of the CSF to under a minute. Unfortunately, the only parameter of the truncated log-parabola fitted to the gold-standard CSF data that could be predicted from the Campbell-Robson chart data was the peak sensitivity for the adult participants. We conclude that the curve visible on the Campbell-Robson chart cannot be used practically to measure the CSF.

Keywords: contrast sensitivity function, Campbell-Robson chart, spatial vision, low-level vision, psychophysics

OPEN ACCESS

Edited by:
Zhong-Lin Lu,
New York University, United States

Reviewed by:
Luca Battaglini,
University of Padua, Italy
David R. Badcock,
University of Western Australia, Australia

****Correspondence:***
Jessica Tardif
jessica.tardif.1@umontreal.ca

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *05 November 2020*
Accepted: *10 February 2021*
Published: *09 March 2021*

Citation:
Tardif J, Watson MR, Giaschi D and Gosselin F (2021) The Curve Visible on the Campbell-Robson Chart Is Not the Contrast Sensitivity Function. Front. Neurosci. 15:626466. doi: 10.3389/fnins.2021.626466

INTRODUCTION

The human visual system analyses the complex luminance modulations that make up the visual stimulus with discrete channels, each tuned to a specific spatial frequency (SF) range that can be expressed in cycles per degree of visual angle (*cpd*; see De Valois and De Valois, 1990, for a review). Low SFs convey coarse information, such as the boundary between the sand and water on a beach, while high SFs can represent fine-grained information, such as the pole of a parasol several meters away (see Morrison and Schyns, 2001, and Ruiz-Soler and Beltran, 2006, for reviews). The threshold luminance contrast – the contrast required to detect simple sine wave gratings at a given level of performance or to discriminate their orientation – varies with SF. The contrast sensitivity function (CSF) depicts contrast sensitivity – the threshold's reciprocal – as a function of SF. It has the shape roughly of an upside-down U (Campbell and Robson, 1964).

A wide variety of researchers measure the CSF today: it is a useful tool for screening and assessing spatial vision in many visual and cognitive impairments (e.g., Regan, 1991). The gold

standard for measuring the CSF is to measure the contrast threshold for detecting the presence or for discriminating the orientation of 5–10 different SFs using the method of constant stimuli or a staircase method such as QUEST (for a review of different staircase methods, see Leek, 2001). When the method of constant stimuli is used to estimate the thresholds, each combination of SF and contrast is repeated 20 times or more, usually in random order. This procedure thus requires a minimum of 500 trials to evaluate five points on the CSF. When QUEST is used to measure the thresholds, each threshold estimate requires 40 trials or more. This method thus requires a minimum of 200 trials to evaluate five points on the CSF. In this article, we used QUEST for our gold-standard evaluation of the CSF. Several quicker and cruder alternatives have been proposed over the years (summarized below).

Why Is It Important to Measure the CSF Quickly?

Contrast sensitivity is often measured in research settings, as an experimental variable, a covariable or to control for subjects' visual health. There are various instances in which measuring the CSF with the gold-standard methods described above might be more difficult because of their duration. One such instance is when the population studied has difficulties to stay focused on a repetitive task for a long period. These populations include young children, older adults, and some persons who have neurological or psychological disorders lowering their attentional capacities, for example Attention Deficit Disorder or Major Depression. Another instance is simply when the other tasks participants complete already take a lot of time and completing a CSF assessment might add too much time. Similarly, if the CSF is used as a screening evaluation for excluding participants with atypical vision, a long evaluation might make this screening impractical.

In screening and regular vision evaluations, visual acuity is typically measured; it is the most useful tool to detect visual conditions, such as myopia. When contrast sensitivity is included in these regular test batteries, it usually consists of a measurement of perception threshold at a single SF (Pelli and Bex, 2013). Because the high SF end of the CSF drops to a null contrast sensitivity (i.e., no perception at 100% contrast) at a SF that approximates visual acuity, it can be estimated by a high-contrast Snellen-type chart. Similarly, the maximum contrast sensitivity of the CSF can be estimated, for example, by the Pelli-Robson chart described below (Pelli et al., 1988). Notwithstanding, these two types of screening tools each evaluate only one point on the CSF.

In some cases, visual acuity remains normal and contrast sensitivity is impaired (e.g., treated amblyopia: Huang et al., 2007; corrected severe myopia: Liou and Chiu, 2001; multiple sclerosis: Regan and Neima, 1983), it is therefore crucial to measure contrast sensitivity to detect these impairments. Further, in many cases, contrast sensitivity is only affected for a subset of SFs. In these cases, the isolated deficit cannot be observed if the method used does not encompass the SFs for which sensitivity is altered, which means that it is not enough to use the Snellen as well as the Pelli-Robson charts. In these cases, measuring the entire CSF is required to screen for visual impairments. For example, these cases include disorders that primarily affect sensitivity to high SF: amblyopia (in all studies 12 cpd and above; Hess and Howell, 1977; Bradley and Freeman, 1981; Sjöstrand, 1981; Howell et al., 1983; Levi et al., 1994; however, Barollo et al., 2017 observed low contrast sensitivity in amblyopia at all SFs), macular degeneration (in all studies 6 cpd and above; Wolkstein et al., 1980; Marmor, 1986; Kleiner et al., 1988; Stangos et al., 1995; Sunness et al., 1997) and high myopia (Thorn et al., 1986; Collins and Carney, 1990; Liou and Chiu, 2001). The SF ranges at which sensitivity is impaired seem to be case-dependent for other disorders. For example, multiple sclerosis usually affects middle SFs, but it also affects high and low SFs in some patients (Nordmann et al., 1987; Ashworth et al., 1989). Cataracts most often affect sensitivity only for intermediate and high SFs – about 2 cpd and above –, while in other cases they also affect lower SFs (Hess and Woo, 1978; Elliott et al., 1989; Elliott and Hurst, 1990; Pardhan and Gilchrist, 1991; Lewis et al., 1992; Superstein et al., 1997; Shandiz et al., 2011).

How Can the CSF Evaluations Be Shortened?

Over the years, CSF evaluation methods have been shortened in many clever ways. Before 1990, a few paper-based charts were introduced as quick and cheap methods for estimating the CSF. The Arden plates (1978), for example, were simple gratings of specific SFs printed on cards, with contrast varying from top to bottom. These cards were at first completely covered by a second opaque gray card which was moved slowly to reveal increasing contrast levels. Subjects verbally reported when they perceived the grating, and completed the procedure with different SFs, allowing the experimenter to trace a threshold curve. The results, although reported as unreliable (Robson, 1993), remain useful as a screening tool (Woods et al., 1998).

The Vistech chart (Ginsburg, 1984) is a cardboard chart featuring a grid of Gabor patches of different SFs and three orientations. Subjects simply report the orientation of the gratings until the contrast is too low for them to perceive it. An advantage of this method is that the task is objective, unlike in the case of the Arden plates. Subjects complete the Vistech chart in about 6 min (Robson, 1993). Threshold estimates tend to be noisy, however, because each SF and contrast combination is only presented once (Robson, 1993). While the Vistech thresholds correlate well with a method which includes a larger number of trials (Leat and Woo, 1997), its reliability is low (Rubin, 1988; Reeves et al., 1991). Furthermore, cardboard charts fade with time, altering the contrast of the printed gabors or letters.

More recently, computerized tests were combined with adaptive algorithms to measure the CSF. The Freiburg Visual Acuity Test (FrACT; Bach, 1996, 2006) is one such example. FrACT contains different tasks to evaluate acuity and contrast sensitivity. In one of these, the participant is asked to identify the orientation of a sinusoidal grating of a specified spatial frequency among four possibilities (horizontal, vertical, and the obliques). The contrast threshold of this stimulus is measured efficiently using the Best-PEST adaptive algorithm

(Lieberman and Pentland, 1982). The CSF of the participant can be estimated by running this task multiple times with gratings of different spatial frequencies. A software implementing FrACT is freely available[1].

In Quick CSF (Lesmes et al., 2010), another example of an adaptive method and the fastest to date, each stimulus is chosen in light of the participant's past accuracy to present the stimulus that has the greatest potential for new information. This method is highly efficient, taking only 10 min for an accurate measurement and 2 min for a broad evaluation (Lesmes et al., 2010). It has been validated for achromatic contrast sensitivity (Dorr et al., 2013), as well as for chromatic contrast sensitivity (Kim et al., 2017).

The CSF can also be measured quickly using electroencephalography (EEG). The fastest method – the sweep VEP (e.g., Seiple et al., 1984; Allen et al., 1986; Norcia et al., 1989) – uses gratings contrast-reversing at 12 Hz, for example, and increasing either in SF or contrast over 10 s. Participants view these stimuli while their electrophysiological scalp response is recorded. A discrete Fourier transform is applied to the signal to measure the amplitude of the electrophysiological response at the frequency of the contrast-reversals. The amplitude associated with a particular SF is related to the contrast threshold at this SF. The measure takes only about 10 s per sweep, in addition to the time needed to install the electrodes – as little as two occipital electrodes and a reference. However, it requires EEG equipment as well as knowledge of how to use it, which are not readily accessible to everybody who might want to measure the CSF.

The Campbell-Robson Chart

The Campbell-Robson chart (Campbell and Robson, 1964; Ratliff, 1965) is a highly popular figure used in several psychophysics and visual perception textbooks to illustrate the CSF (e.g., Lu and Dosher, 2014; Wolfe et al., 2018). The chart depicts a grating which varies logarithmically in spatial frequency (SF) from left to right and in contrast from bottom to top. Campbell and Robson's (1964) intuition was that the boundary between the grating and the homogeneous gray area (below threshold) would trace the shape of the observer's own CSF. Surprisingly, it appears that no one has tested if the Campbell-Robson chart can indeed be used to measure the CSF.

Here, we asked adults and children with normal or corrected-to-normal vision to adjust directly the *truncated log-parabola* to the curve visible on a rendition of the Campbell-Robson chart. The truncated log-parabola was chosen here because of its simplicity and its good fit to the CSF curves of observers with normal vision (e.g., Lesmes et al., 2010). Participants adjusted the four parameters of this curve in only three mouse clicks. Then, we attempted to predict the parameters of the truncated log-parabola adjusted on the contrast thresholds of seven SFs estimated using the QUEST algorithm from those of the truncated log-parabola adjusted directly on the Campbell-Robson chart in adults (N = 100) and in children (N = 81). If successful, this procedure, which requires a mere three clicks on the computer mouse, would speed up the measurement of the CSF to under a minute.

[1]https://michaelbach.de/fract/download.html

MATERIALS AND METHODS

Participants

A total of 103 adults were recruited at Université de Montréal (N = 83) or at The University of British Columbia (N = 20). One participant was excluded due to technical problems during data collection, and data for six additional outlier participants were removed (see section "Results"). The final adult sample comprised 96 individuals between 18 and 35 years of age [median of 22.0 (interquartile range: 3); 34 men]. Ninety-nine children were recruited at The University of British Columbia. An incomplete dataset was obtained for 15 children due to various reasons: technical problems (N = 1), the child not feeling well (N = 1), or general restlessness, distractedness, not following instructions or lack of time (N = 13). Eight outlier children's data were removed and two additional children's data were removed because the truncated log-parabola did not properly adjust to the thresholds measured with the gold-standard method (see section "Results"). The final sample comprised 74 children between 4.9 and 17.7 years old [median of 11.1 (interquartile range: 4.9)]. All participants were neurotypical and received a small monetary compensation for their participation. Participants recruited at Université de Montréal (all adults) reported having normal or corrected-to-normal vision, and participants recruited at The University of British Columbia (all children and some adults) completed the Regan chart to determine monocular distance visual acuity. The study was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), and approved by the Children's and Women's Research Ethics Board at the University of British Columbia as well as the Comité d'éthique de la recherche en éducation et en psychologie (CEREP) at Université de Montréal. Informed consent was obtained from each adult participant, or parent/guardian as well as verbal or written assent from each child participant.

Apparatus

The experimental programs were run on Macintosh computers in the Matlab (MathWorks Inc.) environment, using functions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). Participants were seated in a dim-lighted room. For participants recruited at Université de Montréal (only adults), all stimuli were presented on 27-inch Asus VG278H monitors (1920 × 1080 pixels at 120 Hz), calibrated to allow linear manipulation of luminance. Luminance ranged from 0.33 to 245 cd/m^2 (measured with a Samsung SyncMaster 753df photometer). A chinrest was used to maintain a constant viewing distance. For participants recruited at The University of British Columbia (all children and some adults), stimuli were presented on a 24-inch A1267 Apple Cinema Display (1920 × 1200 pixels at 60 Hz); luminance ranged from 1.6 to 159 cd/m^2 (measured with a Minolta LS-110 photometer).

Procedure

Three-Click CSF Method

Each participant completed three runs of the three-click CSF. On each run, a Campbell-Robson chart was generated to cover the whole computer screen (**Figure 1A**). SFs varied logarithmically

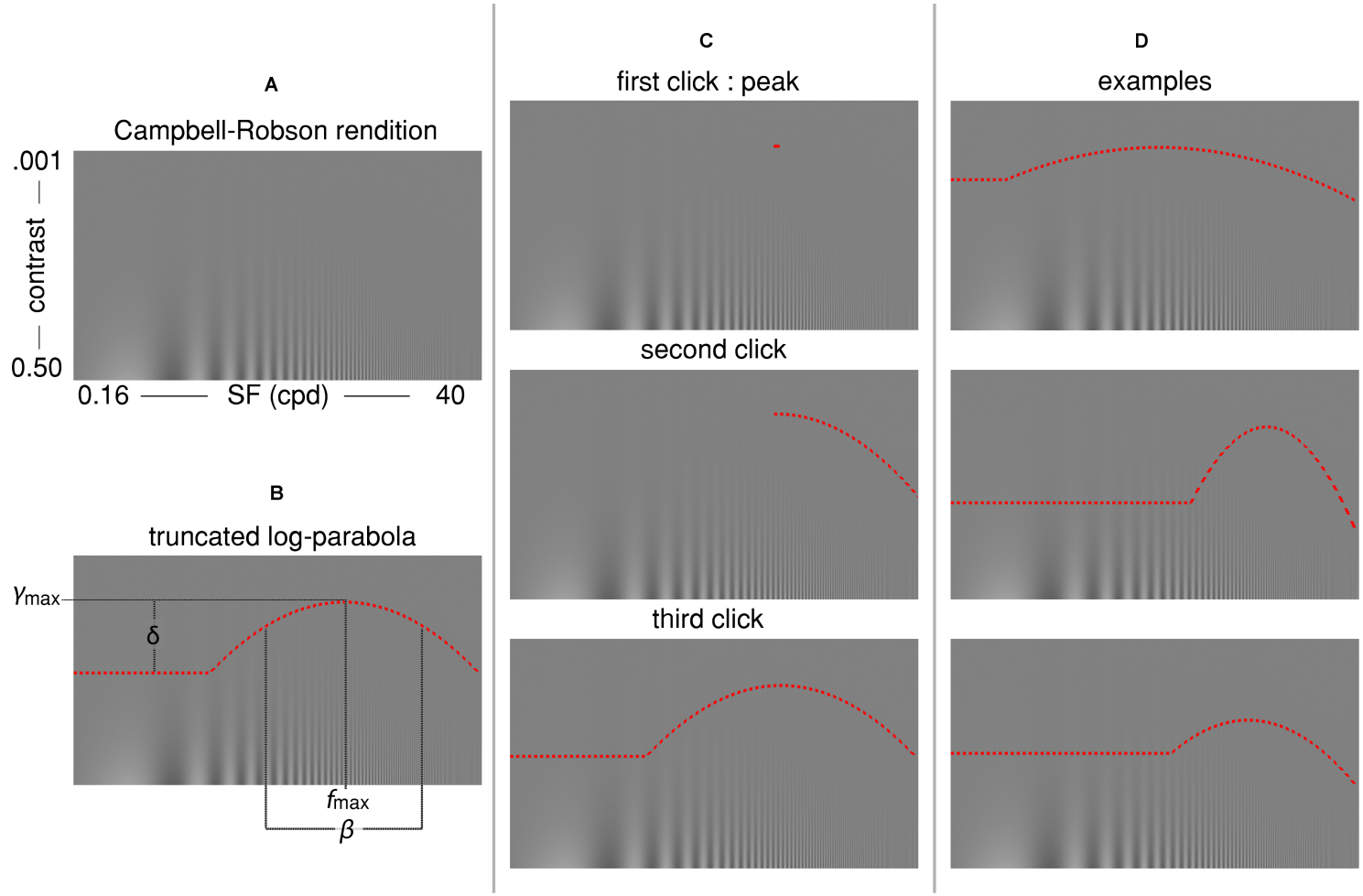


FIGURE 1 | (A) The Campbell-Robson chart. **(B)** The truncated log-parabola is defined by four parameters: f_{max} and y_{max} are the peak of the curve, respectively, the frequency at which it peaks and the sensitivity at this point; β is the Full-Width at Half-Maximum (FWHM); and δ is the truncation parameter in the lower spatial frequency range. **(C)** Three computer mouse clicks are required for adjusting the curve: The first click determines f_{max} and y_{max}, the coordinates of the peak. The second click determines β, the sensitivity for high spatial frequencies. The last click determines the truncation parameter δ, or the sensitivity to low spatial frequencies. **(D)** Three examples of the shape individual curves could take, depending on parameters.

from 0.16 to 40 cycles per degree in 14.4° of visual angle (from left to right) and Michelson contrast varied logarithmically from 0.001 to 0.5 (from bottom to top) in 8.12° of visual angle for the Université de Montréal participants and 9° of visual angle for The University of British Columbia participants. We used the noisy-bit method which uses spatial pooling to increase luminance resolution beyond 8 bits (Allard and Faubert, 2008). Participants were instructed to adjust the truncated log-parabola directly onto the rendition of the Campbell-Robson chart (see **Figure 1D** for examples of the truncated log-parabola with differing parameters). Adults used a computer mouse to adjust the curve's parameters, children used either a computer mouse or a game controller.

The truncated log-parabola is defined as

$$\mathrm{f}(x) = \begin{cases} \log_{10}(y_{\max}) - \delta, \\ \quad \text{if } x < f_{\max} \text{ and } \mathrm{f}(x) < \log_{10}(y_{\max}) - \delta \\ \log_{10}(y_{\max}) - \left(\dfrac{\log_{10} x - \log_{10} f_{\max}}{\beta \cdot \log_{10}(2)/2} \right)^2, \\ \quad \text{otherwise} \end{cases}$$

Equation 1 – truncated log-parabola

where f(x) is the contrast sensitivity threshold, x is the spatial frequency, $y_{\max}$ is the peak sensitivity, $f_{\max}$ is the SF of this peak sensitivity, β is the Full-Width-at-Half-Maximum (FWHM) of the log-parabola in octaves, and δ is the truncation parameter. The curve was specified by the latter four quantities which were determined by three computer mouse clicks (see **Figure 1C**). First, participants were instructed to place a small horizontal line segment at the highest point at which they could see "stripes" on the displayed Campbell-Robson chart. This first click served to establish the peak contrast sensitivity ($y_{\max}$) and the SF at which it peaked ($f_{\max}$). Next, the right half of a log-parabola peaking at ($f_{\max}$, $y_{\max}$) appeared as a dotted red curve on the displayed Campbell-Robson chart. Participants were instructed to adjust the width of this curve so that the "stripes" were visible underneath it but invisible over it by moving the computer mouse along the x-axis, and to click on the mouse button when they were satisfied with the adjustment. This second click defined the FWHM of the log-parabola (β), or the contrast sensitivity for the mid to high SF. Finally, a complete truncated parabola peaking at ($f_{\max}$, $y_{\max}$) and with a FWHM equal to β was overlaid on the Campbell-Robson chart as a dotted red curve. Participants were instructed to adjust the height of the truncated portion of this curve so that the "stripes" were visible underneath it but invisible over it by moving the computer mouse along the y-axis, and to click on the computer mouse button when they were satisfied with the adjustment. This third and last click determined the truncation parameter (δ) of the truncated log-parabola, or the contrast sensitivity for low SFs. These instructions were given to adults while they performed their first of three three-click CSF runs. Children were given adapted instructions in the form of an animated PowerPoint story prior to their three three-click CSF runs. Participants were not given any instruction about fixation and could freely explore the stimulus.

Gold-Standard CSF Method

We compared the four truncated log-parabola parameters adjusted on the Campbell-Robson chart with the four parameters of the truncated log-parabola fitted to the contrast sensitivity thresholds measured for sinusoidal gratings of seven different SFs using the QUEST staircase method (Watson and Pelli, 1983). This psychophysical adaptive method has often been used to measure the CSF (e.g., Burr et al., 1994; Carrasco et al., 2000).

All participants completed 336 trials (48 trials for each SF), divided into four blocks of 84 trials. Michelson contrast thresholds were measured independently for each of seven SFs: 0.5, 0.99, 1.96, 3.87, 7.66, 15.16, and 30 cycles per degree. Gratings were revealed through a Gaussian window with a FWHM equal to 2° of visual angle. Noisy-bit dithering was applied to every stimulus (Allard and Faubert, 2008). The starting values of the contrast threshold estimates were determined using the average of the Gabor data for all subjects from ModelFest (Carney et al., 2000). The order of presentation of the different SFs was randomized. The gratings were presented equiprobably horizontally or vertically on each trial. Participants were instructed to indicate the orientation of the gratings using the arrows on the computer keyboard. As with the three-click CSF, children were given adapted instructions and could answer using the buttons of a game controller. Contrast was adjusted using the QUEST algorithm for each SF, independently, to reach a correct rate of 82%.

RESULTS

Gold-Standard CSF Evaluation

Threshold contrast levels were obtained from the final QUEST contrast threshold estimates for every tested SF. To compare the QUEST to the Campbell-Robson chart adjustment, we fitted a truncated log-parabola (Equation 1 and illustrated in **Figure 1**) to the sensitivity levels measured by QUEST for each participant. Apart from a few exceptions, the function fitted very well to these data [Adults: median R^2 = 0.96 (interquartile range: 0.04); Children: median R^2 = 0.93 (interquartile range: 0.07)]. We excluded two children from further analysis because the R^2 of their fit was lower than 0.5. We also excluded two adults and five children because at least one of their fitted parameters was an outlier ($Z > 3.00$).

The average of all curves measured with this method are presented in gray in **Figure 2** (adult data) and **Figure 3** (child data). To obtain these figures, we adjusted a truncated log-parabola on the sensitivity levels obtained by each participant. Then, we averaged all the curves point-by-point.

Three-Click CSF Method Evaluation

We excluded four adults and three children from further analysis because at least one of their four three-click CSF parameters were outliers ($Z > 3.00$). Average parameters of the truncated log-parabola adjusted on the Campbell-Robson chart for each of the three runs are presented in **Tables 1** and **2**. The point-by-point average of all curves measured with the three-click CSF method

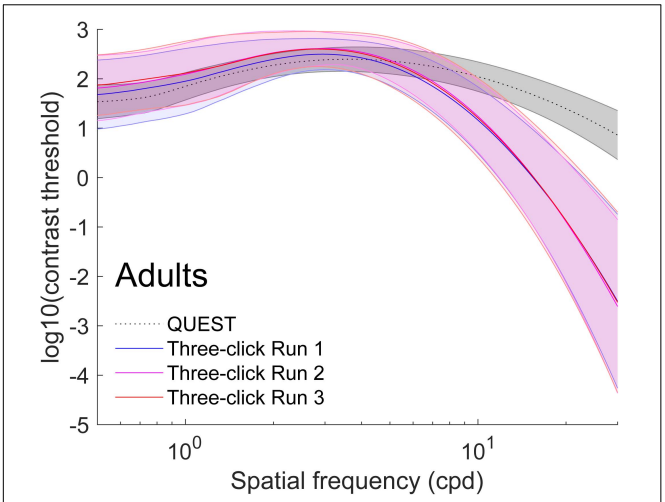


FIGURE 2 | Average (±1 standard deviation in shaded area) contrast sensitivity functions for adults (*N* = 96). To obtain these curves, we first computed four truncated log-parabola curves for each individual: one adjusted on their sensitivity measured using QUEST, and three for each run adjusted onto the Campbell-Robson. Then, each curve was averaged across participants point-by-point. The average as measured using QUEST to find thresholds is presented in dotted gray (±1 standard deviation in shaded area) and the average curve adjustments on the Campbell-Robson chart are presented in blue (run 1), pink (run 2), and red (run 3).

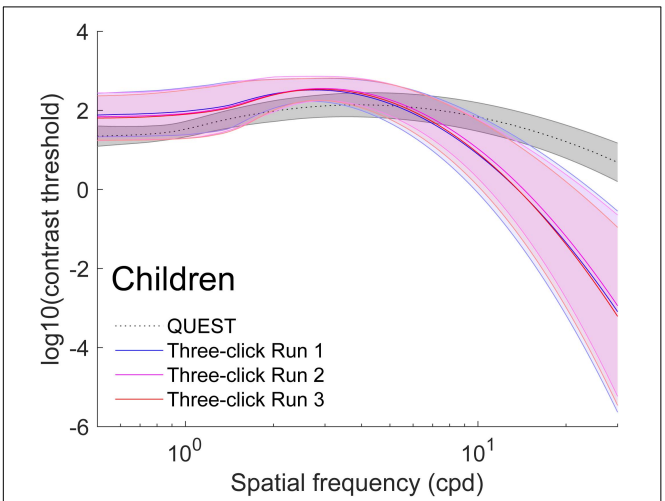


FIGURE 3 | Average (±1 standard deviation in shaded area) contrast sensitivity functions for children (*N* = 74). To obtain these curves, we first computed four truncated log-parabola curve for each individual: one adjusted on their sensitivity measured using QUEST, and three for each run adjusted onto the Campbell-Robson. Then, each curve was averaged across participants point-by-point. The average, as measured using QUEST to find thresholds, is presented in dotted gray (±1 standard deviation in shaded area), and the average curve adjustments on the Campbell-Robson chart are presented in blue (run 1), pink (run 2), and red (run 3).

are plotted for each run, in red, pink and blue on **Figure 2** (adult data) and **Figure 3** (child data).

On average, the Campbell-Robson chart curve adjustment overestimated low-to-mid SF sensitivity and underestimated high SF sensitivity. Specifically, as observed in **Figure 2**, mean peak contrast sensitivity was higher for the Campbell-Robson than the gold-standard measure, as determined using a paired *t* test [y_{max}; Adults: $t(95) = 10.45$; $p < 0.0001$; Children: $t(73) = 10.82$; $p < 0.0001$]. While the difference between the peaks of the two methods seems bigger for children than for adults, the method × age group interaction was not significant [$F(1,168) = 0.37$; $p = 0.55$]. At higher SF, mean SF at which contrast sensitivity peaked [f_{max}; $t(95) = -4.59$; $p < 0.001$; Children: $t(73) = -9.66$; $p < 0.0001$] and mean log-parabola width [β; Adults: $t(95) = -17.25$; $p < 0.0001$; Children: $t(73) = -17.35$; $p < 0.0001$] were lower for the three-click CSF than the gold-standard measure. At low SF, the parameter δ statistically differed for adults but not for children [δ; Adults: $t(95) = 2.31$; $p = 0.02$; Children: $t(73) = 0.32$; $p = 0.75$], but since it measures the difference in sensitivity between the peak and sensitivity at low SF, it is more difficult to interpret. Instead, we compared sensitivity at the lowest spatial frequency (0.5 cpd). At that specific point, it is higher when measured with the three-click method [Adults: $t(95) = 3.60$; $p < 0.001$; Children: $t(73) = 7.76$; $p < 0.0001$].

Extracting Information From the Campbell-Robson Chart

The main objective of this paper was to verify if it is possible to extract CSF information directly from adjusting a specific curve to the Campbell-Robson chart. To investigate this question, we adopted a data-driven approach. More specifically, we used machine learning methods to verify if we can predict the CSF using only the parameters adjusted on the Campbell-Robson chart.

We trained models to predict the four parameters evaluated from the individual gold-standard thresholds measured with QUEST using the four individual parameters adjusted directly on the Campbell-Robson chart with the three-click CSF method. We trained different sets of models for children and for adults and evaluated their performance using 6-fold cross-validation. Specifically, we did 100 iterations of the following. First, the dataset was randomly split into a test set of 16 random observations (12 for children) and a training set. Then, the training set was randomly split evenly into five sets of the same size as the test set. Each of these five sets in turn was set aside and SVMs with Gaussian kernel whose FWHM varied between 0.1 and 5 in increments of 0.1 were trained on the remaining 4/5 sets. Their Root-Mean-Square Errors (RMSE) were evaluated on the 1/5 set that was put aside. The best model (lowest RMSE) was chosen across the five folds and applied to calculate a RMSE and a R^2 on the independent test set. To verify how much information we gain by adjusting the curve on the Campbell-Robson chart more than once, this 100-iteration procedure was repeated three times: (1) using the parameters of the first run, (2) using the parameters of the first two runs, and (3) using the parameters of the three runs.

To compare the RMSE we obtained with the null hypothesis, we also ran a permutation test. The exact same steps were followed in the permutation test, but on each iteration the

TABLE 1 | Average parameters (β, δ, f_{max}, and y_{max}) defining the truncated log-parabola for adult participants (N = 96) when the contrast sensitivity function was measured using the Campbell-Robson chart (Runs 1, 2, and 3 independently), and when it was measured using QUEST to find perceptive thresholds (Q).

Task	Run	β	δ	f_{max}	y_{max}
Campbell-Robson	1	3.24 (SD = 0.96)	1.08 (SD = 0.76)	2.96 (SD = 1.23)	550.04 (SD = 364.52)
	2	3.15 (SD = 0.91)	1.06 (SD = 0.75)	3.03 (SD = 1.33)	729.50 (SD = 461.71)
	3	3.19 (SD = 0.85)	1.01 (SD = 0.70)	2.96 (SD = 1.32)	722.51 (SD = 504.89)
QUEST		5.13 (SD = 1.02)	0.90 (SD = 0.36)	3.62 (SD = 0.82)	294.33 (SD = 135.19)

TABLE 2 | Average parameters (β, δ, f_{max}, and y_{max}) defining the truncated log-parabola for child participants (N = 74) when the contrast sensitivity function was measured using the Campbell-Robson chart (Runs 1, 2, and 3 independently), and when it was measured using QUEST to find perceptive thresholds (Q).

Task	Run	β	δ	f_{max}	y_{max}
Campbell-Robson	1	3.23 (SD = 0.97)	0.83 (SD = 0.49)	2.68 (SD = 0.89)	537.55 (SD = 371.20)
	2	3.17 (SD = 0.89)	0.88 (SD = 0.54)	2.82 (SD = 0.88)	554.14 (SD = 333.73)
	3	3.09 (SD = 0.83)	0.86 (SD = 0.43)	2.81 (SD = 0.89)	527.71 (SD = 295.59)
QUEST		5.01 (SD = 0.75)	0.84 (SD = 0.24)	3.94 (SD = 0.85)	168.18 (SD = 86.48)

dataset was randomly permuted so that independent variables (parameters adjusted on the Campbell-Robson chart) were matched with dependent (parameters measured with QUEST) variables of a different, random participant. In other words, we ran the same analysis but with data for which the independent and dependent variables were not linked.

For adults, the only parameter for which the prediction was better with the non-permuted data than the permuted data is y_{max}, the maximal sensitivity (highest Y point). For children, that parameter is f_{max} (the frequency at which the maximal sensitivity is attained). RMSE and R^2 are presented, respectively, in **Figures 4**, **5** for the prediction of (**Figure 4A**) y_{max} for adults and (**Figure 4B**) f_{max} for children. Progressively adding data from more than one run of fitting the curve on the Campbell-Robson chart (i.e., twice or three times) lowers the error for predicting y_{max} for adults, as can be observed in **Figure 4A**. For predicting y_{max} for adults, RMSE is an average of 0.88 using one run of data, and lowers to 0.78 using two runs of data, and 0.74 using three runs of data. We can predict 38.8% of the variance using three runs of data. Using the best model from the 100 iteration, we predict 45.7% of the variance of y_{max}. For the other three parameters measured with adults, the RMSE after three runs was of 0.88, 0.86, and 0.92, respectively, for beta, delta, and f_{max}, equivalent to predicting 2.3, 14.6, and 6.2% of the variance of new data.

For predicting f_{max} for children, RMSE is of an average of 0.87 using one run of data, and does not lower when adding successive runs of data. Average RMSE using three runs of data is of 0.88. This does not translate to being able to predict much of the variance for f_{max}, the best R^2 is of 0.068 using two runs of data. In **Supplementary Material**, RMSE and R^2 values are presented in the form of graphs for each iteration, each parameter and each number of runs included.

DISCUSSION

In this article, we examined the mainstream belief, originating from Campbell and Robson's (1964) seminal work, that the boundary between the grating and the homogeneous gray area in the Campbell-Robson chart traces the shape of an observer's own CSF. Participants adjusted the truncated log-parabola directly to the curve visible on the Campbell-Robson chart. We call this procedure the "three-click CSF method." They also completed a gold-standard evaluation of their CSF, using QUEST to adjust the contrast of sinusoidal gratings of different spatial frequencies on a trial-to-trial basis. We found that, on average, the three-click CSF method overestimated low-to-mid SF sensitivity and underestimated the high SF sensitivity compared to our gold-standard method for measuring the CSF. We then trained support vector machine models with radial-basis function kernels to predict the gold-standard CSF from the four parameters obtained with the three-click method. Our goal was to extract as much CSF information as possible – linear and non-linear – from the Campbell-Robson chart. We found that, in adults, 42% of the variance of the maximal sensitivity (the y_{max} parameter of the truncated log-parabola) could be predicted using the best support vector machine model. None of the other parameters could be predicted in adults, nor any parameter whatsoever in children.

Dorr et al. (2017) asked their participants to tap on the peak of the curve visible on a Campbell-Robson chart presented on a tablet – the coordinates of this peak should match closely our f_{max} and y_{max}. They found a linear relationship between these coordinates and the area under the curve obtained using the Quick CSF method (r = 0.63 for x coordinate or f_{max} and r = 0.58 for y coordinate or y_{max}). The area under the curve measure can be understood as a global contrast sensitivity measure. For adults, we also found significant correlations between the area under the curve calculated from our gold-standard data and, first, f_{max} measured with the three-click method in all three runs (r = 0.20, p = 0.047; r = 0.32, $p < 0.001$; r = 0.23, p = 0.03, respectively, for the first, second, and third run) and, second, y_{max} measured with the three-click method in the second and third runs (r = 0.17, *n.s.*; r = 28, p = 0.01; r = 0.30, p = 0.003, respectively, for the first, second, and third run). These correlations, however, were not significant for children.

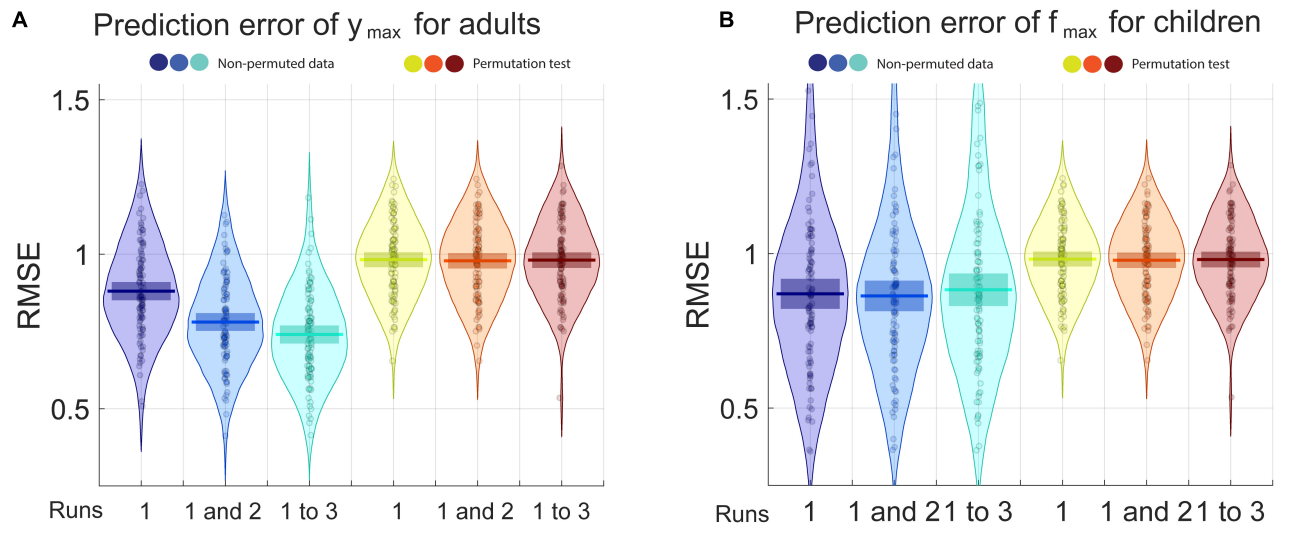


FIGURE 4 | The blue pirate plots show the Root Mean Square Error (RMSE) of the prediction of y_{max}, the truncated log-parabola's maximal height, for adults, and of fmax, the spatial frequency at the curve's maximal height, for children. The darkest, middle and lightest blue pirate plots show the RMSE of the predictions from models based, respectively, on the first, on the first and the second, and on all three three-click CSF runs. The pirate plots with warm colors show the null hypothesis obtained using a permutation test (see also **Supplementary Figures S1, S2**).

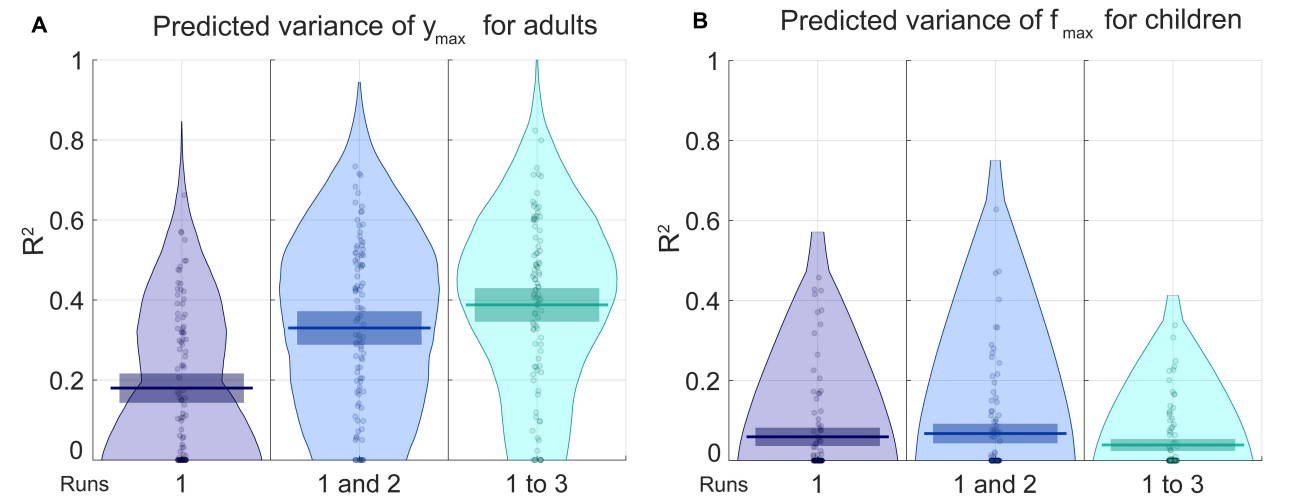


FIGURE 5 | The blue pirate plots show the explained variance (R^2) of the predicted of y_{max}, which defines the truncated log-parabola's maximal height, for adults, and of f_{max}, the spatial frequency at the curve's maximal height, for children. The darkest, middle and lightest blue pirate plots show the R^2 of the predictions from models based, respectively, on the first, on the first and the second, and on all three three-click CSF runs. The pirate plots with warm colors show the null hypothesis obtained using a permutation test (see also **Supplementary Figures S3, S4**).

The discrepancies observed between the three-click CSF and the gold-standard results could be due to one or a combination of the many differences between the methods. For example, while our gold-standard CSF method is objective, the three-click CSF method is subjective. Thus, in the former, QUEST searched for the thresholds associated with a correct rate of 82% for each spatial frequency and participant. In the latter, the correct rate is undefined. The two tasks also differ drastically in stimulus size: 14.42° of visual angle for the Campbell-Robson chart used for the three-click CSF method vs. about 2° of visual angle for the sinusoidal gratings used for the gold-standard method. The three-click CSF method thus required several saccades to foveate the parts of the Campbell-Robson chart relevant to the different adjustment clicks. These eye movements could also have led to confounding grating aftereffects – resulting from the fixation of a contrasted region of the Campbell-Robson chart – for actual gratings in the "gray" regions of the chart. Moreover, each SF occupies a much smaller area in the Campbell-Robson chart (1 pixel or about 0.63 min of arc) than in the gold-standard method

(2° of visual angle). This may also have had an effect on the observed sensitivity, especially for the high SF where the represented SFs change rapidly.

Is it possible that the truncated log-parabola, the curve that we chose to fit onto the Campbell-Robson chart, does not capture the CSF information on the Campbell-Robson chart? This curve fits very well to our gold-standard thresholds [Adults: median R^2 = 0.96 (interquartile range: 0.04); Children: median R^2 = 0.93 (interquartile range: 0.07)]. It also fits very well to the ModelFest thresholds [median R^2 = 0.99 (interquartile range: 0.01)]. In other words, it provides a very good approximation of the gold-standard CSF. In fact, this is one of the reasons it is at the core of the Quick CSF method (Lesmes et al., 2010). Could it be, however, that the truncated log-parabola doesn't fit the curve visible on the Campbell-Robson chart very well but that, nonetheless, this curve contains useful information about the CSF? It is difficult to give objective evidence for or against this possibility. Our participants reported that the curve adjustment was easy, with the possible exception of the third, and last, click that measured the truncation parameter (δ) or the contrast sensitivity for low SFs. A few participants reported dips to which they could not adjust the curve. A more quantitative evidence of the adequacy of the curve is the intra-subject consistency. For all truncated log-parabola parameters and between all pairs of runs, the average Pearson reliability was 0.70 (SD = 0.10; it ranged from 0.54 for δ between childrens' runs 1 and 2 and 0.88 for y_{max} between adults' runs 2 and 3). Thus the truncated log-parabola appears to fit well the Campbell-Robson chart curve at least subjectively. This truncated log-parabola, however, is different from the one that fits the gold-standard CSF.

In sum, the short answer to our opening question is: The Campbell-Robson chart cannot be used to measure the CSF. It does predict contrast sensitivity (y_{max}) but other rapid methods to measure contrast sensitivity are already available and widely used, such as the Pelli-Robson chart. The Campbell-Robson chart should remain a useful educational tool to teach students about the broad shape of the CSF.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Comité d'éthique de la recherche en éducation et en psychologie de l'Université de Montréal. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JT, MW, and FG programmed the experiments. JT recruited and tested the adult participants, drafted the manuscript, and edited by FG and DG. MW recruited and tested the child participants. JT and FG worked on data analysis and interpretation. All authors contributed to the development of the project. DG and FG supervised all aspects of the study.

ACKNOWLEDGMENTS

This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Postgraduate Scholarship – Doctoral program awarded to JT, NSERC Discovery Grants awarded to FG (RGPIN-04777-2014) and to DG (RGPIN-194526-2012), and a Postdoctoral Fellowship granted to MW by the Child and Family Research Institute at the University of British Columbia.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.626466/full#supplementary-material

REFERENCES

Allard, R., and Faubert, J. (2008). The noisy-bit method for digital displays: converting a 256 luminance resolution into a continuous resolution. *Behav. Res. Methods* 40, 735–743. doi: 10.3758/brm.40.3.735

Allen, D., Norcia, A. M., and Tyler, C. W. (1986). Comparative study of electrophysiological and psychophysical measurement of the contrast sensitivity function in humans. *Optom. Vis. Sci.* 63, 442–449. doi: 10.1097/00006324-198606000-00008

Ashworth, B., Aspinall, P. A., and Mitchell, J. D. (1989). Visual function in multiple sclerosis. *Doc. Ophthalmol.* 73, 209–224. doi: 10.1007/bf00155090

Bach, M. (1996). The Freiburg Visual acuity test-automatic measurement of visual acuity. *Optom. Vis. Sci.* 73, 49–53. doi: 10.1097/00006324-199601000-00008

Bach, M. (2006). The Freiburg Visual acuity test-variability unchanged by post-hoc re-analysis. *Graefes Arch. Clin. Exp. Ophthalmol.* 245, 965–971. doi: 10.1007/s00417-006-0474-4

Barollo, M., Contemori, G., Battaglini, L., Pavan, A., and Casco, C. (2017). Perceptual learning improves contrast sensitivity, visual acuity, and foveal crowding in amblyopia. *Restor. Neurol. Neurosci.* 35, 483–496. doi: 10.3233/rnn-170731

Bradley, A., and Freeman, R. D. (1981). Contrast sensitivity in anisometropic amblyopia. *Invest. Ophthalmol. Vis. Sci.* 21, 467–476.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897x00357

Burr, D. C., Morrone, M. C., and Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature* 371, 511–513. doi: 10.1038/371511a0

Campbell, F. W., and Robson, J. G. (1964). Application of Fourier analysis to the modulation response of the eye. *J. Opt. Soc. Am.* 54:581.

Carney, T., Tyler, C. W., Watson, A. B., Makous, W., Beutter, B., Chen, C. C., et al. (2000). "Modelfest: year one results and plans for future years," in *Proceedings of the SPIE International Society for Optics and Photonics: Human Vision and Electronic Imaging V*, Vol. 3959, eds B. E. Rogowitz and T. Pappas (Bellingham, WA: SPIE), 140–152.

Carrasco, M., Penpeci-Talgar, C., and Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. *Vis. Res.* 40, 1203–1215. doi: 10.1016/s0042-6989(00)00024-9

Collins, J. W., and Carney, L. G. (1990). Visual performance in high myopia. *Curr. Eye Res.* 9, 217–223. doi: 10.3109/02713689009044516

De Valois, R. L., and De Valois, K. K. (1990). *Spatial Vision*. Oxford, UK: Oxford University Press.

Dorr, M., Lesmes, L. A., Lu, Z.-L., and Bex, P. J. (2013). Rapid and reliable assessment of the contrast sensitivity function on an iPad. *Invest. Ophthalmol. Vis. Sci.* 54, 7266–7273. doi: 10.1167/iovs.13-11743

Dorr, M., Lesmes, L. A., Elze, T., Wang, H., Lu, Z. L., and Bex, P. J. (2017). Evaluation of the precision of contrast sensitivity function assessment on a tablet device. *Sci. Rep.* 7, 1–11. doi: 10.1038/srep46706

Elliott, D. B., Gilchrist, J., and Whitaker, D. (1989). Contrast sensitivity and glare sensitivity changes with three types of cataract morphology: are these techniques necessary in a clinical evaluation of cataract? *Ophthalmic Physiol. Opt.* 9, 25–30. doi: 10.1111/j.1475-1313.1989.tb00800.x

Elliott, D. B., and Hurst, M. A. (1990). Simple clinical techniques to evaluate visual function in patients with early cataract. *Optom. Vis. Sci.* 67, 822–825. doi: 10.1097/00006324-199011000-00006

Ginsburg, A. P. (1984). A new contrast sensitivity vision test chart. *Optom. Vis. Sci.* 61, 403–407. doi: 10.1097/00006324-198406000-00011

Hess, R., and Woo, G. (1978). Vision through cataracts. *Invest. Ophthalmol. Vis. Sci.* 17, 428–435.

Hess, R. F., and Howell, E. R. (1977). The threshold contrast sensitivity function in strabismic amblyopia: evidence for a two type classification. *Vis. Res.* 17, 1049–1055. doi: 10.1016/0042-6989(77)90009-8

Howell, E. R., Mitchell, D. E., and Keith, C. G. (1983). Contrast thresholds for sine gratings of children with amblyopia. *Invest. Ophthalmol. Vis. Sci.* 24, 782–787.

Huang, C. B., Tao, L. M., Zhou, Y. F., and Lu, Z. L. (2007). Treated amblyopes remain deficient in spatial vision: a contrast sensitivity and external noise study. *Vis. Res.* 47, 22–34. doi: 10.1016/j.visres.2006.09.015

Kim, Y. J., Reynaud, A., Hess, R. F., and Mullen, K. T. (2017). A normative data set for the clinical assessment of achromatic and chromatic contrast sensitivity using a qCSF approach. *Invest. Ophthalmol. Vis. Sci.* 58, 3628–3636. doi: 10.1167/iovs.17-21645

Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in psychtoolbox-3? *Perception* 36(ECVP Abstract Suppl.), 14.

Kleiner, R. C., Enger, C., Alexander, M. F., and Fine, S. L. (1988). Contrast sensitivity in age-related macular degeneration. *Arch. Ophthalmol.* 106, 55–57.

Leat, S. J., and Woo, G. C. (1997). The validity of current clinical tests of contrast sensitivity and their ability to predict reading speed in low vision. *Eye* 11, 893–899. doi: 10.1038/eye.1997.228

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Percept. Psychophys.* 63, 1279–1292. doi: 10.3758/bf03194543

Lesmes, L. A., Lu, Z.-L., Baek, J., and Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J. Vis.* 10, 17.1–21. doi: 10.1167/10.3.17

Levi, D. M., Waugh, S. J., and Beard, B. L. (1994). Spatial scale shifts in amblyopia. *Vis. Res.* 34, 3315–3333. doi: 10.1016/0042-6989(94)90067-1

Lewis, T. L., Maurer, D., Tytla, M. E., Bowering, E. R., and Brent, H. P. (1992). Vision in the "good" eye of children treated for unilateral congenital cataract. *Ophthalmology* 99, 1013–1017. doi: 10.1016/s0161-6420(92)31857-3

Lieberman, H. R., and Pentland, A. P. (1982). Microcomputer-based estimation of psychophysical thresholds: the best PEST. *Behav. Res. Methods Instrum.* 14, 21–25. doi: 10.3758/bf03202110

Liou, S. W., and Chiu, C. J. (2001). Myopia and contrast sensitivity function. *Curr. Eye Res.* 22, 81–84. doi: 10.1076/ceyr.22.2.81.5530

Lu, Z., and Dosher, B. (2014). *Visual Psychophysics*. Cambridge, MA: MIT Press.

Marmor, M. F. (1986). Contrast sensitivity versus visual acuity in retinal disease. *Br. J. Ophthalmol.* 70, 553–559. doi: 10.1136/bjo.70.7.553

Morrison, D. J., and Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychon. Bull. Rev.* 8, 454–469. doi: 10.3758/bf03196180

Norcia, A. M., Tyler, C. W., Hamer, R. D., and Wesemann, W. (1989). Measurement of spatial contrast sensitivity with the swept contrast VEP. *Vis. Res.* 29, 627–637. doi: 10.1016/0042-6989(89)90048-5

Nordmann, J. P., Saraux, H., and Roullet, E. (1987). Contrast sensitivity in multiple sclerosis. *Ophthalmologica* 195, 199–204. doi: 10.1159/000309813

Pardhan, S., and Gilchrist, J. (1991). The importance of measuring binocular contrast sensitivity in unilateral cataract. *Eye* 5, 31–35. doi: 10.1038/eye.1991.6

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897x00366

Pelli, D. G., and Bex, P. (2013). Measuring contrast sensitivity. *Vis. Res.* 90, 10–14. doi: 10.1016/j.visres.2013.04.015

Pelli, D. G., Robson, J. G., and Wilkins, A. J. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clin. Vis. Sci.* 2, 187–199.

Ratliff, F. (1965). *Mach Bands: Quantitative Studies on Neural Networks*, Vol. 2. San Francisco, CA: Holden-Day.

Reeves, B. C., Wood, J. M., and Hill, A. R. (1991). Vistech VCTS 6500 charts–within-and between-session reliability. *Optom. Vis. Sci.* 68, 728–737. doi: 10.1097/00006324-199109000-00010

Regan, D. (1991). The Charles F. Prentice Award Lecture 1990: specific tests and specific blindnesses: keys, locks, and parallel processing. *Optom. Vis. Sci.* 68, 489–512. doi: 10.1097/00006324-199107000-00001

Regan, D., and Neima, D. (1983). Low-contrast letter charts as a test of visual function. *Ophthalmology* 90, 1192–1200. doi: 10.1016/s0161-6420(83)34407-9

Robson, J. G. (1993). "Contrast sensitivity: one hundred years of clinical measurement," in *Contrast Sensitivity*, ed. D. M. Lam (Cambridge, MA: MIT Press), 253–267.

Rubin, G. S. (1988). Reliability and sensitivity of clinical contrast sensitivity tests. *Clin. Vis. Sci.* 2, 169–177.

Ruiz-Soler, M., and Beltran, F. S. (2006). Face perception: an integrative review of the role of spatial frequencies. *Psychol. Res.* 70, 273–292. doi: 10.1007/s00426-005-0215-z

Seiple, W. H., Kupersmith, M. J., Nelson, J. I., and Carr, R. E. (1984). The assessment of evoked potential contrast thresholds using real-time retrieval. *Invest. Ophthalmol. Vis. Sci.* 25, 627–631.

Shandiz, J. H., Derakhshan, A., Daneshyar, A., Azimi, A., Moghaddam, H. O., Yekta, A. A., et al. (2011). Effect of cataract type and severity on visual acuity and contrast sensitivity. *J. Ophthalmic Vis. Res.* 6, 26–31.

Sjöstrand, J. (1981). Contrast sensitivity in children with strabismic and anisometropic amblyopia. A study of the effect of treatment. *Acta Ophthalmol.* 59, 25–34. doi: 10.1111/j.1755-3768.1981.tb06706.x

Stangos, N., Voutas, S., Topouzis, F., and Karampatakis, V. (1995). Contrast sensitivity evaluation in eyes predisposed to age-related macular degeneration and presenting normal visual acuity. *Ophthalmologica* 209, 194–198. doi: 10.1159/000310612

Sunness, J. S., Rubin, G. S., Applegate, C. A., Bressler, N. M., Marsh, M. J., Hawkins, B. S., et al. (1997). Visual function abnormalities and prognosis in eyes with age-related geographic atrophy of the macula and good visual acuity. *Ophthalmology* 104, 1677–1691. doi: 10.1016/s0161-6420(97)30079-7

Superstein, R., Boyaner, D., Overbury, O., and Collin, C. (1997). Glare disability and contrast sensitivity before and after cataract surgery. *J. Cataract Refract. Surg.* 23, 248–253. doi: 10.1016/s0886-3350(97)80349-5

Thorn, F., Corwin, T. R., and Comerford, J. P. (1986). High myopia does not affect contrast sensitivity. *Curr. Eye Res.* 5, 635–640. doi: 10.3109/02713688609015130

Watson, A. B., and Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* 33, 113–120. doi: 10.3758/bf03202828

Wolfe, J. M., Kluender, K. R., Levi, D. M., Bartoshuk, L. M., Herz, R. S., Klatzky, R. L., et al. (2018). *Sensation and Perception*. New York, NY: Sinauer Associates.

Wolkstein, M., Atkin, A., and Bodis-Wollner, I. (1980). Contrast sensitivity in retinal disease. *Ophthalmology* 87, 1140–1149. doi: 10.1016/s0161-6420(80)35112-9

Woods, R. L., Tregear, S. J., and Mitchell, R. A. (1998). Screening for ophthalmic disease in older subjects using visual acuity and contrast sensitivity. *Ophthalmology* 105, 2318–2326. doi: 10.1016/s0161-6420(98)91235-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

CLINICAL TRIAL
published: 09 April 2021
doi: 10.3389/fnins.2021.622729

A Randomized Clinical Trial Comparing Eyetronix Flicker Glass and Patching for Treatment of Amblyopia in Children Reveals Similar Improvements in Vision

Seung Hyun Min[1,2†], Shijia Chen[1†], Jinling Xu[1], Bingzhen Chen[1], Hui Chen[1], Yuwen Wang[1], Jiawei Zhou[1] and Xudong Yu[1*]*

[1] School of Ophthalmology and Optometry and Eye Hospital, State Key Laboratory of Ophthalmology, Optometry and Vision Science, Wenzhou Medical University, Wenzhou, China, [2] Department of Ophthalmology and Visual Sciences, McGill Vision Research, McGill University, Montreal, QC, Canada

OPEN ACCESS

Edited by:
Peter J. Bex,
Northeastern University, United States

Reviewed by:
Simon Grant,
City University of London, United Kingdom
Krista Rose Kelly,
Retina Foundation of the Southwest, United States

***Correspondence:**
Jiawei Zhou
zhoujw@mail.eye.ac.cn
Xudong Yu
yxd@mail.eye.ac.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: 29 October 2020
Accepted: 26 February 2021
Published: 09 April 2021

Citation:
Min SH, Chen S, Xu J, Chen B, Chen H, Wang Y, Zhou J and Yu X (2021) A Randomized Clinical Trial Comparing Eyetronix Flicker Glass and Patching for Treatment of Amblyopia in Children Reveals Similar Improvements in Vision. Front. Neurosci. 15:622729. doi: 10.3389/fnins.2021.622729

Purpose: Recently, Eyetronix Flicker Glass (EFG) has been introduced as a novel treatment for amblyopia. It alternatively deprives the visual input of each eye rapidly (e.g., 7 Hz). However, whether it is comparable with standard patching therapy is unclear. In this randomized clinical trial, we evaluate the efficacy of an EFG therapy as treatment for amblyopia in children and compare it to the patching therapy.

Methods: We tested 31 children (aged 4–13 years) with amblyopia. They were assigned into one of the two treatment groups and were treated for 12 weeks. The first group was treated with EFG for 1 h/day (Flicker Group) and the latter with a standard patch (Patching Group) for 2 h/day. We designated changes from baseline in best-corrected visual acuity (BCVA) of the amblyopic eye as our primary outcome. Changes from baseline in other visual outcomes, such as contrast sensitivity, stereopsis, and fusional vergence range were measured as secondary outcome.

Results: BCVA improved significantly at 12 weeks relative to baseline in both the Flicker (0.13 ± 0.11 logMAR; mean ± SD) and Patching Groups (0.21 ± 0.14 logMAR). However, the improvements were not significantly different between groups ($p = 0.13$). Contrast sensitivity also significantly improved at 3 and 12 cycles/degree between baseline and 12 weeks in both groups (p's < 0.05). However, stereopsis and fusion range did not improve significantly in both groups.

Conclusion: An EFG therapy and patching improved BCVA similarly for children with amblyopia at 12 weeks. Both therapies improved the contrast sensitivity at 3 and 12 cycles per degree (cpd); however, only patching improved the contrast sensitivity at 6 cpd. Both therapies did not benefit binocular visual functions (stereopsis and fusional vergence range). We believe that EFG can be an additional choice for therapy.

Clinical Trial Registration: chictr.org number: ChiCTR2000034436.

Keywords: amblyopia, eyetronix flicker glass, randomized controlled trial, patching, visual acuity

INTRODUCTION

Amblyopia is a neurodevelopmental disorder from abnormal visual experience during the critical period (Hess and Thompson, 2015; Levi et al., 2015). Early treatment is necessary for a proper recovery. In clinics, amblyopic children are prescribed with a patching therapy, which deprives the fellow eye to force the amblyopic eye to work (De Buffon, 1743). Younger children have been reported to benefit more than older counterparts (Fronius et al., 2014). It improves the visual acuity (VA) of the amblyopic eye by more than two lines of logMAR in over 50% of cases (Repka et al., 2003; Scheiman et al., 2005). However, it discomforts the children because only the amblyopic eye is opened (Stewart et al., 2004; Fronius et al., 2014). Alternative therapies, such as atropine penalization and Bangerter filters, have been used, but their premise is the same as patching: the input in the fellow eye is inhibited. These monocular therapies have been shown to bring suboptimal results, poor compliance, and harm to binocular function (Searle et al., 2002; Hess and Thompson, 2013; Wallace et al., 2013; Papageorgiou et al., 2019).

Recently, Eyetronix Flicker Glass (EFG) has been introduced as a novel treatment for amblyopia. The lenses in EFG for both eyes flicker alternatively between opaque ("off") and transparent ("on"). Hence, EFG therapy differs from the patching therapy because it alternatively deprives each monocular input rapidly (e.g., 7 Hz) and enables the patients to see with both eyes throughout the treatment (Spierer et al., 2010; Wang et al., 2016). Moreover, it differs from binocular therapies with a dichoptic training protocol that have been developed in the last decade. A dichoptic training protocol displays stimuli separately (i.e., dichoptically) and simultaneously to each eye in the form of movie viewing (Li et al., 2015b), video gaming (Knox et al., 2012), and perceptual learning (Vedamurthy et al., 2015). On the other hand, patients do not receive binocular input at any moment during an EFG therapy.

When a new amblyopia treatment is developed, one must resolve whether it produces a comparable VA gain as a standard therapy such as patching. Likewise, dichoptic therapies have been extensively studied in randomized controlled trials (RCTs). While some RCTs in children show that their superiority to patching remains elusive (Manh et al., 2018; Holmes et al., 2019), others show that a novel binocular treatment is superior to patching (Kelly et al., 2016). A recent study from Eyetronix® shows that EFG therapy for 12 weeks improves both the VA of the amblyopic eye and stereopsis (Vera-Diaz et al., 2016). However, it does not compare against a control group of patching. We conducted a randomized clinical trial to compare EFG therapy with patching at a non-profit research hospital. Just as in previous RCTs of binocular therapies (Kelly et al., 2016; Birch et al., 2019), we designated changes from baseline in best-corrected VA (BCVA) as our primary outcome. Moreover, VA of the amblyopic eye has been the primary measure in diagnosing and treating amblyopia (Wallace et al., 2018). Changes from baseline in other visual functions, such as contrast sensitivity, stereopsis, and fusional vergence range, are reported as secondary outcome.

MATERIALS AND METHODS

This RCT is listed in the Chinese Clinical Trial Registry[1] with the identifier ChiCTR2000034436. The full trial can be accessed from the Chinese Clinical Trial Registry's website: http://www.chictr.org.cn/showproj.aspx?proj=56030. The research protocol and the informed consent forms were reviewed and approved by the Ethics Committee of the Affiliated Eye Hospital at Wenzhou Medical University (2016-18-Q-11). All data were collected at the Affiliated Eye Hospital of Wenzhou Medical University.

Patients

Thirty-two children (aged 4–13) with untreated, mild-to-severe unilateral amblyopia participated in our study. One of them received another treatment during the study and had to be excluded, so 31 children completed the study. The clinical details of the patients are provided in **Table 1**. We obtained informed consent from their parent or legal guardians. The sample size was determined based on the previous studies of EFG and liquid crystal glasses (LCG) (Spierer et al., 2010; Vera-Diaz et al., 2016). Their guardians volunteered to participate in the study and accepted the random assignment of each patient to either Group 1 (therapy with the EFG) or Group 2 (traditional patching therapy).

We estimated the sample size via power analysis (Hickey et al., 2018). To do so, we extracted from the literature a typical improvement of the amblyopic eye's VA in amblyopic children after patching. Studies show that patching for 12 weeks improves VA by about 0.21–0.35 logMAR (Stewart et al., 2004; Rutstein et al., 2010), which were then averaged for our analysis (i.e., 0.273 logMAR). Also, treatment using flickering goggles improves it by 0.124 logMAR (Vera-Diaz et al., 2014; Vera-Diaz et al., 2016). With a power of 80%, an α level of 0.05, a difference (i.e., absolute effect size) of 0.149 logMAR between the improvements in VA shown by the two treatment methods, and a standard deviation of 0.11 logMAR (Rutstein et al., 2010; Rajavi et al., 2019), we found that the minimum sample size per group would have to be nine patients. To make our study robust, we recruited 16 children for each treatment group, although one of them later withdrew from the study.

Eligibility inclusion criteria: (1) age range of 4 to 13 years; (2) a diagnosis of amblyopia, BCVA of the amblyopic eye equal to or worse than 0.3 (logMAR) in 3- to 5-year age range, and 0.15 (logMAR) in children over 6 years old, and at least a two-line BCVA difference between the two eyes (according to the diagnostic criteria for amblyopia established by the Ophthalmology Branch of the Chinese Medical Association in 2011); (3) myopia of no more than −6.00 diopters (D), hypermetropia of no more than + 9.00 D, astigmatism of less than 3.00 D, anisometropia of at least 1.50 D spherical equivalent, or at least 1.00 D cylindrical equivalent; (4) strabismus of no more than 20 prism diopters (Δ) according to prism and alternate cover test (PACT); and (5) better than a 0.7 logMAR BCVA of the amblyopic eye.

Eligibility exclusion criteria: (1) ocular diseases, such as ptosis, refractive media opacity, fundus disease, and optic neuropathy;

[1] http://www.chictr.org.cn

TABLE 1 | Baseline characteristics of the patients.

	Flicker group (*N* = 15)		Patching group (*N* = 16)	
	N	**%**	***N***	**%**
Gender				
Female	6	40%	8	50%
Age at enrollment (years)				
4 to < 6	9	60%	6	38%
6 to < 8	5	33%	5	31%
8 to ≤13	1	7%	5	31%
Mean ± SD	5.27 ± 1.10		6.38 ± 2.45	
Distance visual acuity of amblyopic eye (logMAR)				
0.15 to < 0.2	1	7%	1	6%
0.2 to < 0.3	3	20%	3	19%
0.3 to < 0.4	3	20%	5	31%
0.4 to < 0.5	3	20%	2	13%
0.5 to < 0.6	3	20%	2	13%
0.6 to < 0.7	1	7%	0	0%
0.7	1	7%	3	19%
Mean ± SD	0.39 ± 0.16		0.38 ± 0.19	
Distance visual acuity of fellow eye (logMAR)				
Mean ± SD	0.03 ± 0.05		0.02 ± 0.09	
Interocular acuity difference (logMAR)				
Mean ± SD	0.37 ± 0.19		0.37 ± 0.17	
Baseline stereoacuity (arcseconds)				
Nil (converted to 4,500)	4	27%	6	40%
800	0	0%	1	7%
400	3	20%	2	13%
200	2	13%	0	0%
140	0	0%	2	13%
100	1	7%	0	0%
80	0	0%	3	20%
60	1	7%	1	7%
50	2	13%	0	0%
40	2	13%	1	7%
Mean ± SD (arcseconds)	1,329.33 ± 1,983.41		1,826.25 ± 2,147.26	
Mean ± SD (log arcseconds)	2.49 ± 0.81		2.71 ± 0.82	
Amblyopic eye's spherical equivalent (diopters)				
Mean ± SD	3.95 ± 2.58		4.05 ± 1.76	
Fellow eye's spherical equivalent (diopters)				
Mean ± SD	0.85 ± 1.50		0.77 ± 1.41	
Squint (Δ) by PACT				
0	8	53%	13	81%
1 to < 10	6	40%	12	13%
10 to < 20	1	7%	1	6%
Mean ± SD (Δ)	−2.40 ± 3.87		−1.63 ± 4.21	

logMAR, logarithm of the minimum angle of resolution; SD, standard deviation; PACT, Prism and Alternate Cover Test.

(2) history of interocular or refractive surgery that affects vision; (3) history of treatment for amblyopia in the last 3 months before screening (except spectacle frames); (4) photosensitivity epilepsy; (5) confirmed or suspected conjunctivitis; (6) allergy or intolerance to the test equipment or patch; (7) history of pharmacological intervention that may affect vision such as atropine; (8) participation in another clinical study/trial within 1 month before the enrollment of the study; and (9) failure to comply with optical adaptation.

Treatment Procedure

After confirming for the eligibility of each patient, we randomly assigned all participants to one of the two treatment groups using a random number generator so that the group assignment was balanced for both groups (see **Figure 1**). For all of the patients in Flicker Group and Patching Group, the flicker and patching treatment were combined with optimal refractive correction. As for tracking the compliance, the EFG automatically recorded the duration and time in which the patients wore them. A child has to wear the glasses for 1 h to be counted as complying for the day. On the other hand, we relied on parental diaries, which are not as objective, for measuring daily compliance in the Patching Group. Throughout the study, we asked parents to report for any signs of adverse reactions, such as nausea and double vision, to patching or EFG to our clinic. However, we did not receive any reports of such side-effects.

During the period of wearing the EFG or the patch, patients performed daily activities such as doing homework or watching television. During their first visit (Day 1), guardians of the patients within the Flicker Group were informed about the specifics of the EFG device, such as recharging and handling. Follow-up visits were scheduled after 3, 6, and 12 weeks from their first visit. During each visit, monocular and binocular visual functions were assessed. We had decided to follow up patients for 12 weeks because the investigators who first reported the benefit of the EFG therapy (Vera-Diaz et al., 2016) also followed up their patients for 12 weeks.

Flicker Group (Eyetronix Flicker Glass Therapy)

Schor et al. (1976) employed a rate for alternate flicker at 7 Hz, and Vera-Diaz et al. (2016) recently observed a gain in VA of amblyopes with 12 weeks of 7-Hz alternative flicker treatment. Similar to Vera-Diaz et al. (2016), we set the flicker rate of EFG at 7 Hz; the lenses were alternatively transparent ("on") and opaque ("off") state for 0.071 + seconds. Also, we set the duty cycle at 50% (50:50 ratio); EFG deprived the normal eye 50% of the time throughout the treatment. Therefore, 1-h treatment with the EFG was equivalent to 30-min deprivation of said eye (50%). The patients were asked to wear the EFG for 1 h/day, 7 days/week throughout the 12-week treatment period.

Patching Group (Control)

Patients were instructed to wear a standard, latex-free and adhesive style patch in front of their normal eyes for 2 h, 7 days/week throughout the 12-week treatment period.

Visual Acuity

Best-corrected visual acuity was measured separately for each eye using a Tumbling E Logarithmic Visual Acuity Chart (xk100-06, China), which follows the protocol of Early Treatment Diabetic

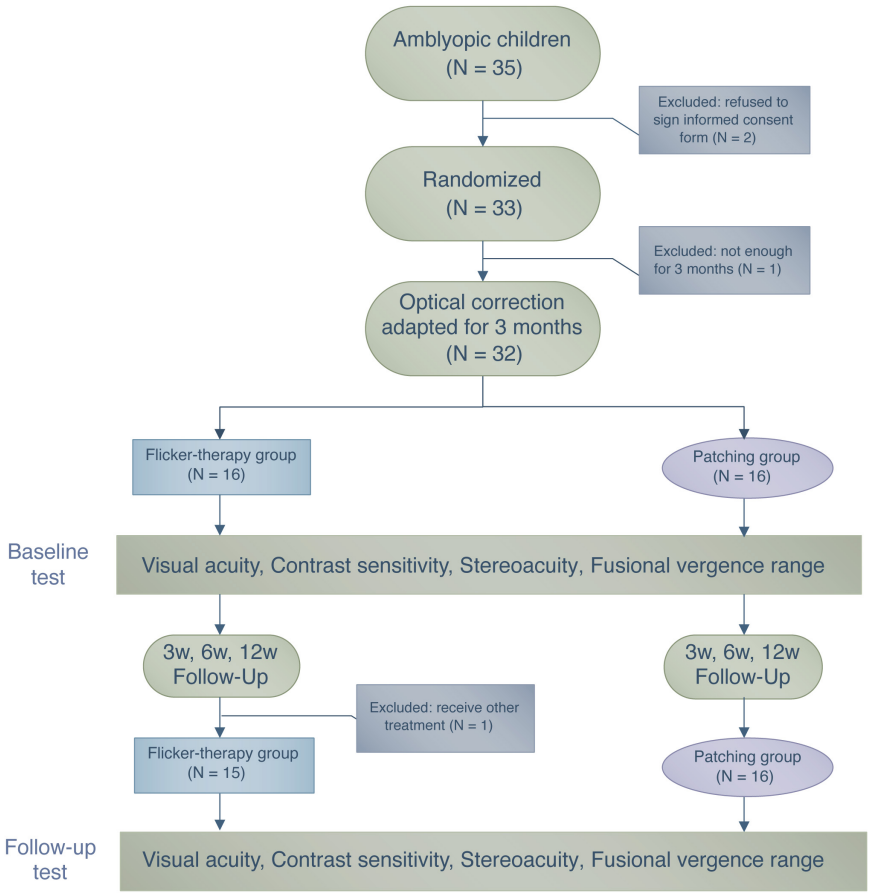


FIGURE 1 | A flowchart illustrating the treatment procedure and the number of patients who participated in this study.

Retinopathy Study. The total score within each line from the Logarithmic Visual Acuity Chart was 0.1 log units. Since there were five letters per line, the correctly read letter was assigned a score of 0.02 log units. The patients were asked to report the orientation of the E letter tested with one eye at 5 m from the chart, and with the non-tested eye occluded throughout the test.

Contrast Sensitivity

A CSV-1000 grating chart (VectorVision® Inc., Greenville, OH 45331, United States) is a printed chart-based method to test contrast sensitivity. It presents sinewave targets at four spatial frequencies [3, 6, 12, and 18 cycles per degree (cpd)] at a distance of 40 cm. Subjects were asked to perform a forced-choice task between two targets at different rows: one had a sinusoidal modulation, whereas the other was a mean gray. There are eight levels of contrast (one per column) per spatial frequency. The lowest contrast in the chart is 0.5%. The contrast of patches is presented in decreasing order from left to right. Using the table in the company's website[2], we translated the scores (range of 1 to 8) to log units for analysis.

[2]http://www.vectorvision.com/csv1000-norms/

Stereopsis

Stereo threshold was measured with a Titmus Stereo test (Stereo Optical Co., Chicago, IL, United States) at a viewing distance of 40 cm under natural light. All patients wore polarizing glasses and, if necessary, additional optical corrections throughout testing. If the patients were not able to perceive the largest disparity given by the Titmus Stereo test, we recorded that the stereoacuity of the patients was "nil" and converted the non-numerical value into 4,500 arcseconds for data analysis.

Fusional Vergence Range

Fusional convergence and divergence amplitudes were measured using Synoptophore L-2510B/L-2510HB (Inami & Co., Ltd., Japan). We measured the extent to which the subjects maintained binocular single vision as we increased the vergence demands. Then we computed the total range of fusion, which is the absolute sum of the convergence and divergence break points. The data were recorded in the unit of degrees (deg).

Data Analysis

Statistical analyses were performed using R software (RStudio, 2016). Data were plotted using Python software (Hunter, 2007). A Shapiro–Wilk test indicated that our dataset of BCVA, contrast

sensitivity, and stereopsis did not assume a normal distribution ($p < 0.05$) and failed to meet the requirement for the use of parametric procedures. Therefore, we performed non-parametric (rank-based) analysis of variance (ANOVA)-like computation of longitudinal data using the package "nparLD" designed for R software (Noguchi et al., 2012). *Post hoc* tests were performed with Bonferroni correction when there was a significant main effect of either treatment group or time. We report ANOVA-type statistics (Brunner et al., 2017); 95% confidence intervals (CIs) were obtained from t-distribution approximations. In addition, dataset of fusional vergence range was normally distributed. So we used a parametric two-way mixed ANOVA (within-subject: time, between-subject: group) to compute statistics.

RESULTS

Baseline Characteristics

We wanted to ensure that all patients were randomly assigned to the two treatment groups. So we looked for possible differences in baseline between the Flickering and Patching groups via a non-parametric Wilcoxon–Mann–Whitney test. We found that the baseline characteristics, including age, BCVA, contrast sensitivity, stereopsis, fusional vergence range, amblyopic eye's spherical equivalent, fellow eye's spherical equivalent, and squint, did not differ between the treatment groups (p-values ranged from 0.17 to 0.96) and concluded that the group assignment was random.

Compliance

We computed compliance by dividing the number of actual days where the patients had undergone the treatment at home from the number of total days for treatment. The mean compliance rate in the Flicker Group was 93.80 ± 0.025%, whereas the mean compliance rate in the Patching Group was 93.97 ± 0.021%. A non-parametric Wilcoxon–Mann–Whitney test revealed no significant difference in compliance rate between the two treatment groups ($W = 116$, $p = 0.89$).

Comparison Between the Eyetronix Flicker Glass Therapy and the Standard Patching Therapy

Improvement of Best-Corrected Visual Acuity (Primary Outcome)

We firstly conducted a non-parametric mixed ANOVA-like test (see *Materials and Methods*), which includes between-subject factor (treatment group) and within-subject factor (time). It revealed no main effect of group ($F_{(1.00,\infty)} = 1.57$, $p = 0.21$) but a significant effect of time ($F_{(2.37,\infty)} = 40.42$, $p < 0.001$). Then we performed a *post hoc* analysis for each pairwise comparison (between two timepoints per group) with Bonferroni correction. We found a significant difference between baseline and 12 weeks in Flicker Group ($p < 0.001$, Cohen's $d = 0.86$, 95% CI: [0.016, 0.24] logMAR). We also found a significant difference between baseline and 12 weeks in the Patching Group ($p < 0.001$, Cohen's $d = 1.08$, 95% CI: [0.068, 0.34] logMAR). These results show that both groups improved BCVA significantly over 12 weeks. However, according to the ANOVA test, there was no significant difference in VA improvement at 12 weeks relative to baseline between the two groups ($F_{(1.00,\infty)} = 2.26$, $p = 0.13$). In particular, the mean improvement (mean ± SD) in BCVA at 12 weeks relative to baseline in the Patching Group (0.21 ± 0.14 logMAR) was slightly (but not significantly) higher than in the Flicker Group (0.13 ± 0.11 logMAR).

Also, we found a significant interaction between the treatment group and time ($F_{(2.29,\infty)} = 4.39$, $p = 0.0091$). This indicates that the rate of BCVA improvement was significantly faster in the Patching Group than in the Flicker Group. **Figure 2B** shows that most points are under the dashed line (unity line); this indicates that most children in both groups experienced an increase in BCVA after 12 weeks of treatment.

Interestingly, we observed a significant negative correlation between baseline VA and VA improvement at 12 weeks in the Flicker Group ($r = -0.56$, $p = 0.029$). This indicates that the worse the baseline VA, the larger the VA improvement at 12 weeks. However, we did not find a significant correlation in the Patching Group. We also examined the relationship between age and VA improvement at 12 weeks. According to a Pearson correlation test, we found no significant relationship in both Flicker (rho = −0.20, $p = 0.48$) and Patching (rho = 0.14, $p = 0.60$) groups.

Contrast Sensitivity

Using the non-parametric ANOVA test (within-subject factor: time, between-subject factor: group), we found no significance difference between Flicker and Patching groups (p's < 0.05; see **Figure 3**). However, we found a significant effect of time at all spatial frequencies (p's < 0.001). So we compared baseline and 12 weeks as *post hoc* analysis at each spatial frequency and treatment group. In the Flicker Group, we observed a significance improvement (p's < 0.05) at 12-week relative to baseline for contrast sensitivity at 3 and 12 cpd but not at 6 and 18 cpd. The mean improvements (mean ± SD) for 3, 6, 12, and 18 cpd between baseline and 12 weeks were 0.14 ± 0.271, 0.23 ± 0.43, 0.28 ± 0.40, and 0.13 ± 0.52 respectively. For the Patching Group, there was a significance improvement (p's < 0.05) at 12-week relative to baseline for 3, 6, and 12 cpd but not at 18 cpd. This means that contrast sensitivity for 3, 6, and 12 cpd improved significantly throughout the treatment. The mean improvements (mean ± SD) for 3, 6, 12, and 18 cpd between baseline and 12 weeks were 0.28 ± 0.46, 0.31 ± 0.33, 0.30 ± 0.44, and 0.25 ± 0.47, respectively.

For all spatial frequencies, we also evaluated the relationship between age and improvement in contrast sensitivity using a Pearson correlation test. For the Flicker Group, there was no correlation between age and improvement (p's > 0.05). For the Patching Group, there was no significant correlation (p's > 0.05) except at 3 cpd (rho = 0.58, $p = 0.018$).

Stereopsis

Non-parametric mixed ANOVA-like test revealed no main effect of group ($F_{(1.00,\infty)} = 0.52$, $p = 0.47$) and no significant effect of time ($F_{(2.46,\infty)} = 2.18$, $p = 0.10$). Also, it revealed no interaction effect between the treatment group and time ($F_{(2.46,\infty)} = 0.56$,

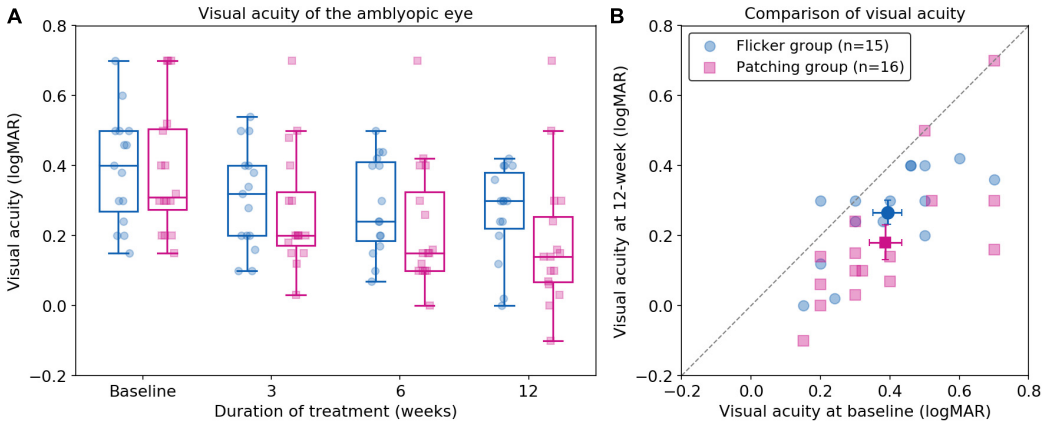


FIGURE 2 | Best-corrected visual acuity following treatment. **(A)** Boxplots of best correct visual acuity (BCVA) of the amblyopic eye (in logMAR) in Flicker and Patching treatment groups. Blue represents the Flicker Group, and pink the Patching Group. Individual data point is represented by a dot. The black solid line within each box represents the median. The box represents the interquartile range (IQR) of the data (25th to 75th percentile). The whisker represents 1.5 × IQR either above the third quartile or below the first quartile. **(B)** Individual data point of visual acuity (BCVA) in patients' amblyopic eyes measured at baseline and at 12 weeks of treatment. The dashed line represents unity. Blue circles represent the Flicker Group, whereas pink triangles represent the Patching Group. Points below the dashed line (unity line) show an improved BCVA at week 12 relative to baseline.

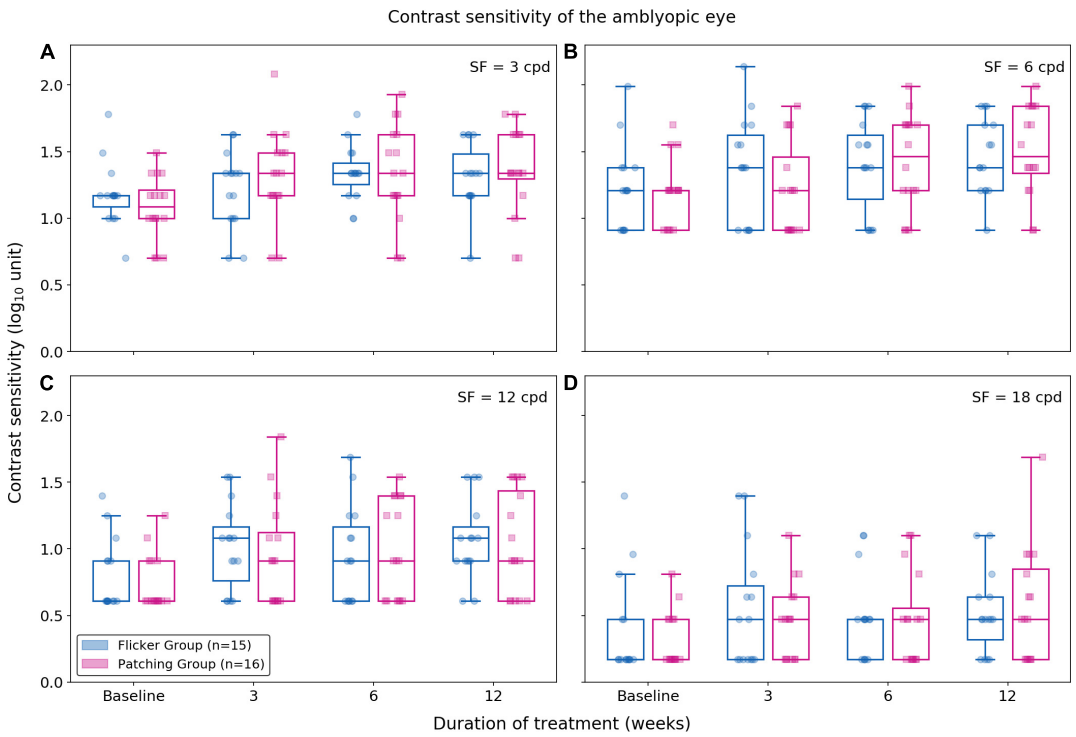


FIGURE 3 | Contrast sensitivity at 3, 6, 12, and 18 cpd (cycles per degree) for each treatment group. Blue plots represent the Flicker Group, and pink plots the Patching Group. Circles represent the contrast sensitivity of patients in the Flicker Group, and squares the contrast sensitivity patients in the Patching Group. **(A)** Contrast sensitivity for 3 cpd. **(B)** Contrast sensitivity for 6 cpd. **(C)** Contrast sensitivity for 12 cpd. **(D)** Contrast sensitivity for 18 cpd.

$p = 0.61$). For the Flicker Group, the mean improvement in stereo threshold was 0.573 ± 1.62 $\log_{10}$ arcsecs (mean ± SD) at 12 weeks. For the Patching Group, the mean improvement in stereo threshold was 0.662 ± 1.93 $\log_{10}$ arcsecs (mean ± SD) at 12 weeks. Both improvements were not significant (p's > 0.05) relative to baseline (see **Figure 4**).

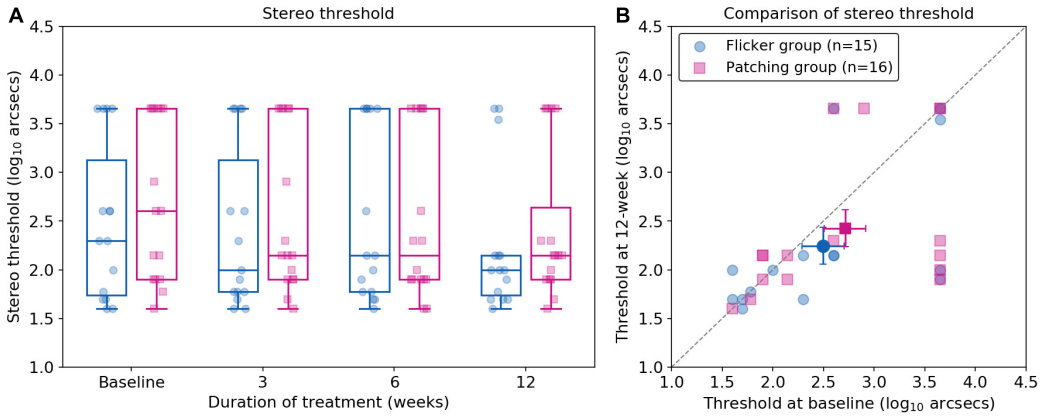


FIGURE 4 | Stereo thresholds (log arcsecs) as measured with the Titmus Stereo test. **(A)** Boxplots of stereo thresholds in Flicker and Patching treatment groups. **(B)** Individual data point of stereo thresholds measured at baseline and at 12 weeks of treatment. Both panels are plotted similarly to **Figure 2**.

Fusional Vergence Range

Five patients (two in the Flicker Group and three in the Patching Group) were not able to complete the test of fusional vergence range and therefore had to be excluded in data analysis (see **Figure 5**). A two-way mixed ANOVA showed no significant effect of group ($F_{(1,24)} = 0.24, p = 0.63$) and no significant effect of time ($F_{(3,72)} = 2.19$, $p = 0.096$). It revealed no significant interaction effect between the treatment group and time ($F_{(3,72)} = 0.77$, $p = 0.34$). In short, for both groups, no significant difference was observed between baseline and 12 weeks. However, the mean changes in fusional vergence range for both groups slightly deteriorated. In the Flicker Group, the mean change (mean ± SD) was −1.85 ± 5.63 deg, whereas in the Patching Group, the mean change was −0.692 ± 2.75 deg.

DISCUSSION

Our results showed that the EFG therapy is not superior to patching for amblyopia in children. The gain in VA between baseline and 12 weeks was slightly larger in patching (0.21 ± 0.14 logMAR) than EFG therapy (0.13 ± 0.11 logMAR). The VA gain in EFG therapy is similar to what a previous EFG study reports at 12 weeks (0.12 ± 0.11 logMAR) (Vera-Diaz et al., 2016). However, the difference in VA gain between the EFG and patching therapies was not statistically significant. In addition, both therapies induced a significant improvement in VA at 12 weeks relative to baseline. Also, both treatment groups showed a gain in contrast sensitivity at 3 and 12 cpd; however, only patching improved the contrast sensitivity at 6 cpd. Stereo threshold and fusional vergence range did not significantly improve in both groups.

The gain in VA from 12-week EFG therapy (0.13 logMAR) reported in our study is similar to that from binocular therapies. For instance, a pediatric study on dichoptic perceptual learning in 14 amblyopic children shows a VA gain of 0.1 logMAR (Knox et al., 2012). However, the dichoptic therapy seems to be more effective than the EFG therapy because the children completed five training sessions (1 h each) total in 1 week rather than undergoing a daily 1-h treatment for 12 weeks (i.e., our EFG therapy). Moreover, studies show that a binocular iPad therapy improves VA by 0.08–0.09 logMAR at 4 weeks (Li et al., 2014; Birch et al., 2015; Birch et al., 2020) and 0.105 logMAR at 16 weeks (Holmes et al., 2016). These improvements are similar to what we found at 3 (0.085 logMAR) and 12 weeks (0.13 logMAR) of EFG treatment. However, other RCTs show that a binocular iPad therapy is more effective than an EFG therapy. They report a VA gain of 0.15 logMAR after 2 to 8 weeks of a binocular iPad therapy (Li et al., 2015a; Kelly et al., 2016). An EFG therapy seems to be superior I-BiT therapy, which preferentially stimulates the amblyopic eye without depriving the input to the fellow eye. To illustrate, an RCT shows that 3 h of weekly treatment using I-BiT reports a gain of 0.07 logMAR in the amblyopic eye at 6 weeks (Foss et al., 2013), whereas we report a gain of 0.11 logMAR at 6 weeks. To evaluate whether an EFG therapy stands in all of these treatment options, a future RCT should test a larger pool of patients and compare the therapies.

We found that the rate of VA improvement in the Patching Group was significantly faster than in the Flicker Group. It seems that the VA gain did not plateau earlier in the Patching Group, since the improvement relative to baseline was larger at 12 weeks (0.21 ± 0.14 logMAR) than 6 weeks (0.16 ± 0.10 logMAR). Therefore, it seems that the VA improvements in EFG therapy were smaller in comparison with patching. There could be a few possible explanations for this difference. First, patching might simply be slightly superior for improving VA to EFG therapy. To illustrate, the age distribution of the patients for the Patching Group (6.38 ± 2.45 years, mean ± SD) was slightly (but not significantly) different than that in the Flicker Group (5.27 ± 1.10 years, mean ± SD). The former group had four more children in the age between 8 and 13 years. Studies have shown that there is an age-dependent response to patching in favor of younger patients (Fronius et al., 2014). Therefore, the rate of VA improvement from patching would have been even faster if the age distribution between the two groups was more similar. Second, the treatment durations of the two therapies differed.

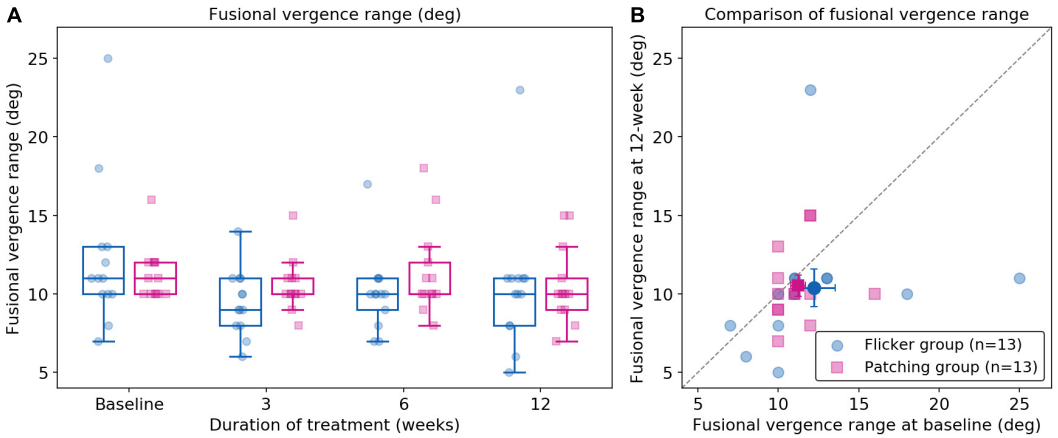


FIGURE 5 | Fusional vergence range (deg). **(A)** Fusional vergence range in Flicker and Patching treatment groups. **(B)** Individual data point of fusional vergence range measured at baseline and at 12 weeks of treatment. Both panels are plotted similarly to **Figure 2**.

We acknowledge that the shorter duration of EFG therapy (i.e., 1 h/day) might have contributed to the slower rate of VA gain than that in patching (2 h/day). To our limited knowledge, there is no direct evidence to show that a longer daily duration of EFG treatment produces more VA gain. However, a patching therapy study reports that dose rates of 2 to 6 h/day produced the same visual outcome, although the rate of visual gain was quicker in a larger dose of patching treatment (Stewart et al., 2004). We had decided to administer EFG therapy for 1 h/day, rather than 2 h/day, for two reasons: 1) Eyetronix® claims that daily dose of 1–2 h is sufficient[3], and (2) the battery life of our EFG lasted for only 1.5 h. Despite the difference in treatment duration between the two groups, we did not find a significant difference in VA improvement at 12 weeks.

Eyetronix® categorizes the EFG therapy as binocular because a dichoptic flicker has been shown to affect interocular interaction (Schor et al., 1976). To illustrate, using a binocular VA test, Schor et al. (1976) showed that the influence of the fellow eye on amblyopic eye's perception was minimized while visual targets were alternatively presented to the eyes at 7 Hz (i.e., dichoptic flicker) (Schor et al., 1976). Another study shows that binocular (but not dichoptic) flicker can affect binocular interaction (Kosovicheva et al., 2019). It claims that EFG can "break suppression and restore normal binocular fusion" (See Text Footnote 3). However, the categorization is open to debate. Unlike EFG, which deprives monocular input at all times, binocular therapies show stimuli to both eyes simultaneously. For example, a binocular therapy, be it movie viewing (Li et al., 2015b) or game playing (Knox et al., 2012), displays stimuli to both eyes simultaneously. Moreover, we do not see an improvement in stereopsis and fusional vergence range as some studies of binocular therapies have shown. For instance, dichoptic anti-suppression therapies, which reduce the suppression by lowering the contrast for the fellow eye, have been reported to bring binocular benefits (Hess et al., 2010; Li et al., 2013; Vedamurthy et al., 2015). Moreover, a laboratory study reports that a dichoptic virtual reality display, which does not reduce the contrast for the fellow eye, reduces suppression in normal adults (Bao et al., 2017). This dichoptic design is analogous to the EFG therapy, which deprives each eye equally without rebalancing the contrast. It is very likely that these different designs of binocular treatment could be accounted for by different mechanisms.

It should be noted that the reported interocular suppression changes were measured while subjects were viewing with dichoptic flicker (Schor et al., 1976). Would there be the same effect when patients are no longer viewing with the flicker? We do not think so because we did not observe binocular benefits after EFG treatment. However, our results disagree with a previous report from Eyetronix® (Vera-Diaz et al., 2016), which shows an improvement in stereopsis. This might be accounted for by the different doses of treatment in each day (2 vs. 1 h) or different stereo measures (Random Dot 2 test vs. Titmus). Nevertheless, one would wonder why we had not measured suppression in our RCT. We had not done so because children have found psychophysical tasks (i.e., dichoptic motion coherence task) that measure suppression challenging (Li et al., 2014; Li et al., 2015b). We are currently working on a study to investigate the effect of the EFG therapy on suppression in adults.

Binocular therapies of amblyopia, such as dichoptic training, perceptual learning, and video gaming, show promise for treating adults (Li et al., 2013; Hess and Thompson, 2015; Vedamurthy et al., 2016). On the other hand, monocular therapies except refractive therapy (Gao et al., 2018) show that younger patients benefit more than older ones (Stewart et al., 2007; Fronius et al., 2014). Therefore, whether EFG therapy, which is effective in improving amblyopic children's VA, is also effective for amblyopic adults should also be explored in the future.

Limitations

Our study has several limitations. First, we did not have a large sample size. Second, the durations of the two therapies were

[3]http://eyetronix.com/

different. Third, we did not perform follow-up assessments of the patients after the end of the treatment. Therefore, whether the monocular benefits from the EFG therapy last remains to be investigated.

CONCLUSION

An EFG therapy and patching brought a similar VA gain for children with amblyopia at 12-week visit. Both therapies resulted in a gain in contrast sensitivity at 3 and 12 cpd; however, only patching improved the contrast sensitivity at 6 cpd. Both therapies did not benefit binocular visual functions (stereopsis and fusional vergence range). We believe that EFG can be an additional choice for therapy.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Affiliated Eye Hospital at Wenzhou Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

XY, JX, YW, and JZ conceived the experiments. BC and HC performed the experiments. SM, SC, and JZ analyzed and interpreted the data, and wrote the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

This research was supported by the National Key Research and Development Program of China Grant 2020YFC2003800, the National Natural Science Foundation of China Grant NSFC 31970975, and the Qianjiang Talent Project (QJD1702021) to JZ and Canadian Institute of Health Research's graduate award to SM. The preprint version is listed in medRxiv (doi: https://doi.org/10.1101/2020.07.20.20157552).

REFERENCES

Bao, M., Dong, B., Liu, L., Engel, S. A., and Jiang, Y. (2017). The best of both worlds: adaptation during natural tasks produces long-lasting plasticity in perceptual ocular dominance. *Psychol. Sci.* 29, 14–33. doi: 10.1177/0956797617728126

Birch, E. E., Jost, R. M., De La Cruz, A., Kelly, K. R., Beauchamp, C. L., Dao, L., et al. (2019). Binocular amblyopia treatment with contrast-rebalanced movies. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* 23, 160.e1–160.e5. doi: 10.1016/j.jaapos.2019.02.007

Birch, E. E., Jost, R. M., Wang, S. X., and Kelly, K. R. (2020). A pilot randomized trial of contrast-rebalanced binocular treatment for deprivation amblyopia. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* S1091-8531:30217-2. doi: 10.1016/j.jaapos.2020.07.009

Birch, E. E., Li, S. L., Jost, R. M., Morale, S. E., De La Cruz, A., Stager, D. Jr., et al. (2015). Binocular iPad treatment for amblyopia in preschool children. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* 19, 6–11. doi: 10.1016/j.jaapos.2014.09.009

Brunner, E., Konietschke, F., Pauly, M., and Puri, M. L. (2017). Rank-based procedures in factorial designs: hypotheses about non-parametric treatment effects. *J. R. Stat. Soc. Ser. B (Statist. Methodol.)* 79, 1463–1485.

De Buffon, C. (1743). *Disseration Sur la Cause du Strabisme ou des Yeux Louches*. Paris: Medical Acadamy Royal Science.

Foss, A. J., Gregson, R. M., MacKeith, D., Herbison, N., Ash, I. M., Cobb, S. V., et al. (2013). Evaluation and development of a novel binocular treatment (I-BiT^{TM}) system using video clips and interactive games to improve vision in children with amblyopia ('lazy eye'): study protocol for a randomised controlled trial. *Trials* 14:145. doi: 10.1186/1745-6215-14-145

Fronius, M., Cirina, L., Ackermann, H., Kohnen, T., and Diehl, C. M. (2014). Efficiency of electronically monitored amblyopia treatment between 5 and 16 years of age: new insight into declining susceptibility of the visual system. *Vis. Res.* 103, 11–19. doi: 10.1016/j.visres.2014.07.018

Gao, T. Y., Anstice, N., Babu, R. J., Black, J. M., Bobier, W. R., Dai, S., et al. (2018). Optical treatment of amblyopia in older children and adults is essential prior to enrolment in a clinical trial. *Ophthalmic. Physiol. Opt.* 38, 129–143. doi: 10.1111/opo.12437

Hess, R. F., Mansouri, B., and Thompson, B. (2010). A binocular approach to treating amblyopia: antisuppression therapy. *Optometr. Vision Sci.* 87, 697–704.

Hess, R. F., and Thompson, B. (2013). New insights into amblyopia: binocular therapy and noninvasive brain stimulation. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* 17, 89–93. doi: 10.1016/j.jaapos.2012.10.018

Hess, R. F., and Thompson, B. (2015). Amblyopia and the binocular approach to its therapy. *Vis. Res.* 114, 4–16. doi: 10.1016/j.visres.2015.02.009

Hickey, G. L., Grant, S. W., Dunning, J., and Siepe, M. (2018). Statistical primer: sample size and power calculations-why, when and how? *Eur. J. Cardiothorac. Surg.* 54, 4–9. doi: 10.1093/ejcts/ezy169

Holmes, J. M., Manh, V. M., Lazar, E. L., Beck, R. W., Birch, E. E., Kraker, R. T., et al. (2016). Effect of a binocular iPad game vs Part-time patching in children aged 5 to 12 years with amblyopia: a randomized clinical trial. *JAMA Ophthalmol.* 134, 1391–1400. doi: 10.1001/jamaophthalmol.2016.4262

Holmes, J. M., Manny, R. E., Lazar, E. L., Birch, E. E., Kelly, K. R., Summers, A. I., et al. (2019). A randomized trial of binocular dig rush game treatment for amblyopia in children aged 7 to 12 years. *Ophthalmology* 126, 456–466. doi: 10.1016/j.ophtha.2018.10.032

Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95.

Kelly, K. R., Jost, R. M., Dao, L., Beauchamp, C. L., Leffler, J. N., and Birch, E. E. (2016). Binocular iPad game vs patching for treatment of amblyopia in children: a randomized clinical trial. *JAMA Ophthalmol.* 134, 1402–1408. doi: 10.1001/jamaophthalmol.2016.4224

Knox, P. J., Simmers, A. J., Gray, L. S., and Cleary, M. (2012). An exploratory study: prolonged periods of binocular stimulation can provide an effective treatment for childhood amblyopia. *Invest. Ophthalmol. Visi. Sci.* 53, 817–824. doi: 10.1167/iovs.11-8219

Kosovicheva, A., Ferreira, A., Vera-Diaz, F. A., and Bex, P. J. (2019). Effects of temporal frequency on binocular deficits in amblyopia. *Vis. Res.* 163, 52–62. doi: 10.1016/j.visres.2019.08.004

Levi, D. M., Knill, D. C., and Bavelier, D. (2015). Stereopsis and amblyopia: a mini-review. *Vis. Res.* 114, 17–30. doi: 10.1016/j.visres.2015.01.002

Li, J., Thompson, B., Deng, D., Chan, L. Y. L., Yu, M., and Hess, R. F. (2013). Dichoptic training enables the adult amblyopic

brain to learn. *Curr. Biol.* 23, R308–R309. doi: 10.1016/j.cub.2013.01.059

Li, S. L., Jost, R. M., Morale, S. E., De La Cruz, A., Dao, L., Stager, D. Jr., et al. (2015a). Binocular iPad treatment of amblyopia for lasting improvement of visual acuity. *JAMA Ophthalmol.* 133, 479–480. doi: 10.1001/jamaophthalmol.2014.5515

Li, S. L., Jost, R. M., Morale, S. E., Stager, D. R., Dao, L., Stager, D., et al. (2014). A binocular iPad treatment for amblyopic children. *Eye (Lond.)* 28, 1246–1253. doi: 10.1038/eye.2014.165

Li, S. L., Reynaud, A., Hess, R. F., Wang, Y.-Z., Jost, R. M., Morale, S. E., et al. (2015b). Dichoptic movie viewing treats childhood amblyopia. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* 19, 401–405. doi: 10.1016/j.jaapos.2015.08.003

Manh, V. M., Holmes, J. M., Lazar, E. L., Kraker, R. T., Wallace, D. K., Kulp, M. T., et al. (2018). A randomized trial of a binocular iPad game versus part-time patching in children aged 13 to 16 years with amblyopia. *Am. J. Ophthalmol.* 186, 104–115. doi: 10.1016/j.ajo.2017.11.017

Noguchi, K., Gel, Y. R., Brunner, E., and Konietschke, F. (2012). nparLD: an R software package for the nonparametric analysis of longitudinal data in factorial experiments. *J. Stat. Softw.* 50, 1–23.

Papageorgiou, E., Asproudis, I., Maconachie, G., Tsironi, E. E., and Gottlob, I. (2019). The treatment of amblyopia: current practice and emerging trends. *Graefes Arch. Clin. Exp. Ophthalmol.* 257, 1061–1078. doi: 10.1007/s00417-019-04254-w

Rajavi, Z., Sabbaghi, H., Amini Sharifi, E., Behradfar, N., and Kheiri, B. (2019). Comparison between patching and interactive binocular treatment in amblyopia: a randomized clinical trial. *J. Curr. Ophthalmol.* 31, 426–431. doi: 10.1016/j.joco.2019.07.004

Repka, M. X., Beck, R. W., Holmes, J. M., Birch, E. E., Chandler, D. L., Cotter, S. A., et al. (2003). A randomized trial of patching regimens for treatment of moderate amblyopia in children. *Arch. Ophthalmol.* 121, 603–611. doi: 10.1001/archopht.121.5.603

RStudio (2016). *RStudio: Integrated Development for R*. Boston, MA: RStudio.

Rutstein, R. P., Quinn, G. E., Lazar, E. L., Beck, R. W., Bonsall, D. J., Cotter, S. A., et al. (2010). A randomized trial comparing Bangerter filters and patching for the treatment of moderate amblyopia in children. *Ophthalmology* 117, 998–1004.e6. doi: 10.1016/j.ophtha.2009.10.014

Scheiman, M. M., Hertle, R. W., Beck, R. W., Edwards, A. R., Birch, E., Cotter, S. A., et al. (2005). Randomized trial of treatment of amblyopia in children aged 7 to 17 years. *Arch. Ophthalmol.* 123, 437–447. doi: 10.1001/archopht.123.4.437

Schor, C., Terrell, M., and Peterson, D. (1976). Contour interaction and temporal masking in strabismus and amblyopia. *Am. J. Optom. Physiol. Opt.* 53, 217–223. doi: 10.1097/00006324-197605000-00001

Searle, A., Norman, P., Harrad, R., and Vedhara, K. (2002). Psychosocial and clinical determinants of compliance with occlusion therapy for amblyopic children. *Eye (Lond.)* 16, 150–155. doi: 10.1038/sj.eye.6700086

Spierer, A., Raz, J., Benezra, O., Herzog, R., Cohen, E., Karshai, I., et al. (2010). Treating amblyopia with liquid crystal glasses: a pilot study. *Inves. Ophthalmol. Vis. Sci.* 51, 3395–3398. doi: 10.1167/iovs.09-4568

Stewart, C. E., Moseley, M. J., Stephens, D. A., and Fielder, A. R. (2004). Treatment dose-response in amblyopia therapy: the Monitored occlusion treatment of amblyopia study (MOTAS). *Invest. Ophthalmol. Vis. Sci.* 45, 3048–3054. doi: 10.1167/iovs.04-0250

Stewart, C. E., Stephens, D. A., Fielder, A. R., and Moseley, M. J. (2007). Objectively monitored patching regimens for treatment of amblyopia: randomised trial. *BMJ* 335:707. doi: 10.1136/bmj.39301.460150.55

Vedamurthy, I., Knill, D. C., Huang, S. J., Yung, A., Ding, J., Kwon, O.-S., et al. (2016). Recovering stereo vision by squashing virtual bugs in a virtual reality environment. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150264. doi: 10.1098/rstb.2015.0264

Vedamurthy, I., Nahum, M., Huang, S. J., Zheng, F., Bayliss, J., Bavelier, D., et al. (2015). A dichoptic custom-made action video game as a treatment for adult amblyopia. *Vis. Res.* 114, 173–187. doi: 10.1016/j.visres.2015.04.008

Vera-Diaz, F. A., Moore, B., Hussey, E., Srinivasan, G., and Johnson, C. (2016). A flicker therapy for the treatment of amblyopia. *Vis. Dev. Rehabil.* 2, 105–114.

Vera-Diaz, F. A., Srinivasan, G., Johnson, C., Hussey, E., Spivey, D., Gleason, W., et al. (2014). Initial evaluation of the novel EFG therapy for the treatment of amblyopia. *Invest. Ophthalmol. Vis. Sci.* 55, 815–815.

Wallace, D. K., Repka, M. X., Lee, K. A., Melia, M., Christiansen, S. P., Morse, C. L., et al. (2018). Amblyopia preferred practice pattern®. *Ophthalmology* 125, 105–142. doi: 10.1016/j.ophtha.2017.10.008

Wallace, M. P., Stewart, C. E., Moseley, M. J., Stephens, D. A., and Fielder, A. R. (2013). Compliance with occlusion therapy for childhood amblyopiacompliance with occlusion therapy. *Invest. Ophthalmol. Vis. Sci.* 54, 6158–6166. doi: 10.1167/iovs.13-11861

Wang, J., Neely, D. E., Galli, J., Schliesser, J., Graves, A., Damarjian, T. G., et al. (2016). A pilot randomized clinical trial of intermittent occlusion therapy liquid crystal glasses versus traditional patching for treatment of moderate unilateral amblyopia. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* 20, 326–331. doi: 10.1016/j.jaapos.2016.05.014

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 13 May 2021
doi: 10.3389/fnins.2021.648648

Do Impairments in Visual Functions Affect Skiing Performance?

*Amritha Stalin, Marieke Creese and Kristine Nicole Dalton**

School of Optometry & Vision Science, University of Waterloo, Waterloo, ON, Canada

Nordic and alpine skiing-related visual tasks such as identifying hill contours, slope characteristics, and snow conditions increase demands on contrast processing and other visual functions. Prospective observational studies were conducted to assess the relationships between skiing performance and a broad range of visual functions in nordic and alpine skiers with vision impairments. The study hypothesized that contrast sensitivity (CS), visual acuity (VA), and visual field (VF) would be predictive of skiing performance. Binocular static VA, CS, light sensitivity, glare sensitivity, glare recovery, dynamic VA, translational and radial motion perception, and VF were assessed in elite Para nordic (n = 26) and Para alpine (n = 15) skiers. Skiing performance was assessed based on skiers' raw race times. Performance on the visual function tests was compared with skiing performances using Kendall's correlations (with and without Bonferroni–Holm corrections) and linear multivariable regressions ($p < 0.05$ considered significant). None of the vision variables were significantly correlated with performance in Para nordic or Para alpine skiing after Bonferroni–Holm corrections were applied. Before applying the corrections, VF extent (ρ = -0.37, p = 0.011), and static VA (ρ = 0.26, p = 0.066) demonstrated the strongest correlations with Para nordic skiing performance; in Para alpine skiing, static VA and CS demonstrated the strongest correlations with downhill (static VA: ρ = 0.54, p = 0.046, CS: ρ = -0.50, p = 0.06), super G (static VA: ρ = 0.50, p = 0.007, CS: ρ = -0.51, p = 0.017), and giant slalom (static VA: ρ = 0.57, p = 0.01, CS: ρ = -0.46, p = 0.017) performance. Dynamic VA and VF were significantly associated with downhill (ρ = 0.593, p = 0.04) and slalom (ρ = -0.49, p = 0.013) performances, respectively. Static VA was a significant predictor of giant slalom [(F(3,11) = 24.71, $p < 0.001$), and R of 0.87], super G [(F(3,9) = 17.34, p = 0.002), and R of 0.85], and slalom [(F(3,11) = 11.8, p = 0.002), and R of 0.80] performance, but CS and VF were not. Interestingly, static VA and CS were highly correlated in both Para nordic (ρ = -0.60, $p < 0.001$) and Para alpine (ρ = -0.80, $p < 0.001$) skiers. Of the vision variables, only static VA and VF were associated with skiing performance and should be included as the in Para nordic and Para alpine classifications. The strong correlations between static VA and CS in these skiers with vision impairment may have masked relationships between CS and skiing performance.

Keywords: contrast sensitivity, visual acuity, visual field, Paralympic alpine skiing, Paralympic nordic skiing

OPEN ACCESS

Edited by:
Fang Hou,
Wenzhou Medical University, China

Reviewed by:
Chang-Bing Huang,
Institute of Psychology, Chinese Academy of Sciences, China
Yukai Zhao,
New York University, United States

***Correspondence:**
Kristine Nicole Dalton
kndalton@uwaterloo.ca

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *01 January 2021*
Accepted: *29 March 2021*
Published: *13 May 2021*

Citation:
Stalin A, Creese M and Dalton KN (2021) Do Impairments in Visual Functions Affect Skiing Performance? Front. Neurosci. 15:648648. doi: 10.3389/fnins.2021.648648

INTRODUCTION

Sports and exercise play significant roles in improving the physical and mental health of individuals with visual disabilities (Gleeson et al., 2014; Koolaee, 2017; Fontenot et al., 2018). While the multidisciplinary rehabilitation strategies integrating sports with traditional methods assist at individual levels, international multi-sport events like Paralympics promote positive changes in societal attitudes toward individuals with disabilities and accelerate advancements in accessibility (Gold and Gold, 2007; Ferrara et al., 2015). Although athletes with visual impairment (VI) have been participating in the Paralympics since 1976 (Tweedy et al., 2014), research exploring the impact of vision impairment on sport performance has been limited.

Nordic skiing and alpine skiing are the only two VI sports in the Winter Paralympics (International Paralympic Committee, 2019). Being highly dynamic in nature, these sports demand rapid processing of visual information. Skiers must make quick decisions and vary their speed, direction, or body position based on their visual feedback (Decroix et al., 2017). Research has demonstrated that the major visual cues that elite, able-sighted alpine skiers rely on are the positions of gates on the course, their pole positions, and terrain cues such as the turn initiation and take-off points, the slope and curve of the hill, and distinctive holes and bumps and remarkable transitions on the course. Skiers also reported that the blue-colored markings on the left and right sides of the courses help them to orient, especially in the presence of fog or shadow (Schläppi et al., 2016). Although a similar study has not been done in nordic skiing, it is possible that some of the visual cues nordic skiers use are similar due to similarities in the sport environments.

Identifying the abovementioned visual cues would require skiers to have reasonable distance visual acuity (VA), contrast sensitivity (CS), depth perception, and peripheral vision while they are static, as well as while moving (Craybiel et al., 1955; Senner et al., 1999; Erickson, 2007). Senner et al. (1999) reported that a 20% decrease in static VA could significantly affect the reaction times of leisure skiers to smaller and low-contrast objects such as ice patches, even though their reaction times to larger obstacles such as standing or moving skiers were not affected. Furthermore, skiers' motion and reduced visibility due to extrinsic factors related to weather, lighting, or snow conditions decrease the relative contrast of the visual information, increasing the demand on contrast processing (Erickson, 2007).

Despite similarities in environment, nordic skiing and alpine skiing are different in terms of terrain and skiing techniques. nordic skiing is practiced on flatter terrains with gently rolling undulating hills, and tracks are often narrow and grooved, while alpine terrains are steeper with sharp changes in direction. Therefore, the visual functions and levels of impairment affecting sports performance might differ between nordic and alpine skiing due to the differences in the visual tasks involved in these sports. Comparatively longer nordic skiing courses might require sustained visual and physical performance for a longer time compared to alpine skiing, which is completed in much shorter durations. However, the visual demands during the competitions could be higher in alpine skiing (albeit for short durations) due to the relatively higher speed involved in alpine skiing compared to nordic skiing (Erickson, 2007). Therefore, nordic and alpine skiing should be considered independently while investigating the sports-specific visual functions.

Preliminary studies conducted with Para nordic and Para alpine skiers using a test battery including a broad range of vision assessments such as static and dynamic VA, CS, low-contrast VA, glare sensitivity (GLS), glare recovery (GLR), and color vision reported that none of these visual functions were individually predictive of Para nordic skiing performance, while static VA was a significant predictor of Para alpine slalom performance (Creese et al., 2017a,c). However, measuring CS with the Pelli-Robson chart was not feasible in these populations due to the limitations in letter size as well as the spatial frequency range of the chart (Creese et al., 2017a,c). GLS and GLR could only be measured monocularly during these preliminary studies as well, and participants with a broad range of experience and skill participated, which could have confounded the true vision–performance relationship because of the variable impact of skill on performance (Creese et al., 2017a,b,c). Thus, it was concluded that future studies with refined test batteries are required to identify the visual functions associated with skiing performance (Creese et al., 2017a,b,c).

The purpose of the two independent prospective studies presented in this manuscript was to reexamine the relationship between vision and sport performance in elite, experienced Para nordic and Para alpine skiers of similar skill using a refined vision test battery. The inferences from these studies were also used to identify the visual functions that should be included in the classification systems for Para nordic and Para alpine skiing. It was hypothesized that CS, static VA, and visual field (VF) were associated with skiing performance.

MATERIALS AND METHODS

These studies used an observational research design and adhered with the tenets of the Declaration of Helsinki. All international level Para nordic and Para alpine skiers were given the opportunity to participate, and informed consent was obtained from all participants. This study was reviewed by and received ethics clearance through a University of Waterloo Research Ethics Committee.

Participants

Elite Para nordic and Para alpine skiers were recruited with the help of the International Paralympic Committee (IPC) at the 2017 Para nordic World Championships (WCH, Finsterau, Germany), 2018 Para nordic World Cup (WC, Oberried, Germany), and the 2017 Para Alpine WCH (Tarvisio, Italy). A total of 26 Para nordic skiers (20 from WCH and 6 from WC) and 15 Para alpine skiers participated in the studies. WCH events are the most prestigious Para nordic and Para alpine events next to the Paralympic Games, with only the elite, most competitive skiers as participants. Recruiting from WCH events ensured that the study participants had comparable levels of skiing skill. The six additional Para nordic skiers recruited at the WC event were also

eligible to participate in the WCHs but had not competed at the WCH event in 2017 for political reasons. For reference, there were only 42 Para nordic skiers (WCH, 2017 = 29; WC, 2018 = 13) and 23 Para alpine skiers who were WCH eligible and competing at the events included in this study, and there were only 46 Para nordic skiers and 34 Para alpine skiers in the world who were WCH eligible and registered with the IPC at the time of the study. Therefore, 61.9% of eligible Para nordic skiers and 65.2% of eligible Para alpine skiers participated in these studies, which accounted for 56.5 and 44.1% of the world's entire population of elite Para nordic and Para alpine skiers, respectively. Although small, the study samples were representative of the elite Para nordic and Para alpine skiers' populations.

Procedure

Each participant in these studies attended a single study visit. During the study visit, participants completed a questionnaire (**Supplementary Appendix A**) about their skiing experience, which included questions about their vision impairment and skiing history as well as their current average annual training routine both on- and off-snow. Participants' visual functions were assessed using a test battery that was determined based on the previous feasibility studies conducted by the research team (Creese et al., 2017a). The test battery included binocular tests of static VA, CS, GLS, GLR, light sensitivity (LS), dynamic VA, translational motion perception (TMP), radial motion perception (RMP), and VF. The participants' performance on the visual function assessments was compared with their skiing performance (described below).

Static VA was measured using an Early Treatment Diabetic Retinopathy Study (ETDRS) chart at 1 m and/or the Berkeley Rudimentary Vision Test (BRVT) at 0.25 to 1 m with an external illumination on 395 ± 10% lux (Ferris et al., 1982). Standard measurement procedures with letter-by-letter scoring were incorporated during the static VA assessments using ETDRS (Klein et al., 1983; Ferris and Bailey, 1996). Participants started reading the letters at the top of the chart and continued to read down until they could no longer identify a minimum of three out of the five letters on a line correctly. To ensure that VA could be calculated at the borders near the limits of measurement of both charts (i.e., where both charts overlap), the single-letter BRVT tumbling E targets were each presented 5 times as there are 5 letters per line on the ETDRS charts, and letter-by-letter scoring was used (each letter correctly identified valued at 0.02 logMAR; Vanden Bosch and Wall, 1997). The ETDRS and BRVT charts were chosen to measure static VA in this study, because these charts used by the IPC for classification of athletes with vision impairments.

Contrast sensitivity was measured using the quick CSF (contrast sensitivity function) procedure on an Adaptive Sensory Technology platform (AST, Germany). The AST platform consisted of a 46″ NEC P463 screen with 1920 × 1080 resolution, calibrated to 90-cd/m^2 background luminance. At a viewing distance of 1 m, the screen allowed a display of stimuli in a spatial frequency range of 0.35 to 9 cycles per degree. It was possible to present contrast levels down to 0.2% reliably (Dorr et al., 2017b). Three letters were presented horizontally during a trial with the left and middle letters displayed at four and two times the contrast of the right letter, respectively. A CSF was calculated after 25 trials. The area under the log CSF curve (AULCSF, logCS units) calculated by the software was used as the summary statistic for the CS assessments (Hou et al., 2010; Jia et al., 2014). AULCSF was chosen as the summary measure of CSF because it has been reported to have better predictive power and test–retest precision compared to peak CSF using fractional rank precision analyses (Dorr et al., 2017a,b).

A novel device (D&zzle, V&mp Vision Suite, University of Waterloo) was used to measure GLS and GLR. GLS was estimated by measuring the static VA of participants immediately after introducing a bright, binocular glare source in the line of sight. GLR was measured by retesting the static VA 1 minute after the glare source was removed. Static VA in the presence of and after removing the glare source was compared to the baseline static VA (no glare) to determine the GLS and GLR, respectively. LS was assessed by measuring the static VA of participants at increased light levels (approximately 1900 lux, both in the surround as well as on the chart). Static VA in the presence of the bright light was compared to the baseline static VA to determine LS in logMAR units. GLS was calculated using the following formula: $\text{GLS} = \text{Static VA in the presence of glare} - \text{Static VA}$. GLR and LS were also calculated similarly. Positive logMAR values for GLS, GLR, and LS indicated that VA worsened compared to baseline during the respective testing conditions.

Dynamic VA was measured using the computer program moV& (V&mp Vision Suite, Waterloo, Canada) with a single moving tumbling *E* letter that moved in a random walk trajectory at a speed of 1 m/s and was presented on a high definition television screen (50″ or 60″ display, 60 Hz refresh rate and 1920 × 1080 resolution, illuminance at 130–150 lx) at a distance of 1 m (Hirano et al., 2017). The initial size of the letter presented was 0.60 log units bigger than the participant's static VA to make sure that the subject started the test from a suprathreshold level, and the maximum letter size presentable on this screen was 2.60 logMAR at a distance of 1 m. Five targets were presented per 0.1 logMAR step, and the display time was set to be unlimited to ensure adequate time to respond to the direction of the letter *E*. The test sequence continued until the participant could no longer correctly identify three out of five targets of the same size. Dynamic VA was also recorded in logMAR units, using a per letter scoring system (each letter correctly identified valued at 0.02 logMAR; Bailey et al., 1991).

Random dot kinematograms consisting of 100 individual, full-contrast, local dots that were equivalent to the size of the target detail of a 2.00 logMAR letter were used to assess two types of global motion tasks: translational (up and down) motion (TMP) and radial (in and out) motion (RMP; Dalton et al., 2017). Stimuli were presented on high-definition television screens (50″ or 60″ displays, 60 Hz refresh rate and 1920 × 1080 resolution, illuminance at 130–150 lx). On each TMP and RMP trial, the stimulus was presented for a maximum of 16 s, and participants were asked to identify the motion direction of the signal dots. The testing followed a staircase method, which was terminated after eight reversals, and the threshold was calculated by averaging the last six reversals.

A VF was binocularly assessed using an Arc perimeter and recorded following the standardized protocol, which was modified to allow binocular measurement (Grosvendor, 2007). VF assessments were performed by the same examiner (AS), moving a 6-mm-diameter target (size IV) from non-seeing areas (starting from 90° eccentricity) in the far periphery to the seeing areas at a speed of approximately 3–5 degrees per second. Once the VF boundary was identified, the target was moved continuously along each axis toward the central fixation point to identify any scotoma, if present. The edges of scotomata were reassessed until the response was consistent and reliable. Testing always started with the horizontal axis, and once the horizontal axis was marked, the arc was rotated to test the entire 360 degree VF in 30-degree intervals (Grosvendor, 2007). The current Paralympic classification uses Humphrey field analyzer, Octopus, and Goldman perimeters for assessing athletes' VFs. However, it was not feasible to use one of these instruments in this study because the study locations were remote on the ski venues; therefore, the Arc perimeter was chosen because of its portability. The VF assessment method used during the data collection for this study was found to be reliable and valid when compared with the Humphrey field analyzer in a separate study (Stalin, 2020). An unbiased modified AMA scoring method (AMA 6E, **Figure 1**) designed by Mann and Ravensbergen was used for functionally scoring VF data of the nordic and alpine participants in the studies in order to ensure no assumptions were made about which aspects of the VF had more importance in skiing performance. The maximum possible VF score was 60, and the scores were converted to percentages [i.e., (AMA 6E score/60) *100] (Mann and Ravensbergen, 2019). The VF scoring method used in the study was also validated and compared with the Humphrey field analyzer (Stalin, 2020).

Skiing Performance

Multiple confounding factors such as fatigue, jetlag, weather conditions, anxiety, or an illness could affect an individual's skiing performance; therefore, we calculated the overall performance points for each participant based on the World Para nordic Skiing (WPNS) and World Para Alpine Skiing (WPAS) scoring systems rather than choosing a single race. The WPNS and WPAS scoring systems award skiers points, based on their best performances in a rolling validity period. In WPNS, skiers' best five performances in a 24-month window are used to determine skiers' performance points, while in WPAS, skiers' best two performances in a 15-year period are used. The performance points in WPAS are discipline specific [downhill (DH), super G (SG), giant slalom (GS), and slalom (SL)], but not in WPNS.

In both Para nordic and Para alpine skiing, skiers with VI compete for one medal, regardless of their class. The WPNS and WPAS scoring systems adjust skiers' race times by a class factor, such that skiers with most severe impairments receive a maximum time bonus. In order to ensure that the skiing performance metric was not impacted by skiers' previous classification, skiers' points across the season were recalculated without the class factor, so that performance was determined based on skiers' raw times. Performance points calculated in these studies are referred to as raw-WPNS or raw-WPAS points to differentiate them from publicly available, published race results.

The formula for calculating unfactored race points was $P = \left(\left(\frac{Tx}{T0}\right) - 1\right) * F$+race pentalty, where P = race points, T_X = raw race time of competitor in seconds, T_0 = raw race time of the overall gender best performer in seconds, and F = discipline factor (determined by IPC and reevaluated once in every 2 years based on competition results). The race penalty is another factor determined by the IPC to account for the quality of competition and ensures that race points from different competitions can be compared equitably. Using this formula, best skiers have the lowest performance points (World Para Nordic Skiing, 2018, 2019; World Para Alpine Skiing, 2019).

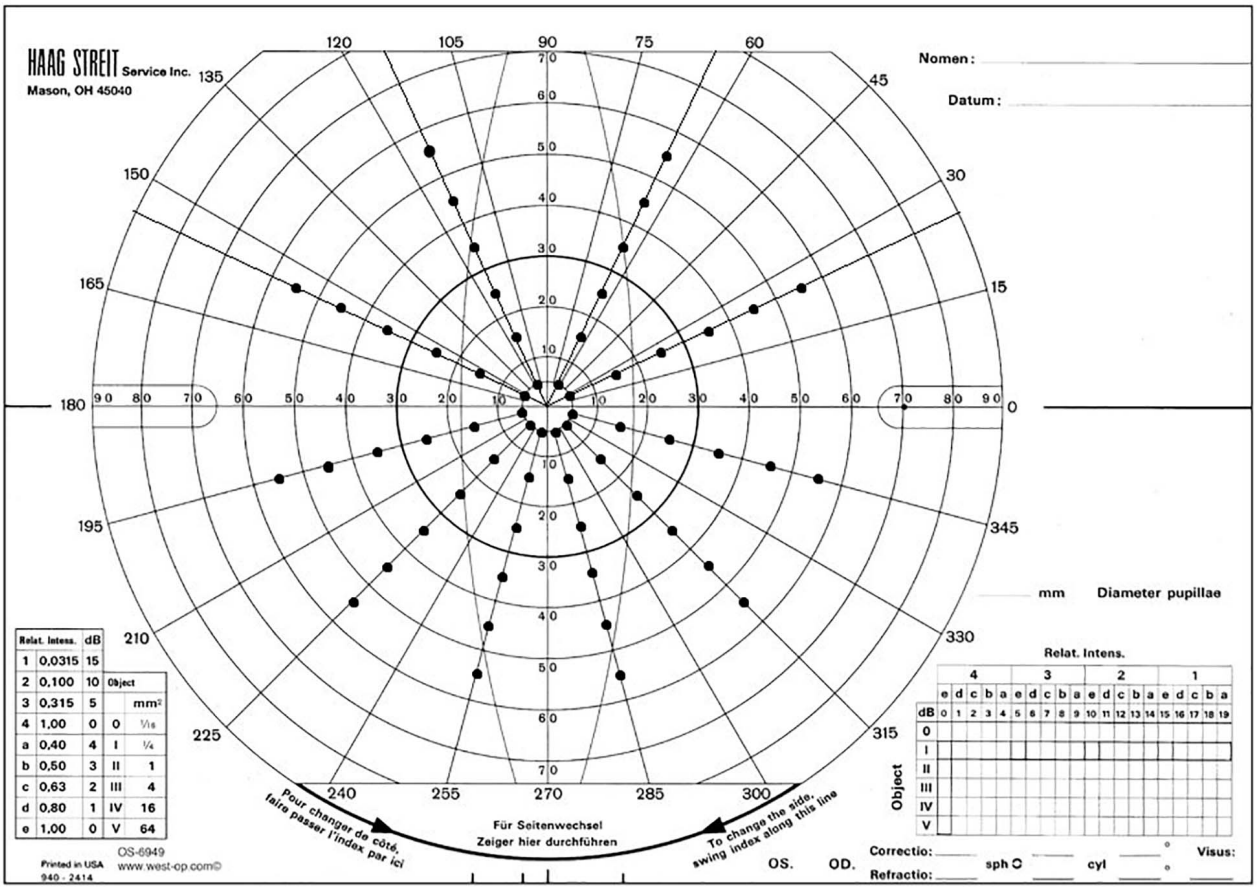
This formula calculates race points relative to the race time of the overall best performer in each race, for each gender. Previous research has also demonstrated that gender does not affect visual functions, such as VA or CS (Beck et al., 1993; Ohlsson and Villarreal, 2005; Grobbel et al., 2016). As performance points were normalized to the best performance in each gender and visual function does not appear to differ between genders, researchers were able to compare performance data between genders, which was important because of the small number of elite alpine and nordic skiers with vision impairment in the world.

For Para nordic skiing, the validity period for the calculation of raw-WPNS points used in this study was from 1 April 2016 to 31 March 2018. For Para alpine skiing, the validity period used was from 1 January 2016 to 31 March 2017. All skiers included in these studies completed the minimum number of races needed to calculate their raw-race points based on the sport rules (five for Para nordic and two for Para alpine). Not all Para alpine skiers in this study competed in each discipline, but all skiers included in the study completed at least two races in at least one discipline.

The use of other tests to quantify confounding factors related to skiing performance, such as visual motor reaction times under different physiological conditions (i.e., fatigue, anxiety) to measure attention, and tests of muscle strength and flexibility was considered and ultimately dismissed. The wide range of vision impairments in the study populations, including athletes with marked VF loss and athletes with no light perception (NLP), made visual motor reaction time testing unfeasible. Tests of muscle strength and flexibility would have certainly provided insight into athlete fitness levels, but male and female athletes would have likely performed differently on these tests and stratifying the performance analysis by gender would have further reduced our sample size and statistical power.

Data Analysis

Data analysis (SPSS for Windows, version 25.0, SPSS, Inc.) focused on (1) determining the associations between skiing performance with vision-related and non-vision-related variables such as skiers' age, age started skiing, age of onset of impairment, total lifetime hours of skiing, and number of races completed in the period that the skiing performance points were calculated for, and (2) identifying the visual function assessments, which could be predictive of skiing performances



Relat. Intens.		dB
1	0,0315	15
2	0,100	10
3	0,315	5
4	1,00	0
a	0,40	4
b	0,50	3
c	0,63	2
d	0,80	1
e	1,00	0

Object	mm²
0	1/16
I	1/4
II	1
III	4
IV	16
V	64

FIGURE 1 | Modified AMA scoring grid on a Goldmann VF scoring sheet for functionally scoring VF. This figure was adapted from the unpublished report of Mann and Ravensbergen. Protocol for AMA-Style Analysis of Visual Field, 2019 (Mann and Ravensbergen, 2019).

taking into account the non-vision variables. Based on the recommendations of the Joint Position Stand on Paralympic classification, correlation models and regression analyses were chosen to identify the visual functions that are predictive of sports performance and to compare the visual function measures and the performance measures (Mann and Ravensbergen, 2018). Kendall τ was chosen for correlation analysis as it guards against outliers among the marginal distributions and is reported to have smaller gross error sensitivity and asymptotic variance, making it more robust and efficient compared to the Spearman correlation (Croux and Dehon, 2010; Wilcox, 2016). Bonferroni–Holm corrections were used to account for multiple comparisons in the correlation analyses as it is a more powerful sequentially rejective multiple-testing approach that strongly controls the family-wise error rate compared to the traditional Bonferroni corrections (Holm, 1979; Simes, 1986). Considering the small sample sizes, including all 14 independent variables would have resulted in overfitting and increased variation inflation factors (VIF; Hawkins, 2004; Kim, 2019). Therefore, Kendall τ correlations were used as a guideline for choosing variables for multivariable regressions. Any variables that demonstrated significant ($p < 0.05$) or near-significant ($p < 0.1$) correlations with skiing performance were included in the multivariable regression models, conducted using the enter method. Multicollinearity and VIFs were assessed before finalizing the variables. There were no outliers in the data, and the assumptions of normality and homoscedasticity were met in both Para nordic and Para alpine data (Olive, 2017). In Para alpine, each discipline was analyzed separately.

Seven Para nordic and two Para alpine participants had light perception (LP) or NLP vision, and values of 3.8 logMAR and 4.2 logMAR were arbitrarily assigned for their static VA, respectively, so they could be included in the correlation and regression analyses on the same continuous scale as the other participants. Similarly, values of 0.00 log CS and 0.0% were assigned for these participants' CS and VF measures, respectively. Dynamic VA has been reported to be between 0.20 and 0.30 logMAR worse than the static VA in individuals with normal vision (Hirano et al., 2017), but it was impossible to predict how much worse dynamic VA would be relative to static VA for each individual with vision impairment. Assigning 0.00 logMAR values for GLS, GLR, or LS for these participants would indicate that their static VA did not change with glare or increased light intensity rather than they were unable to

do the task. Similarly, assigning a 100% value for their TMP and RMP would indicate that they were able to perceive the motion at 100% coherence, not that they were unable to do the task. Therefore, it was not appropriate to assign the same, or adjusted, arbitrary values for dynamic VA, GLS, GLR, LS, TMP, or RMP and no values were substituted for these visual function parameters in the Para nordic and Para alpine participants with LP or NLP vision.

A priori power analysis (G*power 3.1.9.7) indicated that sample sizes of 85 would be required to obtain a minimum level of power of.80 (Cohen, 2013) with an alpha of.05 with medium effect size (0.15; Faul et al., 2007). However, it was impossible to recruit 85 skiers for each study as there were only 46 Para nordic and 34 Para alpine elite skiers with vision impairment in the world.

Post hoc power analysis (G*power 3.1.9.7) indicated that the power to detect the obtained effects at the effect size of 0.15 and alpha error probability of 0.05 were 0.36 in the Para nordic, 0.20 in GS and SL, and 0.28 in SG for the regression analyses in prediction of skiing performance (Faul et al., 2007). These analyses suggest that the Para nordic and Para alpine studies did not have sufficient power to support the analysis results but considering the uniqueness of study population and the fairly representative samples in the studies, the results are still meaningful.

RESULTS

Twenty-six Para nordic skiers from 13 nations and 15 Para alpine skiers from 10 nations who competed at the events where testing took place participated in these studies (**Table 1**). Summary visual function data for each sport are found in **Tables 2**, **3**. The arbitrarily assigned values for static VA, CS, and VF were not included in the calculation of means and standard deviations presented in these summary tables because they were not actual measured values of the participants. The arbitrary values were only included in the correlation and regression analyses.

Among the Para nordic participants, five had NLP and two had LP vision. Among the Para alpine participants, one had NLP and one had LP vision. One Para alpine participant also had very good static VA (-0.04 logMAR); however, this participant had qualified for the competition based on the extent of their VF (7.5°radius).

Both the Para nordic and Para alpine skiers had a broad range of ocular pathologies. Ocular diseases affecting the central retina, peripheral retina, and total retina were most common among Para nordic and Para alpine skiers. 62% of the Para nordic participants and 53% of Para alpine participants had onset of VI after age 2. Forty percent (40%) of the Para nordic and 63% of the Para alpine skiers had VI conditions that were progressive. The most common VF defect among both Para nordic and Para alpine skiers was a peripheral VF constriction. Further details on the types of VF defects can be found in **Supplementary Appendix B**.

The average raw-WPNS points of Para nordic participants was 58.73 ± 52.44 (range: 0.00 to 172.07, N = 26). The average raw-WPAS points of Para alpine participants for DH discipline was 155.81 ± 66.36 (range: 33.99 to 254.19, N = 9), GS was 226.98 ± 212.13 (range: 51.11 to 854.02, N = 15), SG was 336.20 ± 341.34 (range: 50.09 to 1299.41, N = 13), and SL was 193.40 ± 185.03 (range: 66.77 to 722.13, N = 15).

TABLE 1 | Participant details and summary statistics of their non-vision variables by sport.

	Para nordic	Para alpine
Number of athletes	26	15
Gender	18 Male; 8 female	8 Male; 7 female
Number of nations	13	10
Age (years)	26.0 ± 6.3	28.1 ± 11.6
Age range (years)	18 to 43	16 to 58
Age started skiing (years)	12.8 ± 8.2	16.2 ± 8.2
Age of onset of impairment (years)	6.8 ± 8.1	5.3 ± 7.1
Total lifetime hours of skiing	4545.5 ± 3883.5	4239.3 ± 4094.0
Number of races during the validity period	12.2 ± 4.9	DH: 6.8 ± 2.1 (N = 9)
		GS: 8.9 ± 3.4 (N = 15)
		SG: 7.4 ± 3.4 (N = 13)
		SL: 13.7 ± 5.0 (N = 15)

The summary statistics presented in this table include the mean ± standard deviation of all the athletes.

TABLE 2 | Summary of visual function assessments of Para nordic skiing participants.

Visual function tests	*N*	Mean ± SD	Median	Range
Static visual acuity (logMAR)	19	1.71 ± 0.40	1.60	1.18 to 2.68
Contrast sensitivity (logCS)	19	0.21 ± 0.26	0.12	0.00 to 0.82
Glare sensitivity (change in logMAR)	19	0.20 ± 0.31	0.10	–0.19 to 0.98
Glare recovery (change in logMAR)	19	0.06 ± 0.20	0.00	–0.20 to 0.79
Light sensitivity change in logMAR)	19	0.00 ± 0.09	0.00	–0.15 to 0.16
Dynamic visual acuity (logMAR)	16	1.80 ± 0.31	1.80	1.20 to 2.20
Translational motion perception (%)	15	59.8 ± 26.9	61.8	19.2 to 100.0
Radial motion perception (%)	15	62.8 ± 28.5	61.2	26.5 to 100.0
Visual field (%)	19	63.9 ± 26.9	71.7	3.3 to 100.0

Only the participants' data with measurable results on each test are included. The data presented do not include participants with LP or NLP vision.

Associations of Visual Functions and Skiing Performance

In Para nordic skiing, participants' raw-WPNS points were significantly correlated with the AMA 6E scoring of VFs (p = 0.011). There were also trends toward significance for static VA (p = 0.066) being correlated with raw-WPNS points (**Table 4** and **Figure 2**).

TABLE 3 | Summary of visual function assessments of Para alpine skiing participants.

Visual function tests	*N*	Mean ± SD	Median	Range
Static visual acuity (logMAR)	13	1.20 ± 0.51	1.40	0.04 to 1.64
Contrast sensitivity (logCS)	13	0.53 ± 0.59	0.40	0.00 to 1.90
Glare sensitivity (change in logMAR)	13	0.19 ± 0.17	0.14	0.02 to 0.54
Glare recovery (change in logMAR)	13	0.05 ± 0.08	0.02	–0.06 to 0.18
Light sensitivity change in logMAR)	10	0.09 ± 0.14	0.04	–0.08 to 0.34
Dynamic visual acuity (logMAR)	11	1.48 ± 0.57	1.40	0.50 to 2.20
Translational motion perception (%)	12	56.4 ± 31.9	53.3	9.3 to 100.0
Radial motion perception (%)	12	56.8 ± 29.0	55.3	12.8 to 100.0
Visual field (%)	13	53.5 ± 28.5	55.0	16.7 to 100.0

Only the participants' data with measurable results on each test are included. The data presented do not include participants with LP or NLP vision.

In Para alpine, static VA was significantly associated with raw-WPAS points in DH (p = 0.046), GS (p = 0.010), and SG (p = 0.007) and was nearly significant in the SL discipline (p = 0.074). TMP was significantly associated with raw-WPAS points in SG (p = 0.041) and VF was significantly associated with raw-WPAS points SL (p = 0.013). TMP also demonstrated a trend toward significance with raw-WPAS points in DH (p = 0.095) and GS (p = 0.084). CS was significantly associated with raw-WPAS points in GS (p = 0.017) and SG (p = 0.017) and was nearly significant in DH (p = 0.06) and SL (p = 0.06; **Table 4** and **Figures 3**–**6**).

None of the correlations in the Para nordic or Para alpine data were significant after applying the Bonferroni–Holm correction. The summary of the correlation analyses of the vision-related variables with the adjusted, Bonferroni–Holm-corrected, p-values in the Para nordic and Para alpine data are provided in **Table 5**.

Associations of Non-Vision-Related Variables and Skiing Performance

Participants' number of races during the point calculation period was significantly correlated with the raw-WPNS points (p = 0.010) and SG raw-WPAS points (p = 0.031). Total hours of skiing in lifetime was also nearly significantly correlated (p = 0.098) with raw-WPNS points (**Table 6**). None of the correlations in the Para nordic or Para alpine data were significant after applying the Bonferroni–Holm correction (**Table 7**).

Associations Between Visual Functions and Non-Vision Related Variables

In the Para nordic study population, one of the non-vision variables—number of races—had significant correlations with static VA (τ_b = -0.45, p = 0.002) and VF (τ_b = 0.314, p = 0.032). None of the other non-vision variables had significant correlations with any of the other visual functions in the Para nordic study population ($p > 0.05$). There were no significant correlations between any of the visual functions and non-vision variables in the Para alpine study population ($p > 0.05$).

Visual Functions Predictive of Skiing Performances

Multivariable regression analysis was used to look at whether or not skiing performances could be predicted based on any of the individual visual functions measured. Static VA and VF were the only two visual function variables that demonstrated strong enough correlations ($p < 0.1$) with Para nordic skiing performance to be considered in the model. Based on the correlation analyses, static VA, VF, number of races, and total hours of skiing were included in the modeling for Para nordic skiing performance, and a significant regression equation was found $F(4,21) = 7.12$, p = 0.001, and R^2 = 0.58. Para nordic predicted raw-WPNS points were equal to 130.484 – 3.981 (number of races) – 0.006 (Total hours of skiing; Stalin et al., 2019). In other words, a participant's Para nordic skiing performance improved by 3.981 points for each race competed by the participant during the points calculation period and by 0.006 points for each hour of skiing.

Static VA, dynamic VA, CS, TMP, and VF showed a significant association or were nearly significantly associated with

TABLE 4 | Summary of correlations of visual functions with skiing performances; p-values are presented in the table with sample sizes, and significant correlations are provided in bolded text.

Variable	Raw-WPNS points	DH Raw-WPAS points	GS Raw-WPAS points	SG Raw-WPAS points	SL Raw-WPAS points
Static VA (logMAR)	τ_b = 0.26, p = 0.066 (26)	**τ_b = 0.54, p = 0.046 (9)**	**τ_b = 0.50, p = 0.010 (15)**	**τ_b = 0.57, p = 0.007 (13)**	τ_b = 0.35, p = 0.074 (15)
CS (logCS)	τ_b = −0.23, p = 0.124 (26)	τ_b = −0.50, p = 0.061 (9)	**τ_b = −0.46, p = 0.017 (15)**	**τ_b = −0.51, p = 0.017 (13)**	τ_b = −0.37, p = 0.059 (15)
GLS (change in logMAR)	τ_b = 0.18, p = 0.301 (19)	τ_b = 0.31, p = 0.206 (9)	τ_b = 0.21, p = 0.357 (13)	τ_b = −0.02, p = 1.000 (11)	τ_b = 0.08, p = 0.759 (13)
GLR (change in logMAR)	τ_b = 0.21, p = 0.225 (19)	τ_b = 0.48, p = 0.075 (9)	τ_b = −0.01, p = 0.951 (13)	τ_b = −0.13, p = 0.583 (11)	τ_b = 0.12, p = 0.668 (13)
LS (change in logMAR)	τ_b = −0.06, p = 0.724 (19)	τ_b = −0.20, p = 0.543 (7)	τ_b = −0.21, p = 0.417 (10)	τ_b = −0.33, p = 0.262 (8)	τ_b = −0.16, p = 0.528 (10)
Dynamic VA (logMAR)	τ_b = −0.22, p = 0.238 (16)	τ_b = 0.59, p = 0.044 (8)	τ_b = 0.25, p = 0.306 (11)	τ_b = 0.46, p = 0.092 (9)	τ_b = −0.06, p = 0.813 (11)
TMP (%)	τ_b = −0.10, p = 0.728 (15)	τ_b = 0.44, p = 0.095 (9)	τ_b = 0.39, p = 0.084 (12)	**τ_b = 0.49, p = 0.041 (11)**	τ_b = 0.23, p = 0.299 (12)
RMP (%)	τ_b = −0.24, p = 0.317 (15)	τ_b = 0.03, p = 0.917 (9)	τ_b = 0.08, p = 0.731 (12)	τ_b = −0.04, p = 0.876 (11)	τ_b = −0.02, p = 0.945 (12)
VF (%)	**τ_b = −0.37, p = 0.011 (26)**	τ_b = 0.09, p = 0.753 (9)	τ_b = −0.33, p = 0.090 (15)	τ_b = −0.34, p = 0.110 (13)	**τ_b = −0.49, p = 0.013 (15)**

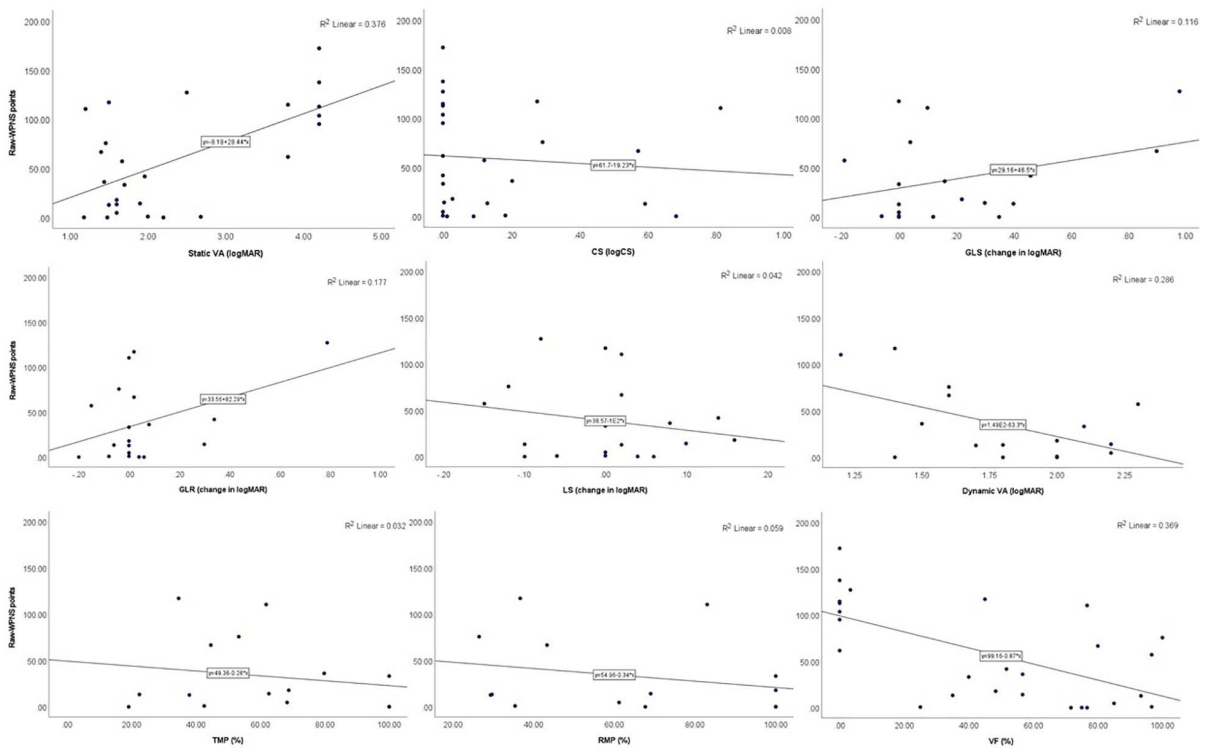


FIGURE 2 | Scatter plots showing the relationships between raw-WPNS points and visual functions.

performance in one or more of the Para alpine disciplines and were considered for inclusion in the multivariable regression. Static VA had strong significant correlations with skiing performances in most of the Para alpine disciplines, and static VA was also strongly correlated with dynamic VA, CS, and TMP, likely due to the wide range of VI among participants in the study populations (Dalton et al., 2019; Stalin et al., 2019). Including all these variables in modeling resulted in multicollinearity and high VIF (Kim, 2019). To avoid the instability and the overfitting due to multicollinearity in the multivariable regression model, static VA was chosen for inclusion in the model as static VA had the strongest correlations overall with each of the other visual functions and skiing performance. Thus, the final regression model for Para alpine included static VA, VF, skier's age, and the number of races.

There was no significant regression equation for DH [$F(4,4) = 0.46$, $p = 0.76$, and $R^2 = 0.32$]. For GS, a significant regression equation was found $F(4,10) = 14.36$, $p < 0.001$, and $R^2 = 0.85$. Para alpine participants' predicted GS raw-WPAS points were equal to -74.472 + 166.991 (static VA) +5.557 (age), where static VA was measured in logMAR units and age was measured in years. Participants' Para alpine GS skiing performance points deteriorated by 166.991 points for each 1.00 logMAR increase in static VA (worsening) and by 5.557 points for an increase in each year of age.

For SG, a significant regression equation was found $F(4,8) = 8.71$, $p = 0.05$, and $R^2 = 0.81$. Para alpine participants' predicted SG raw-WPAS points were equal to 13.714 + 217.007 (static VA), where static VA was measured in logMAR units. Participants' Para alpine SG skiing performance points deteriorated by 217.007 points for each 1.00 logMAR increase in static VA.

Similar to the GS results, a significant regression equation was found $F(4,10) = 14.66$, $p < 0.01$, and $R^2 = 0.85$ for SL performance points. Para alpine participants' predicted SL raw-WPAS points were equal to -164.532 + 145.066 (static VA) +5.739 (age), where static VA was measured in logMAR units and age was measured in years. Participants' Para alpine SL skiing performance deteriorated by 145.066 points for each 1.00 logMAR increase in static VA and by 5.739 points for an increase in each year of age.

DISCUSSION

These studies were conducted to identify visual functions associated with, and predictive of, Para nordic, and Para alpine skiing performance. Ideally, large study populations would have been recruited to assess the significance of a broad range of vision functions on the skiing performance, as was done in these studies, because the high variation in the other non-vision factors could mask the effects of vision on performance. To ensure that the variations in non-vision factors such as skill development, training, and coaching levels were as small as possible between the participants, only elite skiers were recruited for these studies.

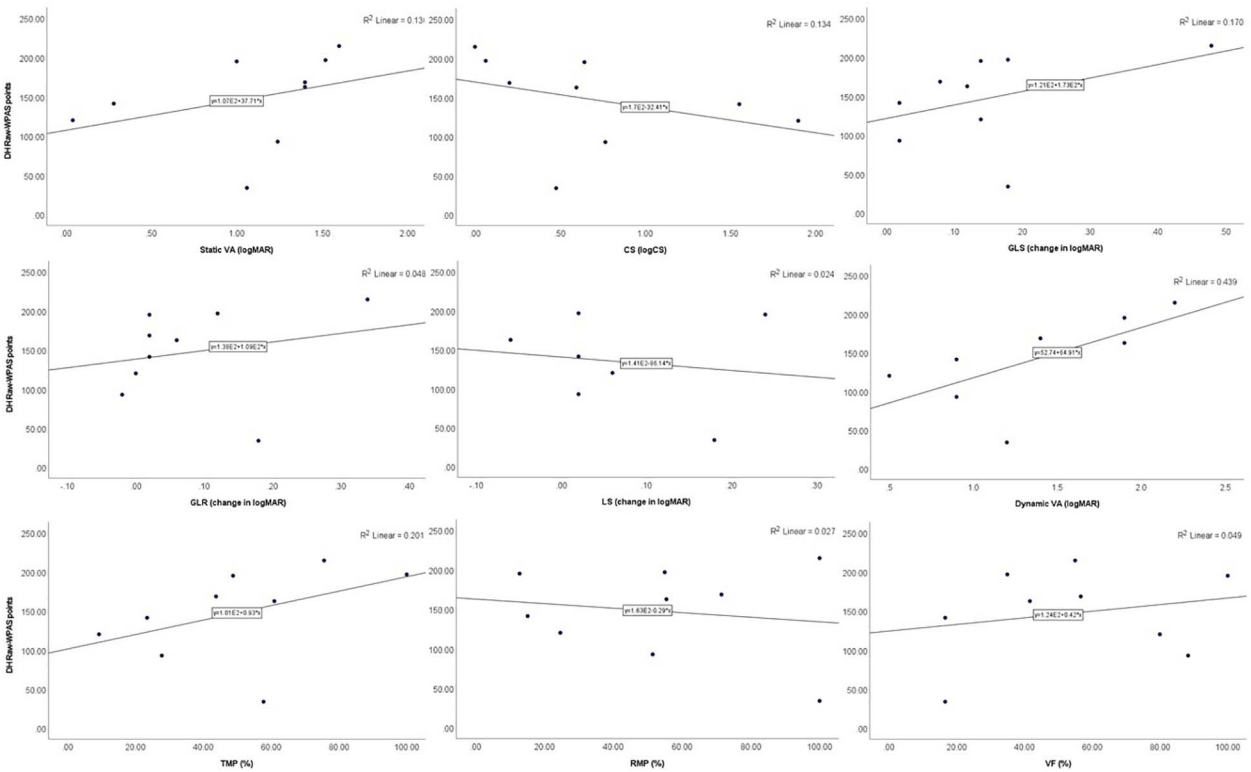


FIGURE 3 | Scatter plots showing the relationships between DH raw-WPAS points and visual functions.

Additionally, in recognition that the sample sizes were small, robust statistical analysis methods such as Kendall τ correlations were conducted and Bonferroni–Holm *post hoc* adjustments were done. *Post hoc* power analyses suggested that these studies did not have the power to support the study results due to the limited sample sizes. However, the Para nordic and Para alpine skiers' populations in the world are unique and small, making it impossible to obtain large sample populations for the studies. These studies recruited approximately half of the entire world's World Championship eligible Para nordic and Para alpine populations registered to compete at the time of the study and over 60% of the athletes in each sport who were competing at the events where testing took place. Therefore, despite the small sample sizes, it can be seen that the study populations in these studies were very representative of the populations of Para nordic and Para alpine skiers with vision impairment.

While various competition events within Para nordic skiing sport (e.g., sprint, middle distance, and long distance) differ mainly in terms of only the length of the courses, disciplines within the Para alpine skiing sport differ in terms of the terrain characteristics and the skiing techniques used. As a result, these Para alpine disciplines also differ in terms of the skiers' participation. Slalom and GS are the two most popular Para alpine disciplines, with participation from skiers with the most severe impairments (B1 class). DH had the least number of participants and rarely has participation from skiers in the B1 class, probably due to the increased speed (maximum speed of 150 km/h) and high visual demands involved in DH (Gilgien et al., 2018). Additionally, DH courses are steeper and longer compared to the courses used in other alpine disciplines, which limits their availability to some athletes for training. The limited availability of well-groomed DH courses for training might also have reflected in reduced participation. It is also reported that DH is the alpine discipline that is reported to have the highest injury incidence rates (Flørenes et al., 2011). These variations in skiers' participation were reflected in the Para alpine population participated in this study as well.

The results of the Para nordic study suggested that even though static VA had possible association and VF had strong association with Para nordic skiing performance, neither individual vision function was predictive of Para nordic skiing performance. Multiple factors such as training, skill development, and coaching levels influence the performance of skiers in addition to various physical and psychological factors, which are unique to each individual. Participants were chosen from an elite population to specifically minimize the impact of variations in such non-vision factors on the skiing performance of the study populations. However, even within such an elite Para nordic study population, the only predictive factors of skiing performance were the number of races that the participants competed in and skiers' total lifetime hours. Therefore, the training variables seemed to have a more significant impact on Para nordic skiing performance than static VA and VF.

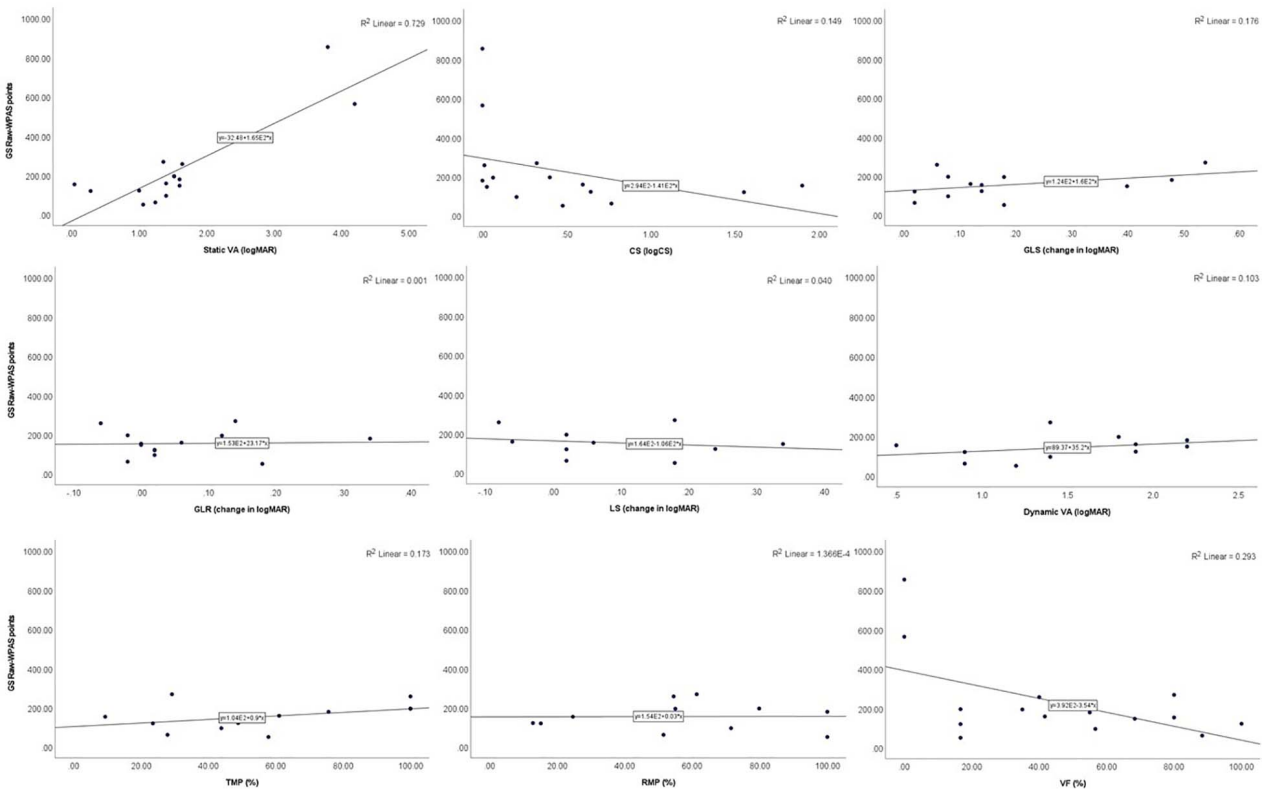


FIGURE 4 | Scatter plots showing the relationships between GS raw-WPAS points and visual functions.

None of the other vision variables, including CS, seemed to be associated with Para nordic skiing performance, contrary to expectations.

Consistent with the previous reports on the visual demands associated with alpine skiing, performance in the Para alpine technical disciplines, which have participation from skiers with a wide range of VI, seemed to be predicted by the static VA and associated with VF when the age of skiers was taken into account (Craybiel et al., 1955; Senner et al., 1999). Better static VA was predictive of better GS, SG, and SL performance points, which require more technical skill and less speed than DH. Similarly, better CS also appeared to be associated with, though not predictive of, skiers' performance in the Para alpine skiing technical disciplines. Though not predictive, better VF was also associated with better SL performances. Therefore, static VA and CS appeared to be associated with the performance of Para alpine skiers in all the disciplines except in DH. Slalom is the most technical alpine discipline, requiring athletes sometimes to look several gates ahead and shift their gaze frequently between different gate positions (Decroix et al., 2017). The need to attend to multiple gates might be the reason behind the association of VF with the SL performance. Better performance in the DH discipline, which requires more speed, was associated with better dynamic VA. TMP, RMP, GLS, GLR, and LS were not significantly associated with performance in either sport.

Static VA is one of the most common assessments of spatial vision and is extremely useful to detect deficits in the visual system. Although the static VA does not seem to directly provide information about the perception of low-contrast images or objects in motion, previous research had reported significant strong correlations between the static VA and measures of CS and dynamic VA (Burg and Hulbert, 1961; Burg, 1966; Lesmes et al., 2012). Consistent with previous literature, static VA of both Para nordic and Para alpine skiers was significantly strongly associated with dynamic VA and CS in current studies (Dalton et al., 2019; Stalin et al., 2019) and similar to those reported in low vision populations, especially when CS was measured using the qCSF method (Lesmes et al., 2012; Stellmann et al., 2015). Correlations between CS and static VA have been shown to be the strongest in heterogeneous populations, such as the population of athletes with diverse types of vision impairments studied here (Haegerstrom-Portnoy et al., 2000). Thus, it is possible that the strong correlations between static VA and CS found in this population of skiers could have masked the relationships between CS and skiing performance in these studies.

In addition to the small study sample size, these studies had some limitations. Depth perception has the potential to be one of the vision functions associated with skiing performance; however, due to the lack of a feasible instrument to measure distance stereopsis in low vision populations, it was not possible to include an assessment for depth perception in the test battery. Para nordic

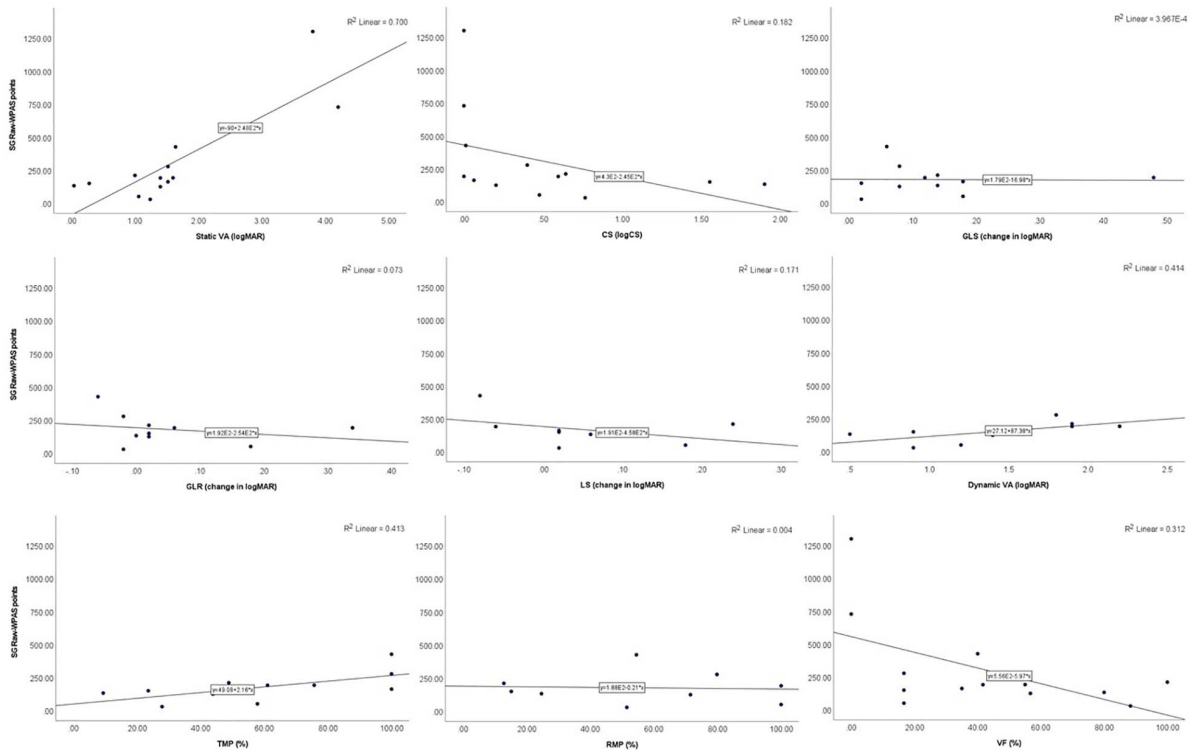


FIGURE 5 | Scatter plots showing the relationships between SG raw-WPAS points and visual functions.

and Para alpine skiers also wear a wide range of ski goggles when competing for protection from external elements such as wind and glare. The impact of these ski goggles on the Para skiers' performance is currently not known, but none of the glare or light sensitivity metrics (GLS, GLR, or LS) measured in the absence of tints and filters were significantly associated with skiing performance. As the use of tints and filters could affect various aspects of vision including CS and glare, future studies could explore the effect of using different tints or filters in the skiing sports.

The test battery for the study was chosen based on consideration of multiple factors such as the test availability, portability, current sport classification rules, and each test's precision and accuracy. While we used the qCSF, which generated automated CSFs with high precision, traditional manual methods were used for measuring the static VA and VF. The current Paralympic classification rules rely on printed charts (ETDRS and BRVT) to assess static VA and standard automated perimetry to assess VF. Printed charts are more prone to random and systematic errors (i.e., effect of inter/intra observer changes in instructions, termination rules, etc.; Rosser et al., 2003). However, the charts used in these studies are used by the sports for classification purposes, so the tradeoff in precision and accuracy is balanced with the increased utility for the sports. Data collection was all done at the ski venues which were in remote locations away from hospitals and optometric clinics in larger city centers. Therefore, an Arc perimeter was used for measuring the peripheral VFs in the studies as it was not feasible to use standard automated perimeters due to their lack of availability at the study locations and the limited portability of these instruments. The Arc perimeter was portable, was easy to use, and has been shown to be reasonably accurate when compared to a Humphrey automated perimeter to determine areas of seeing vs. non-seeing, as long as measurements are conducted by a single, trained researcher (Stalin, 2020).

Novel assessment methods such as the qVA (Adaptive Sensory Technology, Inc.) and qVFM (Adaptive Sensory Technology, Inc.) are potential tests that could be used in similar future studies. The qVA (Lesmes and Dorr, 2019; Zhao et al., 2021) and qVFM (Xu et al., 2019, 2020) are based on the Bayesian active learning and Bayesian adaptive method, respectively, and were reported to be high in precision and accuracy. However, the adoption of new tests would need to be carefully considered as tests included in classification need to be accessible globally and language independent so they can accommodate athletes from countries all over the world.

A modified AMA scoring system was used in these studies for functionally scoring the VFs of participants, to ensure no prior assumptions were made about which aspects of the VF were most important for skiing performance. Using this modified scoring system means the VF related results from these studies might not be directly comparable to future studies using different scoring methods, but the VF scoring system used here is being used consistently across all Paralympic classification research,

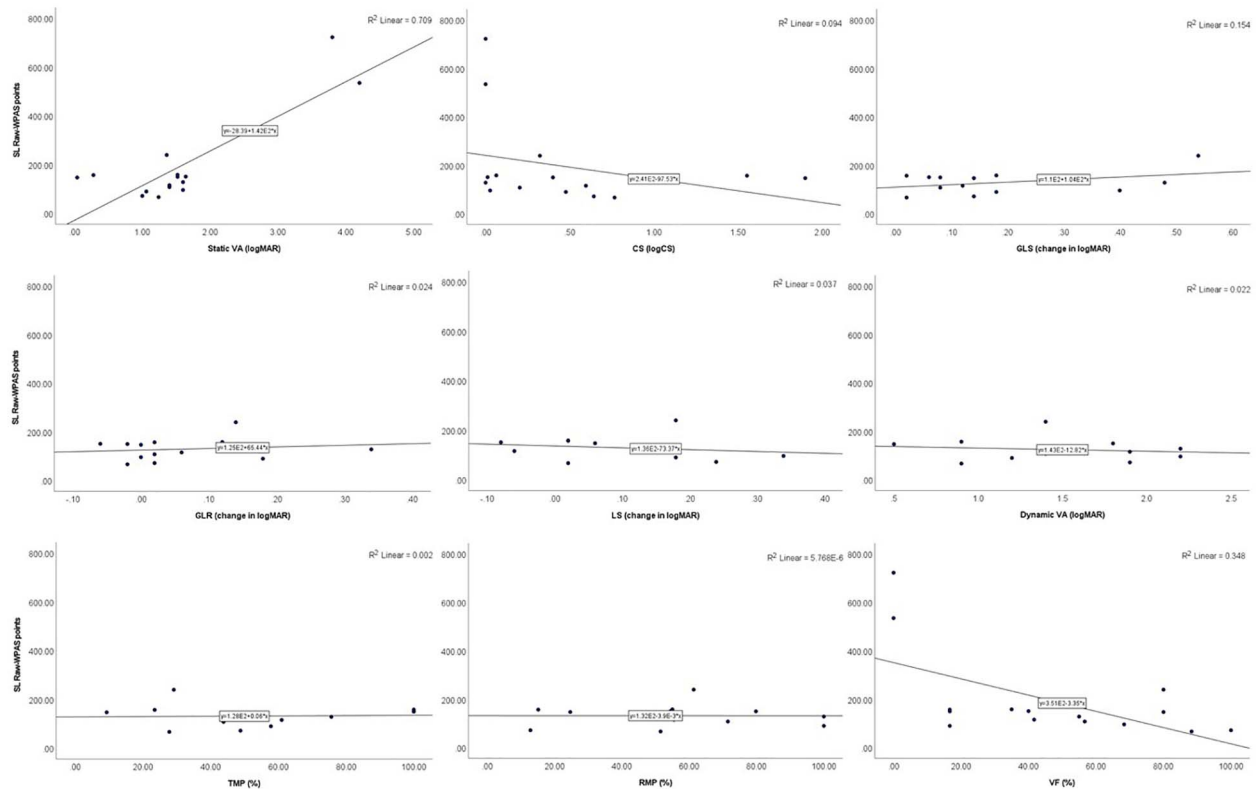


FIGURE 6 | Scatter plots showing the relationships between SL raw-WPAS points and visual functions.

TABLE 5 | Summary of correlations of visual functions with skiing performances.

Variable	Raw-WPNS points	DH Raw-WPAS points	GS Raw-WPAS points	SG Raw-WPAS points	SL Raw-WPAS points
Static VA (logMAR)	$\tau_b = 0.26$, $p = 0.792$ (26)	$\tau_b = 0.54$, $p = 0.598$ (9)	$\tau_b = 0.50$, $p = 0.140$ (15)	$\tau_b = 0.57$, $p = 0.098$ (13)	$\tau_b = 0.35$, $p = 0.77$ (15)
CS (logCS)	$\tau_b = -0.23$, $p = 1.000$ (26)	$\tau_b = -0.50$, $p = 0.72$ (9)	$\tau_b = -0.46$, $p = 0.221$ (15)	$\tau_b = -0.51$, $p = 0.221$ (13)	$\tau_b = -0.37$, $p = 0.708$ (15)
GLS (change in logMAR)	$\tau_b = 0.18$, $p = 1.000$ (19)	$\tau_b = 0.31$, $p = 1.000$ (9)	$\tau_b = 0.21$, $p = 1.000$ (13)	$\tau_b = -0.02$, $p = 1.000$ (11)	$\tau_b = 0.08$, $p = 1.000$ (13)
GLR (change in logMAR)	$\tau_b = 0.21$, $p = 1.000$ (19)	$\tau_b = 0.48$, $p = 1.000$ (9)	$\tau_b = -0.01$, $p = 1.000$ (13)	$\tau_b = -0.13$, $p = 1.000$ (11)	$\tau_b = 0.12$, $p = 1.000$ (13)
LS (change in logMAR)	$\tau_b = -0.06$, $p = 1.000$ (19)	$\tau_b = -0.20$, $p = 1.000$ (7)	$\tau_b = -0.21$, $p = 1.000$ (10)	$\tau_b = -0.33$, $p = 1.000$ (8)	$\tau_b = -0.16$, $p = 1.000$ (10)
Dynamic VA (logMAR)	$\tau_b = -0.22$, $p = 1.000$ (16)	$\tau_b = 0.59$, $p = 0.616$ (8)	$\tau_b = 0.25$, $p = 1.000$ (11)	$\tau_b = 0.46$, $p = 0.90$ (9)	$\tau_b = -0.06$, $p = 1.000$ (11)
TMP (%)	$\tau_b = -0.10$, $p = 1.000$ (15)	$\tau_b = 0.44$, $p = 1.000$ (9)	$\tau_b = 0.39$, $p = 0.96$ (12)	$\tau_b = 0.49$, $p = 0.492$ (11)	$\tau_b = 0.23$, $p = 1.000$ (12)
RMP (%)	$\tau_b = -0.24$, $p = 1.000$ (15)	$\tau_b = 0.03$, $p = 1.000$ (9)	$\tau_b = 0.08$, $p = 1.000$ (12)	$\tau_b = -0.04$, $p = 1.000$ (11)	$\tau_b = -0.02$, $p = 1.000$ (12)
VF (%)	$\tau_b = -0.37$, $p = 0.169$ (26)	$\tau_b = 0.09$, $p = 1.000$ (9)	$\tau_b = -0.33$, $p = 1.000$ (15)	$\tau_b = -0.34$, $p = 1.000$ (13)	$\tau_b = -0.49$, $p = 0.182$ (15)

The adjusted p-values based on the Bonferroni–Holm corrections are presented in the table with sample sizes.

TABLE 6 | Summary of correlations of non-vision variables with skiing performances.

Variable	Raw-WPNS points ($N = 26$)	DH Raw-WPAS points ($N = 9$)	GS Raw-WPAS points ($N = 15$)	SG Raw-WPAS points ($N = 13$)	SL Raw-WPAS points ($N = 15$)
Age (years)	$\tau_b = 0.05$, $p = 0.707$	$\tau_b = 0.06$, $p = 0.833$	$\tau_b = 0.17$, $p = 0.371$	$\tau_b = 0.12$, $p = 0.581$	**$\tau_b = 0.39$, $p = 0.047$**
Age started skiing (years)	$\tau_b = 0.19$, $p = 0.422$	$\tau_b = 0.31$, $p = 0.249$	$\tau_b = 0.17$, $p = 0.371$	$\tau_b = 0.03$, $p = 0.903$	$\tau_b = 0.31$, $p = 0.112$
Age of onset of impairment (years)	$\tau_b = -0.06$, $p = 0.657$	$\tau_b = -0.03$, $p = 0.914$	$\tau_b = -0.17$, $p = 0.411$	$\tau_b = -0.28$, $p = 0.218$	$\tau_b = 0.08$, $p = 0.681$
Total hours of skiing	$\tau_b = -0.23$, $p = 0.098$	$\tau_b = -0.06$, $p = 0.835$	$\tau_b = -0.11$, $p = 0.586$	$\tau_b = -0.18$, $p = 0.393$	$\tau_b = 0.03$, $p = 0.882$
Number of races	**$\tau_b = -0.37$, $p = 0.010$**	$\tau_b = -0.12$, $p = 0.669$	$\tau_b = -0.19$, $p = 0.343$	**$\tau_b = -0.46$, $p = 0.031$**	$\tau_b = -0.17$, $p = 0.371$

The unadjusted p-values are presented in the table with sample sizes, and significant correlations are provided in bolded text.

TABLE 7 | Summary of correlations of non-vision variables with skiing performances.

Variable	Raw-WPNS points (*N* = 26)	DH Raw-WPAS points (*N* = 9)	GS Raw-WPAS points (*N* = 15)	SG Raw-WPAS points (*N* = 13)	SL Raw-WPAS points (*N* = 15)
Age (years)	$\tau_b = 0.05$, $p = 1.000$	$\tau_b = 0.06$, $p = 1.000$	$\tau_b = 0.17$, $p = 1.000$	$\tau_b = 0.12$, $p = 1.000$	$\tau_b = 0.39$, $p = 0.611$
Age started skiing (years)	$\tau_b = 0.19$, $p = 1.000$	$\tau_b = 0.31$, $p = 1.000$	$\tau_b = 0.17$, $p = 1.000$	$\tau_b = 0.03$, $p = 1.000$	$\tau_b = 0.31$, $p = 1.000$
Age of onset of impairment (years)	$\tau_b = -0.06$, $p = 1.000$	$\tau_b = -0.03$, $p = 1.000$	$\tau_b = -0.17$, $p = 1.000$	$\tau_b = -0.28$, $p = 1.000$	$\tau_b = 0.08$, $p = 1.000$
Total hours of skiing	$\tau_b = -0.23$, $p = 1.000$	$\tau_b = -0.06$, $p = 1.000$	$\tau_b = -0.11$, $p = 1.000$	$\tau_b = -0.18$, $p = 1.000$	$\tau_b = 0.03$, $p = 1.000$
Number of races	$\tau_b = -0.37$, $p = 0.130$	$\tau_b = -0.12$, $p = 1.000$	$\tau_b = -0.19$, $p = 1.000$	$\tau_b = -0.46$, $p = 0.403$	$\tau_b = -0.17$, $p = 1.000$

The adjusted p-values based on the Bonferroni–Holm corrections are presented in the table with sample sizes. None of the values retained significance after applying the Bonferroni–Holm corrections.

so at least the results presented here can be compared to other Paralympic sports with vision impairments.

Finally, an ideal outcome measure for these studies would have been the raw race times of all participants on the same nordic or alpine course (depending on their discipline), under the same environmental and experimental conditions. The practicality issues involved in getting all these skiers from around the world to ski a single course on the same day made the ideal outcome measure impossible to achieve. However, choosing race times from a particular race instead of calculating the race performance points over a period of time as was done in this study may have increased the impact of confounding factors such as fatigue, poor weather, or anxiety on individual skiers performances. The point system used in this study ensured that the best performances of each elite skier in the validity period were considered for calculating the outcome measure, minimizing the effect of the abovementioned confounding factors. In addition, point calculations were based on the International Ski Federation (FIS) formula, which calculates race points relative to the race time of the overall best performer in each race, for each gender. Normalizing the performance points to the best performance in each gender allowed the researchers to compare performance data between genders. While additional tests to assess each skier's physiological and psychological factors, such as attention and physical fitness, which could affect skiing performance were considered, ultimately, they were not included. Visual motor reaction time assessments of attention have limited utility in a population with diverse vision impairments and stratifying the population by muscle power and/or flexibility would have made comparison between genders more difficult, thereby reducing the study sample size. However, quantifying these confounding factors would have increased the validity of this study and should be considered for inclusion in similar future studies (Tweedy et al., 2014).

CONCLUSION

In consideration of the correlation and regression analyses from these studies, static VA and VF were the only visual functions associated with both Para nordic and Para alpine skiing performance. Even though CS was associated with the performance in SG, GS, and SL disciplines and dynamic VA was associated with performance in DH, CS, and dynamic VA which were also strongly associated with static VA, CS, and dynamic VA do not appear to add any additional information for classification of skiers' performance in either Para nordic or Para alpine skiing. From a Paralympic classification research point of view, a test should only be incorporated into classification if its addition to the test battery improves the ability of a classification system to minimize the impact of impairments on the outcome of a competition (Mann and Ravensbergen, 2018). Thus, these studies concluded that static VA and VF should be included as visual functions in Para nordic and Para alpine classification. Further research needs to be done in order to determine if other visual functions should also be included in classification.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the data was collected in collaboration with the International Paralympic Committee and some aspects of the data set cannot be shared broadly. Please contact the corresponding author for more details if interested in the data. Requests to access the datasets should be directed to KD (kristine.dalton@uwaterloo.ca).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Office of Research Ethics, University of Waterloo. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

The studies were designed by KD, and the data were collected by KD, MC, and AS. AS organized the database and performed the statistical analysis and interpretation presented in this manuscript. AS wrote the manuscript and KD revised, read, and approved the submitted version. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by an Agitos Foundation grant in collaboration with World Para Snow Sport and the International Paralympic Committee.

ACKNOWLEDGMENTS

The authors would like to thank the International Paralympics Committee, World Para Nordic Skiing, and World Para Alpine Skiing for their support and funding throughout the study. The authors would also like to thank the Classification Research and Development Center for athletes with vision impairment at Vrije Universiteit Amsterdam, for providing the questionnaire used in the studies and their continuous support throughout the study. Our sincere thanks are given to Adaptive Sensory Technology and Dr. Michael Dorr (CTO, Adaptive Sensory Technology) for providing the test equipment and assistance with some aspects of data analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.648648/full#supplementary-material

REFERENCES

Bailey, I. L., Bullimore, M. A., Raasch, T. W., and Taylor, H. R. (1991). Clinical grading and the effects of scaling. *Invest. Ophthalmol. Vis. Sci.* 32, 422–432.

Beck, R. W., Diehl, L., and Cleary, P. A. (1993). The Pelli-Robson letter chart: normative data for young adults. *Clin. Vision Sci.* 8, 207–210.

Burg, A. (1966). Visual acuity as measured by dynamic and static tests: a comparative evaluation. *J. Appl. Psychol.* 50:460. doi: 10.1037/h0023982

Burg, A., and Hulbert, S. (1961). Dynamic visual acuity as related to age, sex, and static acuity. *J. Appl. Psychol.* 45:111. doi: 10.1037/h0044200

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Cambridge, MA: Academic press.

Craybiel, A., Jokl, E., and Trapp, C. (1955). Notes: Russian studies of vision in relation to physical activity and sports. *Res. Q. Am. Assoc. Health Phys. Educ. Recreation* 26, 480–485. doi: 10.1080/10671188.1955.10612840

Creese, M., Leat, S., Thompson, B., and Dalton, K. (2017a). Visual function of visually impaired Paralympic skiers. *Invest. Ophthalmol. Vis. Sci.* 58:4668.

Creese, M., Stalin, A., Leat, S., and Dalton, K. (2017b). *Comparing Visual Function of Visually Impaired Skiers in Different Lighting Conditions*, Ph.D. thesis. Toronto, ON: *VISTA*.

Creese, M., Stalin, A., Susan, L., Thompson, B., and Dalton, K. (2017c). *Visual Function of Para Nordic Skiers with Visual Impairment*. Ph.D. thesis. Toronto, ON: VISTA.

Croux, C., and Dehon, C. (2010). Influence functions of the Spearman and Kendall correlation measures. *Stat. Methods Appl.* 19, 497–515. doi: 10.1007/s10260-010-0142-z

Dalton, K., Ravensbergen, R., Roberts, J., Chakraborty, A., Leat, S., Thompson, B., et al. (2017). Global motion perception with low vision. *Invest. Ophthalmol. Vis. Sci.* 58:4693.

Dalton, K., Stalin, A., and Creese, M. (2019). Can vision functions predict Para alpine skiing performance? *Invest. Ophthalmol. Vis. Sci.* 60:1045.

Decroix, M., Wazir, M. R. W. N., Zeuwts, L., Deconinck, F. F. J. A., Lenoir, M., and Vansteenkiste, P. (2017). Expert–non-expert differences in visual behaviour during alpine Slalom skiing. *Hum. Mov. Sci.* 55, 229–239. doi: 10.1016/j.humov.2017.08.012

Dorr, M., Elze, T., Wang, H., Lu, Z.-L., Bex, P. J., and Lesmes, L. A. (2017a). New precision metrics for contrast sensitivity testing. *IEEE J. Biomed. Health Inform.* 22, 919–925. doi: 10.1109/jbhi.2017.2708745

Dorr, M., Lesmes, L. A., Elze, T., Wang, H., Lu, Z. L., and Bex, P. J. (2017b). Evaluation of the precision of contrast sensitivity function assessment on a tablet device. *Sci. Rep.* 7:46706. doi: 10.1038/srep46706

Erickson, G. (2007). *Sports Vision*. Amsterdam: Elsevier Inc, doi: 10.1016/B978-0-7506-7577-2.X5001-1

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G* power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/bf03193146

Ferrara, K., Burns, J., and Mills, H. (2015). Public attitudes toward people with intellectual disabilities after viewing olympic or paralympic performance. *Adapt. Phys. Activ. Q.* 32, 19–33. doi: 10.1123/apaq.2014-0136

Ferris, F. L., Kassoff, A., Bresnick, G. H., and Bailey, I. (1982). New visual acuity charts for clinical research. *Am. J. Ophthalmol.* 94, 91–96. doi: 10.1016/0002-9394(82)90197-0

Ferris, F. L. III., and Bailey, I. (1996). Standardizing the measurement of visual acuity for clinical research studies: guidelines from the eye care technology forum. *Ophthalmology* 103, 181–182. doi: 10.1016/S0161-6420(96)30742-2

Flørenes, T. W., Nordsletten, L., Heir, S., and Bahr, R. (2011). Recording injuries among world cup skiers and snowboarders: a methodological study. *Scand. J. Med. Sci. Sports* 21, 196–205. doi: 10.1111/j.1600-0838.2009.01048.x

Fontenot, J. L., Bona, M. D., Kaleem, M. A., McLaughlin, W. M., Morse, A. R., Schwartz, T. L., et al. (2018). Vision rehabilitation preferred practice pattern§. *Ophthalmology* 125, 228–278. doi: 10.1016/j.ophtha.2017.09.030

Gilgien, M., Reid, R., Raschner, C., Supej, M., and Holmberg, H.-C. (2018). The training of olympic alpine ski racers. *Front. Physiol.* 9:1772. doi: 10.3389/fphys.2018.01772

Gleeson, M., Sherrington, C., and Keay, L. (2014). Exercise and physical training improve physical function in older adults with visual impairments but their effect on falls is unclear: a systematic review. *J. Physiother.* 60, 130–135. doi: 10.1016/j.jphys.2014.06.010

Gold, J. R., and Gold, M. M. (2007). Access for all: the rise of the paralympic games. *J. R. Soc. Promot. Health* 127, 133–141. doi: 10.1177/1466424007077348

Grobbel, J., Dietzsch, J., Johnson, C. A., Vonthein, R., Stingl, K., Weleber, R. G., et al. (2016). Normal values for the full visual field, corrected for ageand reaction time, using semiautomated kinetic testing on the octopus 900 perimeter. *Transl. Vis. Sci. Technol.* 5:5. doi: 10.1167/tvst.5.2.5

Grosvendor, T. (2007). *Primary Care Optometry*. Amsterdam: Elsevier Inc, 3–41.

Haegerstrom-Portnoy, G., Schneck, M. E., Lott, L. A., and Brabyn, J. A. (2000). The relation between visual acuity and other spatial vision measures. *Optom. Vis. Sci.* 77, 653–662. doi: 10.1097/00006324-200012000-00012

Hawkins, D. M. (2004). The problem of overfitting. *J. Chem. Inform. Comput. Sci.* 44, 1–12. doi: 10.1021/ci0342472

Hirano, M., Hutchings, N., Simpson, T., and Dalton, K. (2017). Validity and repeatability of a novel dynamic visual acuity system. *Optom. Vis. Sci.* 94, 616–625. doi: 10.1097/OPX.0000000000001065

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Hou, F., Huang, C. B., Lesmes, L., Feng, L. X., Tao, L., Zhou, Y. F., et al. (2010). QCSF in clinical application: efficient characterization and classification of contrast sensitivity functions in Amblyopia. *Invest. Ophthalmol. Vis. Sci.* 51, 5365–5377. doi: 10.1167/iovs.10-5468

International Paralympic Committee (2019). *Introduction to Sports Classification*. Bonn: International Paralympic Committee.

Jia, W., Yan, F., Hou, F., Lu, Z.-L., and Huang, C.-B. (2014). QCSF in clinical applications: efficient characterization and classification of contrast sensitivity functions in aging. *ARVO Meet. Abstr.* 55:762.

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean J. Anesthesiol.* 72, 558–569. doi: 10.4097/kja.19087

Klein, R., Klein, B. E. K., Moss, S. E., and DeMets, D. (1983). Inter-observer variation in refraction and visual acuity measurement using a standardized protocol. *Ophthalmology* 90, 1357–1359. doi: 10.1016/S0161-6420(83)34382-7

Koolaee, A. K. (2017). Sport as an effective goal to increase self-concept and hope: a comparison study between athlete and non-athlete women with visual impairments. *Soc. Determ. Health* 3, 98–103. doi: 10.22037/sdh.v3i2.18486

Lesmes, L. A., and Dorr, M. (2019). "Active learning for visual acuity testing," in *Proceedings of the 2nd International Conference on Applications of Intelligent Systems*, (New York, NY: ACM), doi: 10.1145/3309772.3309798

Lesmes, L. A., Wallis, J., Lu, Z.-L., Jackson, M. L., and Bex, P. (2012). Clinical application of a novel contrast sensitivity test to a low vision population: the quick CSF method. *Invest. Ophthalmol. Vis. Sci.* 53:4358.

Mann, D. L., and Ravensbergen, H. J. C. (2018). International Paralympic Committee (IPC) and International Blind Sports Federation (IBSA) joint position stand on the sport-specific classification of athletes with vision impairment. *Sports Med.* 8, 2011–2023. doi: 10.1007/s40279-018-0949-6

Mann, D. L., and Ravensbergen, R. H. J. C. (2019). *Protocol for AMA-Style Analysis of Visual Field*. Amsterdam.

Ohlsson, J., and Villarreal, G. (2005). Normal visual acuity in 17-18 Year Olds. *Acta Ophthalmol. Scand.* 83, 487–491. doi: 10.1111/j.1600-0420.2005.00516.x

Olive, D. J. (2017). *Linear Regression*. Cham: Springer, doi: 10.1007/978-3-319-55252-1

Rosser, D. A., Cousens, S. N. I, Murdoch, E., Fitzke, F. W., and Laidlaw, D. A. H. (2003). How sensitive to clinical change are ETDRS LogMAR visual acuity measurements? *Invest. Ophthalmol. Vis. Sci.* 44, 3278–3281. doi: 10.1167/iovs.02-1100

Schläppi, O., Urfer, J., and Kredel, R. (2016). Visual perception in alpine ski racing: aqualitative analysis based on interviews with top-level athletes. *Sportwissenschaft* 46, 201–212. doi: 10.1007/s12662-016-0400-9

Senner, V., Jendrusch, G., Schaff, P., and Heck, H. (1999). Vision–an essential factor for safety in Skiing: perception, reaction, and motion control aspects. *ASTM Spec. Tech. Publication* 1345, 11–22. doi: 10.1520/stp12353s

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika* 73, 751–754. doi: 10.1093/biomet/73.3.751

Stalin, A. (2020). *Development of Sports-Specific Classification for Paralympic Skiers with Visual Impairment*. Ph.D. thesis. Waterloo, ON: UWSpace.

Stalin, A., Creese, M., and Dalton, K. (2019). Do vision functions predict Para Nordic Skiing performance? *Invest. Ophthalmol. Vis. Sci.* 60:1044.

Stellmann, J. P., Young, K. L., Pöttgen, J., Dorr, M., and Heesen, C. (2015). Introducing a new method to assess vision: computer-adaptive contrast-sensitivity testing predicts visual functioning better than charts in multiple sclerosis patients. *Mult Scler J. Exp. Transl. Clin.* 1:205521731559618. doi: 10.1177/2055217315596184

Tweedy, S. M., Beckman, E. M., and Connick, M. J. (2014). Paralympic classification: conceptual basis, current methods, and research update. *PM and R* 6(8 Suppl), S11–S17. doi: 10.1016/j.pmrj.2014.04.013

Vanden Bosch, M. E., and Wall, M. (1997). Visual acuity scored by the letter-by-letter or probit methods has lower retest variability than the line assignment method. *Eye* 11, 411–417. doi: 10.1038/eye.1997.87

Wilcox, R. R. (2016). *Introduction to Robust Estimation and Hypothesis Testing*, 4th Edn. Amsterdam: Elsevier.

World Para Alpine Skiing (2019). *Rules and Regulations*. Bonn: World Para Alpine Skiing.

World Para Nordic Skiing (2018). *Rules and Regulations*. Bonn: World Para Nordic Skiing.

World Para Nordic Skiing (2019). *World Para Nordic Skiing Percentages*. Bonn: World Para Nordic Skiing.

Xu, P., Lesmes, L. A., Yu, D., and Lu, Z. L. (2019). A novel Bayesian adaptive method for mapping the visual field. *J. Vis.* 19:16. doi: 10.1167/19.14.16

Xu, P., Lesmes, L. A., Yu, D., and Lu, Z. L. (2020). Mapping the contrast sensitivity of the visual field with Bayesian adaptive QVFM. *Front. Neurosci.* 14:665. doi: 10.3389/fnins.2020.00665

Zhao, Y., Lesmes, L. A., Dorr, M., Bex, P. J., and Lu, Z. L. (2021). Psychophysical validation of a novel active learning approach for measuring the visual acuity behavioral function. *Transl. Vis. Sci. Technol.* 10:1. doi: 10.1167/tvst.10.1.1

Conflict of Interest: MC is currently employed by the International Paralympic Committee, and any views/conclusions expressed in the manuscript are her own and do not represent the views/conclusions of the International Paralympic Committee.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 25 May 2021
doi: 10.3389/fnins.2021.667329

Influence of Lenslet Configuration on Short-Term Visual Performance in Myopia Control Spectacle Lenses

Xue Li[1,2,3†], Chenglu Ding[1,2†], Yuhao Li[1,2], Ee Woon Lim[3,4], Yi Gao[3,4], Bruno Fermigier[5], Adeline Yang[3,4], Hao Chen[1,2,3] and Jinhua Bao[1,2,3*]*

[1] Eye Hospital and School of Ophthalmology and Optometry, Wenzhou Medical University, Wenzhou, China, [2] National Clinical Research Center for Ocular Diseases, Wenzhou, China, [3] Wenzhou Medical University–Essilor International Research Center (WEIRC), Wenzhou, China, [4] R&D AMERA, Essilor International, Singapore, Singapore, [5] R&D Essilor International, Créteil, France

Purpose: This study aimed to evaluate short-term visual performance and optical quality of three different lenslet configurations on myopia control spectacle lenses.

Materials and Methods: This study utilized a cross-over design. Distance visual acuity (VA) was measured in 50 myopic children; contrast sensitivity (CS) was measured in 36 myopic children. For each test, four spectacle lenses were evaluated in a random order: single-vision lens (SVL), lens with concentric rings of highly aspherical lenslets (HAL), lens with concentric rings of slightly aspherical lenslets (SAL), and lens with honeycomb configuration of spherical lenslets (HC). The modulation transfer function (MTF) and MTF area (MTFa) were used to determine optical quality. All tests were performed monocularly on the right eye with full correction.

Results: HAL and SAL had larger MTFa than HC. VA in lenses with lenslets was significantly reduced compared to SVL (all $p < 0.01$). The reduction in VA was worse with HC than with SAL ($p = 0.02$) and HAL ($p = 0.03$); no effect of lenslet asphericity was found ($p > 0.05$). VA changes induced by lenslets showed no correlation with spherical equivalent refraction (all $p > 0.05$) and were weakly positively associated with age for SAL ($r = 0.36$, $p = 0.01$) and HC ($r = 0.31$, $p = 0.03$), but not for HAL ($p = 0.30$). The area under the log contrast sensitivity function (AULCSF) decreased with HAL and HC (all $p < 0.001$) in all illumination levels, and AULCSF with HAL was higher than that with HC in a photopic condition (1.17 ± 0.10 vs. 1.10 ± 0.13, $p = 0.0004$). The presence of lenslets did not affect CS at 3 cycles per degree (cpd) ($p = 0.80$). At 6 to 18 cpd, CS was significantly reduced by HAL and HC (all $p < 0.05$), but not SAL ($p > 0.05$) compared to SVL. At high spatial frequencies (>12 cpd) both SAL and HAL reduced CS significantly less than HC (all $p < 0.01$).

Conclusion: Short-term visual performance was minimally impaired by looking through the lenslet structure of myopia control spectacle lenses. Concentric rings with aspherical lenslets had a significantly lower impact on both VA and CS than honeycomb configuration with spherical lenslets.

Keywords: myopia control, visual performance, visual acuity, contrast sensitivity, modulation transfer function, lenslets

Edited by:
Peter J. Bex, Northeastern University, United States

Reviewed by:
Bin Zhang, Nova Southeastern University, United States
Pablo De Gracia, Midwestern University, United States

****Correspondence:***
Jinhua Bao
baojessie@163.com
Hao Chen
chenhao@mail.eye.ac.cn

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *12 February 2021*
Accepted: *08 April 2021*
Published: *25 May 2021*

Citation:
Li X, Ding C, Li Y, Lim EW, Gao Y, Fermigier B, Yang A, Chen H and Bao J (2021) Influence of Lenslet Configuration on Short-Term Visual Performance in Myopia Control Spectacle Lenses. Front. Neurosci. 15:667329. doi: 10.3389/fnins.2021.667329

INTRODUCTION

The prevalence of myopia is predicted to be 50% globally by the year 2050, with 10% being highly myopic (Fricke et al., 2018). This growing epidemic is a concern as the risk for myopia-related pathology is as high as 28.7% in the highly myopic population (Wong et al., 2018). Moreover, these pathologies can lead to vision impairment and heavy economic burdens (Zheng et al., 2013). As such, it is of public health interest to control myopia progression through efficient interventions (Wildsoet et al., 2019).

There are several optical interventions such as orthokeratology, bifocal spectacles, and multifocal contact lenses (Huang et al., 2016; Wildsoet et al., 2019) available in the clinic for myopia control. Recently, spectacle lens designs using lenslets to create a myopia control signal in the periphery, for example, the Defocus Incorporated Multiple Segments (DIMS) (Lam et al., 2019), spectacle lenses with slightly aspherical lenslets (SAL), and spectacle lenses with highly aspherical lenslets (HAL), have shown a promising myopia control effect (Bao et al., 2021). The efficacy of these spectacle lenses was comparable to orthokeratology (Li et al., 2016; Santodomingo-Rubido et al., 2017) and 0.01% atropine (Diaz-Llopis and Pinazo-Duran, 2018; Kinoshita et al., 2018). Moreover, spectacle lenses are non-invasive and safer than contact lenses or drugs.

However, spectacle lenses with lenslets face similar visual performance issues like multifocal contact lenses used for myopia control, especially in the peripheral part of the visual field. Lenses designed for myopia control were found to affect low-contrast visual acuity under low illuminance, while distant high-contrast vision acuity was rarely affected (Kang et al., 2017; Diec et al., 2018; García-Marqués et al., 2020; Lu et al., 2020). Lu et al. (2020) found that the DIMS showed no effect on visual acuity (VA) through the central clear zone but reduced VA by three optotypes in the defocus area with lenslets. Pauling et al., found that the multifocal soft contact lenses affected the low- and high-contrast VA on initial insertion and advocated that the effects on vision should be communicated when dispensing these lenses (Kang et al., 2017). However, Jennie et al. stated that visual acuity did not adequately reflect visual performance for multifocal contact lens. Contrast sensitivity (CS), in contrast, is a more sensitive measure, especially when the lens was significantly decentered (Fedtke et al., 2016). One study found worse visual performance with a higher addition power lens (Przekoracka et al., 2020), while another found no difference (Walline et al., 2020).

In normal, straight viewing conditions, children using spectacle lenses with lenslets in the lens periphery will look through the central clear zone, which has been shown to have no impact on VA (Lu et al., 2020). However, eye movements and possible position shifts of the spectacle frame make it possible for the visual axis to pass through the peripheral zone with lenslets. Thus, it is necessary to evaluate the visual performance through the lenslet zone to understand the impact of the lenslets to provide guidance for clinical practice. This study aimed to evaluate the optical quality and visual performance through various lenslet configurations and compare them with single-vision lenses (SVL) in children. VA and CS were used to evaluate the visual quality subjectively, and modulation transfer function (MTF) was used to estimate the optical property.

MATERIALS AND METHODS

Subjects

This was a cross-over design study. For the VA test, 50 myopic children [mean age 12.7 ± 1.7 years, age range 10 to 15 years, mean spherical equivalent refraction (SER) −3.22 ± 1.57 D, SER range −6.50 to −0.38 D] participated; for the CS test, 36 myopic children (mean age 13.2 ± 1.2 years, age range 10 to 16 years, mean SER −3.20 ± 1.67 D, SER range −7.25 to −0.75 D) were enrolled. Subjects had no ocular pathology or former history of using myopia control interventions. During the experiment, each subject was fully corrected using a trial frame. Testing lenses with lenslets were added to the right eye while the left eye was occluded. All tests were performed immediately after fitting the lenses without any adaptation. This study adhered to the tenets of the Declaration of Helsinki and was approved by the Ethics Committee of the Eye Hospital of Wenzhou Medical University (no. 2019-091-K-87). Written informed consent was obtained from both children and their legal custodian before the study.

Apparatus

All spectacle lenses were made of polycarbonate in this study. There were four designs: (1) traditional single-vision lens (SVL) as control, (2) concentric ring configuration with highly aspherical lenslets (HAL) (**Figure 1**, left), (3) concentric ring configuration with slightly aspherical lenslets (SAL), and (4) honeycomb configuration of spherical lenslets (HC) (**Figure 1**, right). For HAL and SAL, the surface of the lens without lenslets provides distance correction. The geometry of the aspheric lenslets (1.12 mm in diameter) was calculated to generate a volume of myopic defocus ranging from 1.1 to 1.9 mm (HAL) and from 1.0 to 1.3 mm (SAL) in front of the retina at any eccentricity, serving as a myopia control signal. The lenslets (1.03 mm in diameter) of HC introduce myopic defocus at a plane in front of the retina by a relative positive power (+3.50 D) (Lu et al., 2020; Zhang et al., 2020). The surface of the lens without lenslets provides distance correction. The lenslets of two configurations, concentric rings and honeycomb, provide a similar density of lenslets that was approximately 40% of the total surface area of each lens.

Each of the three lenses with lenslets was mounted into a trial lens ring to maximize the lenslet zone, in which the central clear zone was on the edge of the cut lens (**Figure 2**). To ensure viewing only through the lenslets zone, the 9 mm of central clear zone (the small black circle) and the area beyond a distance of 12 mm from the central zone (the black crescent-shaped area) were patched by non-light-permeable tapes. The SVL was edged and covered up in the same way to ensure the same size and shape of the visual field among the lenses. During the experiments, the subjects wore a trial frame and performed the visual tests by looking through the lenslet zone. Four types of lenses were tested in random order.

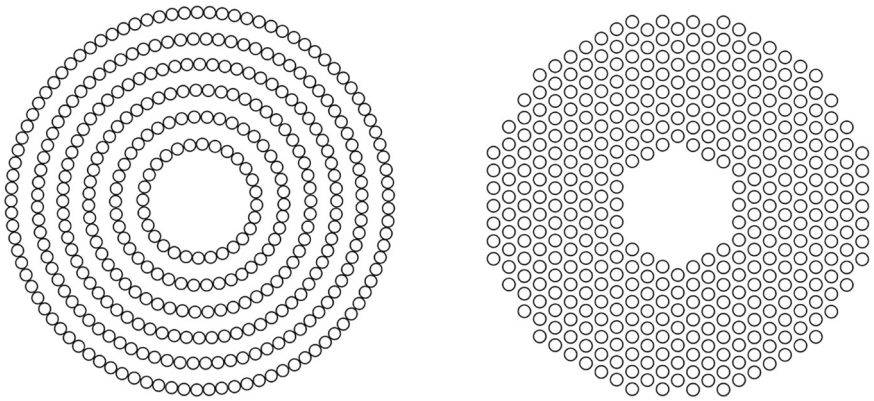


FIGURE 1 | Pictorial representation of concentric rings (left) and honeycomb (right) configurations of lenslets.

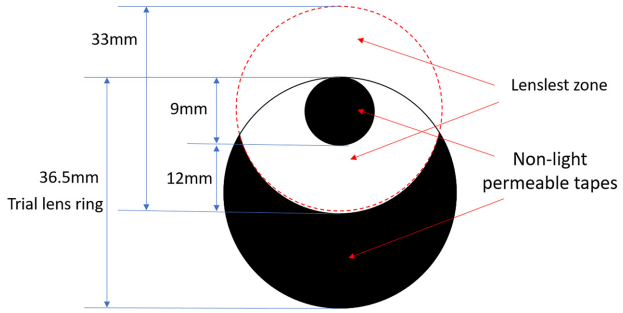


FIGURE 2 | Pictorial representation of patched lenses. The large solid circle of 36.5-mm diameter represents the actual edged lens (trial lens). The small black circle of 9-mm diameter represents the patched central clear zone of the original lens, and the black crescent-shaped area represents the patched peripheral clear zone of the original lens. All four lenses were patched in identical ways.

Modulation Transfer Function

Modulation transfer function (Pieh et al., 2002; Son et al., 2017) has been widely used to quantify the optical quality of the lens design, and MTFa (modulation transfer function area) can be used to predict the VA and CS outcome (Fernández et al., 2019; Armengol et al., 2020). In summary, to calculate MTF, one evaluates the complex amplitude in the pupil plane, then using fast Fourier transform (FFT) calculates the point spread function (PSF) and finally the MTF, using one center wavelength (λ = 550 nm) and assuming the pupil position directly on the glass (Voelz, 2011). The MTFa of optical simulation of three lens designs with lenslets was calculated within the spatial frequency range of 0–15 cycles per degree (cpd) (Vega et al., 2018; Jaskulski et al., 2020) on 4, 6, and 8 mm apertures by a 550-nm light source. MTFs were computed for the same pupil apertures at 5, 10, and 15 cpd.

Contrast Sensitivity Function

The contrast sensitivity (CS) and glare disability with the test lenses were measured with CSV-1000 (Vector Vision Carp, United States; **Figure 3**). The test was performed at a distance of 2.5 m in a dark room; the translucent chart presented four spatial frequencies: 3, 6, 12, and 18 cpd, with contrast levels reduced in steps corresponding to 0.15 logCS. The testing illuminance levels from the light box included photopic (85 cd/m^2) and mesopic (3 cd/m^2) conditions with and without glare (Pomerance and Evans, 1994). The area under the log contrast sensitivity function (AULCSF) was calculated by summing the area under the CSF obtained from the data measured (Applegate et al., 1998) in each condition.

Before the test, two practice trials were implemented to eliminate the effect of familiarity. Then, subjects adapted to each illuminance level for 5 min before testing. The four testing lenses were applied in a random order with a short interval for approximately 1 min. The total testing lasted approximately 1 h.

Visual Acuity

Visual acuity was evaluated using the Freiburg Vision Test (FrACT) (Bach, 1996, 2006). Compared with the Snellen VA chart, the computerized and automated FrACT tool is free of examiner's bias (Ma et al., 2013). A single Landolt C represented the stimulus with the opening at one of eight cardinal directions enclosed in a crowding square on a Mac screen of 21.5-in screen dimension and 1920 × 1080 resolution. The average screen luminance was 75 cd/m^2 (**Figure 4**). An eight-alternative forced-choice paradigm (8-AFC) was used, in which the task was to determine the opening direction of the Landolt C among the eight possible cardinal directions (four cardinal directions and four oblique directions).

During the experiment, the testing distance was 3 m, and the illumination at the eye plane was 200 lux. Before the test, two practice trials were implemented to eliminate the effect of familiarity. Then, each of the four lenses was imposed on the right eye of the subject in a random order for testing, with a short break of 1 min in between. The total testing time was within 30 min. Landolt Cs were presented at 100% contrast, and the measurement procedure was described in detail in previous literature (Bach, 2007; Bach and Schäfer, 2016).

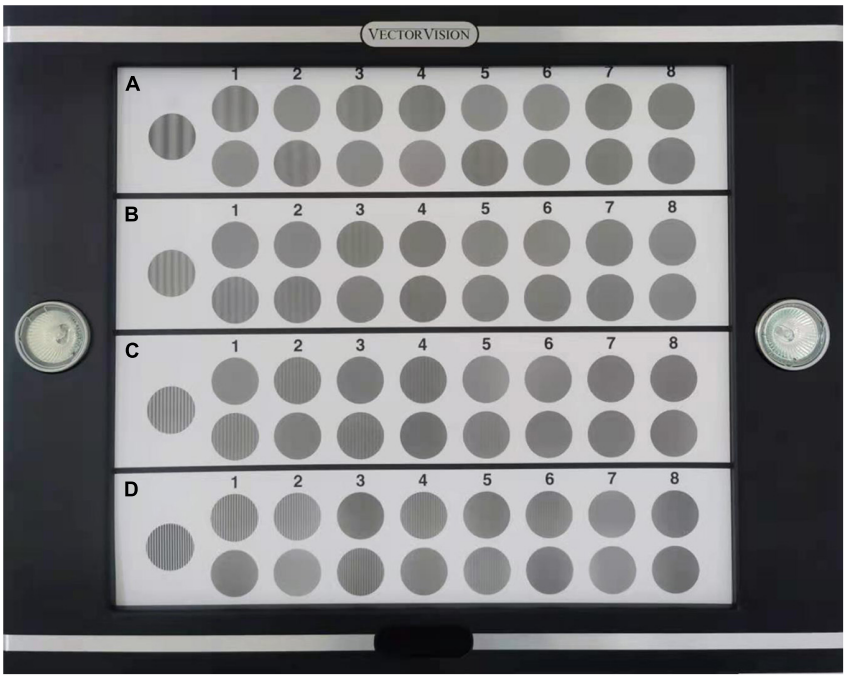


FIGURE 3 | Test card of CSV-1000 for distant contrast sensitivity (CS). **(A-D)** represent four spatial frequencies, from low to high, 3, 6, 12, 18 cycles per degree (cpd).

FIGURE 4 | An example of the stimuli (FrACT test with high contrast, 100%).

Statistical Analysis

Statistical analysis was performed using the SPSS (version 25.0, SPSS, Inc.) software. Repeated-measures ANOVAs were used to test intergroup differences, if significant, followed by *post hoc* Bonferroni tests for pairwise comparisons. The statistical significance threshold was set at $p < 0.05$.

RESULTS

Modulation Transfer Function

A quantitative analysis of the optical performance of honeycomb and concentric ring configurations was performed using MTF and MTFa simulation through aperture sizes 4, 6, and 8 mm (**Figure 5**).

The MTF curves of three lenses showed similar patterns for all pupil apertures, revealing a decrease in image modulation from 0 to 20 cpd. The effect of the spherical lenslets in the honeycomb configuration was similar to that of aspherical lenslets in concentric ring configuration at low spatial frequencies (< 5 cpd). Between 5 and 35 cpd, HC decreased image modulation compared with HAL and SAL, then sharply increased it approximately 47 cpd (**Figure 5A**).

MTFa of honeycomb configuration was less than that of the concentric ring configurations (**Figure 5B**), indicating that the lenslets of concentric ring configurations would provide better optical performance.

Contrast Sensitivity Function

The mean AULCSF and CS values of the four tested lenses across subjects in different illuminance conditions are shown in **Figure 6**.

Repeated-measures ANOVAs found significant effects of lens in each condition (all $p < 0.05$). Further *post hoc* Bonferroni tests showed a pairwise difference or not between the four testing lenses. The difference between SVL and the three lenses with lenslets indicated the impact of lenslets on CS. At the low

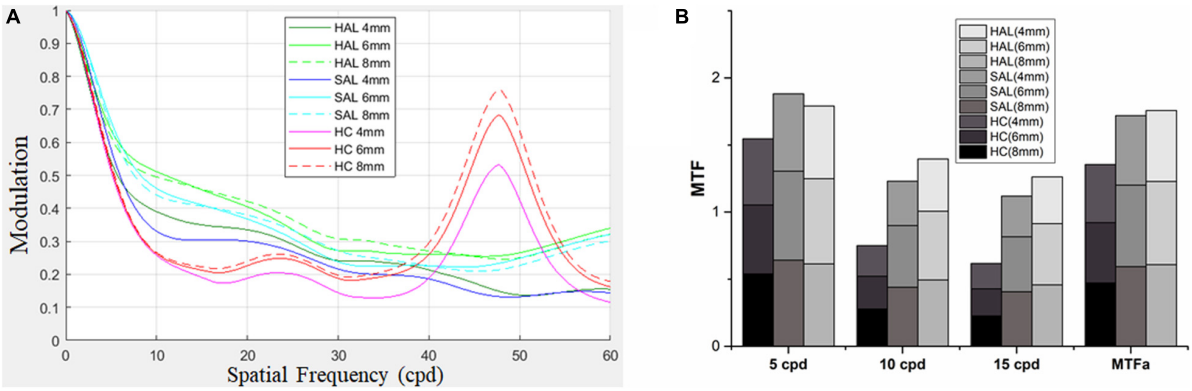


FIGURE 5 | For lenses of HAL, SAL, and HC, MTFs were computed for three pupil apertures of 4, 6, and 8 mm **(A)**. MTFs at 5, 10, and 15 cpd, and MTFa from 0 up to 15 cpd at the three pupil apertures are shown in **(B)**. HAL, highly aspherical lenslets; SAL, slightly aspherical lenslets; HC, honeycomb configuration of spherical lenslets; MTF, modulation transfer function; MTFa, MTF area; cpd, cycle per degree.

spatial frequency of 3 cpd, CS was not significantly affected by lens configurations except in the photopic condition with glare (**Figure 6B**), where HAL reduced CS compared to SVL. At the mid spatial frequency of 6 cpd, SAL did not significantly affect CS compared to SVL in any illuminance conditions, while HAL and HC reduced CS significantly; there was no significant difference between all the three lenslet configurations. At high spatial frequencies (12 and 18 cpd), SAL only reduced CS in the mesopic conditions (**Figures 6C,D**), while HAL and HC reduced CS significantly compared to SVL in most conditions. It is worth noting that HAL did not reduce CS significantly at the very high spatial frequency (SF) of 18 cpd under the photopic conditions, whereas HC did (**Figures 6A,B**).

Comparisons between the three lenses with lenslets found that the two concentric ring configurations HAL and SAL generated a significantly smaller impact on CS than HC at high SFs in most conditions. In the mesopic condition (**Figure 6C**), in contrast to the photopic condition (**Figure 6A**), CS at high spatial frequencies were generally reduced, and the difference between HAL and HC became less significant while SAL still showed significantly higher CS than HC. Adding glare did not reduce the general CS levels as low illuminance did, but caused the difference between the lenslet configurations at high spatial frequencies to become less significant in the photopic condition. HAL and SAL showed no significant difference in CS under any illuminance condition.

Both HAL and HC resulted in significantly lower AULCSF than SVL in all illuminance conditions (**Figures 6E–H**), with and without glare (all $p < 0.001$). SAL did not cause any significant change in AULCSF compared to SVL (all $p > 0.05$). Comparisons between the spectacle lenses with lenslets revealed that AULCSF of HAL was significantly higher than that of HC in the photopic condition (1.17 ± 0.10 vs. 1.10 ± 0.13, $p = 0.0004$, **Figure 6E**), but not in other illuminance conditions (**Figures 6F–H**).

Visual Acuity

The mean VA through four lenses was 0.07 ± 0.09 logMAR (SVL), 0.15 ± 0.10 logMAR (HAL), 0.13 ± 0.09 logMAR (SAL), and 0.17 ± 0.09 logMAR (HC), respectively (**Figure 7A**). Repeated-measures one-way ANOVA found a significant effect of lens design on VA [$F_{(2.8,134.9)} = 23.52$, $p < 0.001$]. *Post hoc* Bonferroni tests showed that, compared with SVL, VA in lenses with lenslets significantly decreased (all $p < 0.001$). VA through SAL was significantly higher than through HC ($p = 0.004$).

The reduction in VA caused by lenslets relative to SVL was 0.07 ± 0.09 for HAL, 0.06 ± 0.09 for SAL, and 0.09 ± 0.07 logMAR for HC, respectively. The drop in VA caused by aspherical lenslets in concentric rings was significantly less than that caused by spherical lenslets in honeycomb configuration (all $p < 0.05$). No significant difference was found between the two lenses with aspherical lenslets ($p > 0.99$, **Figure 7B**).

Correlation Between VA and CS Changes and Refractive Errors and Age

To test whether any individual factors influenced the relative reduction in VA and CS compared to SVL caused by lenslets, a correlation analysis was performed on changes in VA and CS of three lenses with lenslets and factors including the spherical equivalent refraction (SER) and age of subjects. In **Figure 8**, VA changes were plotted as a function of SER (**Figure 8A**) and age (**Figure 8B**) of subjects.

CS changes of all three lenses (HAL, SAL, and HC) in all spatial frequencies were not significantly correlated with age or SER in each illumination condition (all $p > 0.05$).

VA changes of HAL, SAL, and HC from SVL were not significantly correlated with SER (all $p > 0.05$, **Figure 8A**). Age was positively correlated with VA loss in SAL ($r = 0.36$, $y = -0.18 + 0.02 \times x$, $p = 0.01$) and HC ($r = 0.31$, $y = -0.07 + 0.01 \times x$, $p = 0.03$), but not in HAL ($r = 0.16$, $P = 0.27$) (**Figure 8B**).

DISCUSSION

This study presented the optical quality through simulation and short-term visual performance through clinical testing using three configurations of lenslets (HAL, SAL, and HC) on spectacle

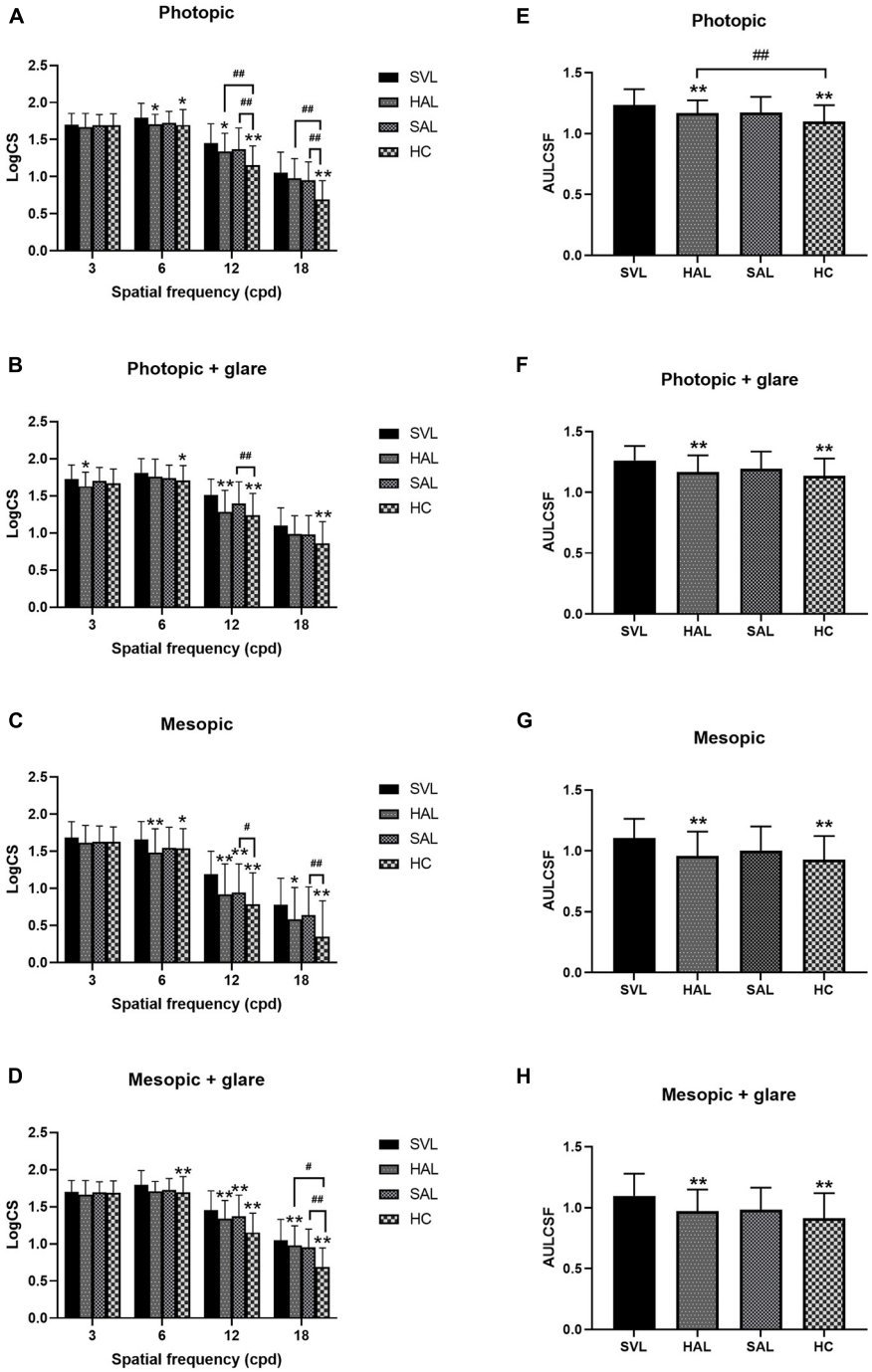


FIGURE 6 | Mean log contrast sensitivity **(A–D)** and area under the log contrast sensitivity function (AULCSF) **(E–H)** with standard deviations of four tested spectacle lenses in photopic **(A,B,E,F)** and mesopic **(C,D,G,H)** conditions, with **(B,D,F,H)** and without **(A,C,E,G)** glare. (SVL for single-vision lens, HAL for spectacle lenses with concentric rings of highly aspherical lenslets, SAL for spectacle lenses with concentric rings of slightly aspherical lenslets, and HC for lenses with spherical lenslets in honeycomb configuration) $N = 36$. Asterisk (*) and number sign (#) represent significance in the Bonferroni *post hoc* test following the repeated measures ANOVA. $*p < 0.05$, $**p < 0.01$, data compared with SVL; $\#p < 0.05$, $\#\#p < 0.01$, data compared between pairs of spectacle lenses with lenslets.

lenses that were designed for myopia control and compared them with single-vision lenses in 10–16 years old children.

We found that optical simulation could be used to predict visual performance with a spectacle lens design, and the pupil size affected the outcome. Ravikumar et al. also found that the change in VA was highly correlated with the change of MTF (Ravikumar et al., 2012). Studies on multifocal contact lenses (Kawamorita and Uozato, 2005; Madrid-Costa et al., 2012;

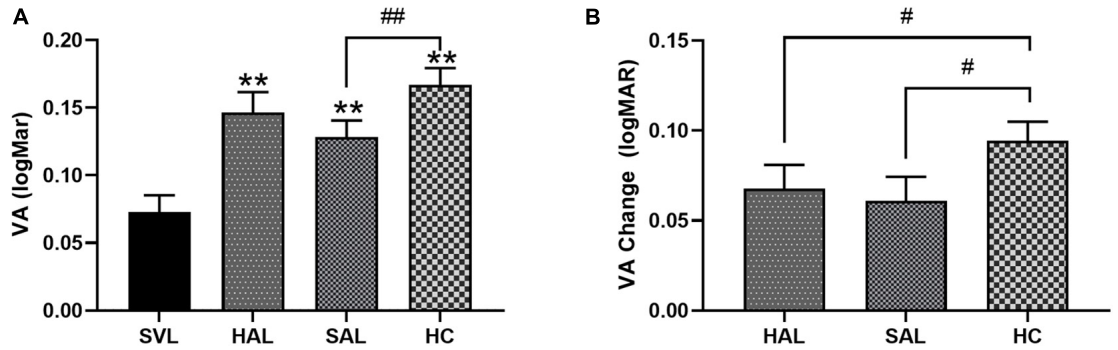


FIGURE 7 | (A) Mean visual acuity (VA) with standard errors of four lenses (SVL for single-vision lens, HAL for spectacle lenses with concentric rings of highly aspherical lenslets, SAL for spectacle lenses with concentric rings of slightly aspherical lenslets, and HC for lenses with spherical lenslets in honeycomb configuration) and **(B)** relative VA changes from SVL of three lenses with lenslets in logMAR unit. $N = 50$. Asterisk (*) and number sign (#) represent significance in the Bonferroni *post hoc* test following the repeated measures ANOVA. $^{**}p < 0.01$, comparisons of each of the three lenses with lenslets to SVL; $\#p < 0.05$, $\#\#p < 0.01$, comparisons between each pair of the three lenses with lenslets.

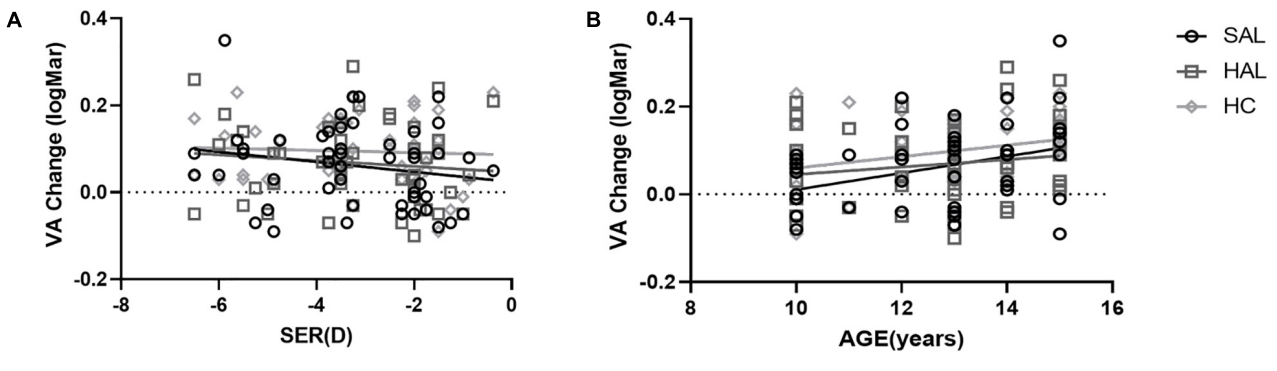


FIGURE 8 | Plots of the relative changes in visual acuity (logMAR) of HAL, SAL, and HC compared to SVL as a function of SER **(A)** and age **(B)** for all subjects ($N = 50$). The lines correspond to linear regressions. Age was significantly correlated to VA changes for SAL and HC. All other correlations were not significant.

Fernández et al., 2019) found the similar results. MTFa was also used to predict VA and CS of multifocal intraocular lenses (Vega et al., 2018; Armengol et al., 2020). In the current study, MTFs and MTFa showed that a concentric ring design impacted less visual performance than the honeycomb design. By testing VA and CS through the lenslet zones of the three spectacle lenses, we confirmed that the visual performance was aligned with the outcome of optical simulation. Moreover the CS under photopic condition was higher than that in mesopic illumination. Higher light levels induce smaller pupils by increasing the depth of focus and minimize the effects of higher-order aberrations by reducing the size of the blurred circle on the retina (Holladay et al., 1991), resulting in an increase of VA (Lombardo and Lombardo, 2010) and improved discrimination of fine stimuli (Xu et al., 2017; Mathôt and Ivanov, 2019).

Contrast sensitivity at high spatial frequencies, which reflects the ability to see fine details, was reduced by all three configurations of lenslets. The loss in CS caused by HC was significantly higher than that caused by both HAL and SAL. At both photopic and mesopic conditions, HC reduced CS at high spatial frequencies significantly more than HAL or SAL. Adding glare did not reduce the general CS as low illuminance did (Hohberger et al., 2007). Glare reduced the difference between the lenslet configuration at high spatial frequencies. CS at low spatial frequencies was not affected significantly by lenslets. The concentric rings of lenslet configuration provided better visual performance than the honeycomb configuration. Other than the configuration of lenslets, which resulted in less fragmented optics due to small aperture, the diameter of lenslets was also a factor impacting optical performance. HAL and SAL had a slightly larger lenslet diameter than HC (1.03 vs. 1.12 mm), which reduced diffraction caused by fragmented optics due to smaller aperture (Jaskulski et al., 2020).

The fact that lenslet design affects CS at high spatial frequencies suggests that lenslets also impact VA, which should be worst in HC according to MTF simulation. The results of VA tested using FrACT in the current study were consistent with the findings of CS. Jaskulski et al. also found that the DIMS decreased contrast sensitivity at high spatial frequencies (Jaskulski et al., 2020). For VA in high luminance and high contrast, HC induced the most vision loss by approximately 0.09 logMAR, followed by HAL and SAL. The lack of difference between HAL and SAL on VA in any condition indicated that the magnitude of the asphericity of the lenslets has little effect

on visual performance. The loss in VA caused by aspherical lenslets in concentric rings was about half a line on a typical VA chart. However, the VA loss caused by the spherical lenslets in the honeycomb configuration was about one whole line on the VA chart. Note that the losses in CS and VA found in the current study was obtained by testing central vision through the lenslet zones. Normally, the lenslets should be located in the periphery, and central vision should be aligned with the central clear zone. Studies have found that VA was not affected when looking through the central clear zone (Lam et al., 2019; Zhang et al., 2020).

VA loss while looking through the lenslet structures was likely caused by less light focusing on the retina (Fedtke et al., 2016), similar to the simulations of MTFa with smaller aperture sizes. However, the VA changes had a positive, weak, but significant correlation with the age of subjects in SAL and HC, but not in HAL. No correlation was found between VA change and refractive error. In other words, lenslets in SAL and HC have a larger impact on visual quality in older children, but not in HAL. Although only short-term visual performance was tested in the current study, the correlation with age suggests that younger children may have an easier or faster adaptation to the lenses, which could compensate for the optical disturbance induced by lenslets while looking through the peripheral parts of the lenses. Better adaptation of blur and acceptance of the lenses were also found in younger children wearing DIMS (Lu et al., 2020) and orthokeratology lenses (Chang and Cheng, 2019) compared to their older counterparts.

Note that only short-term effect of lenslets on visual performance was tested in the current study. Any changes in VA and CS found were immediate effects without adaptation. The impact of lenses on vision often diminishes after an adaptation period. For example, multifocal soft contact lenses designed for myopia control were found to induce reduction in high-contrast VA immediately after fitting, which subsequently recovered after 2 weeks (Kang et al., 2017) or significantly improved by over 0.10 logMAR after 8 days of adaptation (Fedtke et al., 2016). However, the impact of VA may not completely disappear as was found after an adaptation period of 1 week wearing the DIMS lenses (Lu et al., 2020). However in that study, the VA through lenslets was measured with rotating eyes to different angles, rather than looking straight forward straight as in the current study. Therefore, the small impact on VA and CS on children found in the current study is likely to reduce, but persist following adaptation.

The real-life implications of slight VA and CS losses on a child's vision are minimal. First, the measurements were performed through the lenslet zones. In the normal way of wearing spectacle lenses with lenslets, wearers look through the central clear zone that covers the visual field from zero to approximately 18° of eccentricity. The amount of time spent looking through the central clear zone will be is significantly larger than that spent in the lenslet zones. Second, the short-term loss of VA in SAL and HAL was merely approximately 0.06–0.07 logMAR, and 0.09 logMAR in HC, which were not considered clinically significant since the 95% confidence interval of repeatability of VA tests was found to be approximately 0.10 logMAR (Raasch et al., 1998) or 0.15 logMAR (Siderov and Tiu, 1999). Third, we tested only central visual performance. The impact of spectacle lenses with lenslets on the peripheral vision and performance on daily tasks in children's life, such as reading and writing, needs further investigation.

In summary, short-term testing results on visual performance were consistent with the simulation findings. Lenslets reduced short-term visual performance manifested in lower VA and contrast sensitivity at high spatial frequencies compared with SVL. The impact varied with the characteristics and configuration of the lenslets. Spherical lenslets in the honeycomb configuration induced larger loss in VA and CS than aspherical lenslets in concentric rings. However, the level of asphericity of the lenslets showed no significant effect on visual performance. The positive correlation between the impact on VA and the subjects' age for SAL and HC suggests better adaptation in younger children.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Eye Hospital of Wenzhou Medical University (no. 2019-091-K-87). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

All authors listed have made substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by the Leading Science and Technology Innovation Talent of Zhejiang Provincial Ten Thousand Talent Project (Grant No. 2017R52050) and the collaborative research project with Essilor International (Wenzhou Medical University Grant Numbers 95013006 and 95016010).

REFERENCES

Applegate, R. A., Howland, H. C., Sharp, R. P., Cottingham, A. J., and Yee, R. W. (1998). Corneal aberrations and visual performance after radial keratotomy. *J. Refract. Surg.* 14, 397–407. doi: 10.3928/1081-597x-19980701-05

Armengol, J., Garzón, N., Vega, F., Altemir, I., and Millán, M. S. (2020). Equivalence of two optical quality metrics to predict the visual acuity of

multifocal pseudophakic patients. *Biomed. Opt. Express* 11, 2818–2829. doi: 10.1364/boe.388531

Bach, M. (1996). The Freiburg Visual Acuity test–automatic measurement of visual acuity. *Optom. Vis. Sci.* 73, 49–53. doi: 10.1097/00006324-199601000-00008

Bach, M. (2006). *Homepage of the Freiburg Visual Acuity and Contrast Test ('FrACT')*. URL: http://www.michaelbach.de/fract.html

Bach, M. (2007). The Freiburg Visual Acuity Test-variability unchanged by post-hoc re-analysis. *Graefes Arch. Clin. Exp. Ophthalmol.* 245, 965–971. doi: 10.1007/s00417-006-0474-4

Bach, M., and Schäfer, K. (2016). Visual Acuity Testing: feedback Affects Neither Outcome nor Reproducibility, but Leaves Participants Happier. *PLoS One* 11:e0147803. doi: 10.1371/journal.pone.0147803

Bao, J., Yang, A., Huang, Y., Li, X., Pan, Y., Ding, C. et al. (2021). One-year myopia control efficacy of spectacle lenses with aspherical lenslets. Br J Ophthalmol. bjophthalmol-2020-318367. doi: 10.1136/bjophthalmol-2020-318367

Chang, C. F., and Cheng, H. C. (2019). Effect of Orthokeratology Lens on Contrast Sensitivity Function and High-Order Aberrations in Children and Adults. *Eye Contact Lens* 46, 375–380. doi: 10.1097/icl.0000000000000667

Diaz-Llopis, M., and Pinazo-Duran, M. D. (2018). Superdiluted atropine at 0.01% reduces progression in children and adolescents. A 5 year study of safety and effectiveness. *Arch. Soc. Esp. Oftalmol.* 93, 182–185. doi: 10.1016/j.oftal.2017.12.015

Diec, J., Tilia, D., Thomas, V., and Bakaraju, R. C. (2018). Predicting Short-Term Subjective Vision Performance of Contact Lenses Used in Myopia Control. *Eye Contact Lens* 44, 308–315. doi: 10.1097/icl.0000000000000460

Fedtke, C., Bakaraju, R. C., Ehrmann, K., Chung, J., Thomas, V., and Holden, B. A. (2016). Visual performance of single vision and multifocal contact lenses in non-presbyopic myopic eyes. *Cont. Lens Anterior Eye* 39, 38–46. doi: 10.1016/j.clae.2015.07.005

Fernández, J., Rodríguez-Vallejo, M., Martínez, J., Burguera, N., and Piñero, D. P. (2019). Prediction of Visual Acuity and Contrast Sensitivity From Optical Simulations With Multifocal Intraocular Lenses. *J. Refract. Surg.* 35, 789–795. doi: 10.3928/1081597x-20191024-01

Fricke, T. R., Jong, M., Naidoo, K. S., Sankaridurg, P., Naduvilath, T. J., Ho, S. M., et al. (2018). Global prevalence of visual impairment associated with myopic macular degeneration and temporal trends from 2000 through 2050: systematic review, meta-analysis and modelling. *Br. J. Ophthalmol.* 102, 855–862. doi: 10.1136/bjophthalmol-2017-311266

García-Marqués, J. V., Macedo-De-Araújo, R. J., Cerviño, A., García-Lázaro, S., McAlinden, C., and González-Méijome, J. M. (2020). Comparison of short-term light disturbance, optical and visual performance outcomes between a myopia control contact lens and a single-vision contact lens. *Ophthalmic Physiol. Opt.* 40, 718–727. doi: 10.1111/opo.12729

Hohberger, B., Laemmer, R., Adler, W., Juenemann, A. G., and Horn, F. K. (2007). Measuring contrast sensitivity in normal subjects with OPTEC 6500: influence of age and glare. *Graefes Arch. Clin. Exp. Ophthalmol.* 245, 1805–1814. doi: 10.1007/s00417-007-0662-x

Holladay, J. T., Lynn, M. J., Waring, G. O. III., Gemmill, M., Keehn, G. C., and Fielding, B. (1991). The relationship of visual acuity, refractive error, and pupil size after radial keratotomy. *Arch. Ophthalmol.* 109, 70–76. doi: 10.1001/archopht.1991.01080010072036

Huang, J., Wen, D., Wang, Q., McAlinden, C., Flitcroft, I., Chen, H., et al. (2016). Efficacy Comparison of 16 Interventions for Myopia Control in Children: a Network Meta-analysis. *Ophthalmology* 123, 697–708. doi: 10.1016/j.ophtha.2015.11.010

Jaskulski, M., Singh, N. K., Bradley, A., and Kollbaum, P. S. (2020). Optical and imaging properties of a novel multi-segment spectacle lens designed to slow myopia progression. *Ophthalmic Physiol. Opt.* 40, 549–556. doi: 10.1111/opo.12725

Kang, P., McAlinden, C., and Wildsoet, C. F. (2017). Effects of multifocal soft contact lenses used to slow myopia progression on quality of vision in young adults. *Acta Ophthalmol.* 95, e43–e53. doi: 10.1111/aos.13173

Kawamorita, T., and Uozato, H. (2005). Modulation transfer function and pupil size in multifocal and monofocal intraocular lenses in vitro. *J. Cataract Refract. Surg.* 31, 2379–2385. doi: 10.1016/j.jcrs.2005.10.024

Kinoshita, N., Konno, Y., Hamada, N., Kanda, Y., Shimmura-Tomita, M., and Kakehashi, A. (2018). Additive effects of orthokeratology and atropine 0.01% ophthalmic solution in slowing axial elongation in children with myopia: first year results. *Jpn. J. Ophthalmol.* 62, 544–553. doi: 10.1007/s10384-018-0608-3

Lam, C. S. Y., Tang, W. C., Tse, D. Y., Lee, R. P. K., Chun, R. K. M., Hasegawa, K., et al. (2019). Defocus Incorporated Multiple Segments (DIMS) spectacle lenses slow myopia progression: a 2-year randomised clinical trial. *Br. J. Ophthalmol.* 104, 363–368. doi: 10.1136/bjophthalmol-2018-313739

Li, S. M., Kang, M. T., Wu, S. S., Liu, L. R., Li, H., Chen, Z., et al. (2016). Efficacy, Safety and Acceptability of Orthokeratology on Slowing Axial Elongation in Myopic Children by Meta-Analysis. *Curr. Eye Res.* 41, 600–608. doi: 10.3109/02713683.2015.1050743

Lombardo, M., and Lombardo, G. (2010). Wave aberration of human eyes and new descriptors of image optical quality and visual performance. *J. Cataract Refract. Surg.* 36, 313–331. doi: 10.1016/j.jcrs.2009.09.026

Lu, Y., Lin, Z., Wen, L., Gao, W., Pan, L., Li, X., et al. (2020). The Adaptation and Acceptance of Defocus Incorporated Multiple Segment Lens for Chinese Children. *Am. J. Ophthalmol.* 211, 207–216. doi: 10.1016/j.ajo.2019.12

Ma, D. J., Yang, H. K., and Hwang, J. M. (2013). Reliability and validity of an automated computerized visual acuity and stereoacuity test in children using an interactive video game. *Am. J. Ophthalmol.* 156, 195.e–201.e. doi: 10.1016/j.ajo.2013.02.018

Madrid-Costa, D., Tomás, E., Ferrer-Blasco, T., García-Lázaro, S., and Montés-Micó, R. (2012). Visual performance of a multifocal toric soft contact lens. *Optom. Vis. Sci.* 89, 1627–1635. doi: 10.1097/OPX.0b013e31826b0934

Mathôt, S., and Ivanov, Y. (2019). The effect of pupil size and peripheral brightness on detection and discrimination performance. *PeerJ* 7:e8220. doi: 10.7717/peerj.8220

Pieh, S., Marvan, P., Lackner, B., Hanselmayer, G., Schmidinger, G., Leitgeb, R., et al. (2002). Quantitative performance of bifocal and multifocal intraocular lenses in a model eye: point spread function in multifocal intraocular lenses. *Arch. Ophthalmol.* 120, 23–28. doi: 10.1001/archopht.120.1.23

Pomerance, G. N., and Evans, D. W. (1994). Test-retest reliability of the CSV-1000 contrast test and its relationship to glaucoma therapy. *Invest. Ophthalmol. Vis. Sci.* 35, 3357–3361.

Przekoracka, K., Michalak, K., Olszewski, J., Zeri, F., Michalski, A., Paluch, J., et al. (2020). Contrast sensitivity and visual acuity in subjects wearing multifocal contact lenses with high additions designed for myopia progression control. *Cont. Lens Anterior Eye* 43, 33–39. doi: 10.1016/j.clae.2019.12.002

Raasch, T. W., Bailey, I. L., and Bullimore, M. A. (1998). Repeatability of visual acuity measurement. *Optom. Vis. Sci.* 75, 342–348. doi: 10.1097/00006324-199805000-00024

Ravikumar, A., Sarver, E. J., and Applegate, R. A. (2012). Change in visual acuity is highly correlated with change in six image quality metrics independent of wavefront error and/or pupil diameter. *J. Vis.* 12:11. doi: 10.1167/12.10.11

Santodomingo-Rubido, J., Villa-Collar, C., Gilmartin, B., Gutiérrez-Ortega, R., and Sugimoto, K. (2017). Long-term Efficacy of Orthokeratology Contact Lens Wear in Controlling the Progression of Childhood Myopia. *Curr. Eye Res.* 42, 713–720. doi: 10.1080/02713683.2016.1221979

Siderov, J., and Tiu, A. L. (1999). Variability of measurements of visual acuity in a large eye clinic. *Acta Ophthalmol. Scand.* 77, 673–676. doi: 10.1034/j.1600-0420.1999.770613.x

Son, H. S., Tandogan, T., Liebing, S., Merz, P., Choi, C. Y., Khoramnia, R., et al. (2017). In vitro optical quality measurements of three intraocular lens models having identical platform. *BMC Ophthalmol.* 17:108. doi: 10.1186/s12886-017-0460-0

Vega, F., Millán, M. S., Garzón, N., Altemir, I., Poyales, F., and Larrosa, J. M. (2018). Visual acuity of pseudophakic patients predicted from in-vitro measurements of intraocular lenses with different design. *Biomed. Opt. Express* 9, 4893–4906. doi: 10.1364/boe.9.004893

Voelz, D. G. (2011). *Computational Fourier Optics: A MATLAB Tutorial*. Bellingham: SPIE Press.

Walline, J. J., Walker, M. K., Mutti, D. O., Jones-Jordan, L. A., Sinnott, L. T., Giannoni, A. G., et al. (2020). Effect of High Add Power, Medium Add Power, or Single-Vision Contact Lenses on Myopia Progression in Children: the BLINK Randomized Clinical Trial. *JAMA* 324, 571–580. doi: 10.1001/jama.2020.10834

Wildsoet, C. F., Chia, A., Cho, P., Guggenheim, J. A., Polling, J. R., Read, S., et al. (2019). IMI - Interventions Myopia Institute: interventions for Controlling

Myopia Onset and Progression Report. *Invest. Ophthalmol. Vis. Sci.* 60, M106–M131. doi: 10.1167/iovs.18-25958

Wong, Y. L., Sabanayagam, C., Ding, Y., Wong, C. W., Yeo, A. C., Cheung, Y. B., et al. (2018). Prevalence, Risk Factors, and Impact of Myopic Macular Degeneration on Visual Impairment and Functioning Among Adults in Singapore. *Invest. Ophthalmol. Vis. Sci.* 59, 4603–4613. doi: 10.1167/iovs.18-24032

Xu, R., Wang, H., Thibos, L. N., and Bradley, A. (2017). Interaction of aberrations, diffraction, and quantal fluctuations determine the impact of pupil size on visual quality. *J.Opt. Soc. Am. A Opt. Image Sci. Vis.* 34, 481–492. doi: 10.1364/josaa.34.000481

Zhang, H. Y., Lam, C. S. Y., Tang, W. C., Leung, M., and To, C. H. (2020). Defocus Incorporated Multiple Segments Spectacle Lenses Changed the Relative Peripheral Refraction: a 2-Year Randomized Clinical Trial. *Invest. Ophthalmol. Vis. Sci.* 61:53. doi: 10.1167/iovs.61.5.53

Zheng, Y. F., Pan, C. W., Chay, J., Wong, T. Y., Finkelstein, E., and Saw, S. M. (2013). The economic cost of myopia in adults aged over 40 years in Singapore. *Invest. Ophthalmol. Vis. Sci.* 54, 7532–7537. doi: 10.1167/iovs.13-12795

Conflict of Interest: BF is employed by company R&D Essilor International, Créteil, France. JB is an Associate Director of Wenzhou Medical University–Essilor International Research Centre. YB, AY, and EWL are employees of Essilor International. This company supplied the study devices and holds the following patent applications related to this work: WO2019166653, WO2019166654, and WO2019166655.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 28 June 2021
doi: 10.3389/fnins.2021.668116

Suppression of Luminance Contrast Sensitivity by Weak Color Presentation

Ippei Negishi[1,2]* *and Keizo Shinomori*[1,3]*

[1] School of Information, Kochi University of Technology, Kami, Japan, [2] Department of Media Informatics, College of Informatics and Human Communication, Kanazawa Institute of Technology, Hakusan, Japan, [3] Vision and Affective Science Integrated Laboratory, Research Institute, Kochi University of Technology, Kami, Japan

The results of psychophysical studies suggest that color in a visual scene affects luminance contrast perception. In our brain imaging studies we have found evidence of an effect of chromatic information on luminance information. The dependency of saturation on brain activity in the visual cortices was measured by functional magnetic resonance imaging (fMRI) while the subjects were observing visual stimuli consisting of colored patches of various hues manipulated in saturation (Chroma value in the Munsell color system) on an achromatic background. The results indicate that the patches suppressed luminance driven brain activity. Furthermore, the suppression was stronger rather than weaker for patches with lower saturation colors, although suppression was absent when gray patches were presented instead of colored patches. We also measured brain activity while the subjects observed only the patches (on a uniformly black background) and confirmed that the colored patches alone did not give rise to differences in brain activity for different Chroma values. The chromatic information affects the luminance information in V1, since the effect was observed in early visual cortices (V2 and V3) and the ventral pathway (hV4), as well as in the dorsal pathway (V3A/B). In addition, we conducted a psychophysical experiment in which the ability to discriminate luminance contrast on a grating was measured. Discrimination was worse when weak (less saturated) colored patches were attached to the grating than when strong (saturated) colored patches or achromatic patches were attached. The results of both the fMRI and psychophysical experiments were consistent in that the effects of color were greater in the conditions with low saturation colors.

Keywords: luminance-color interaction, fMRI, visual psychophysics, luminance contrast, visual attention

OPEN ACCESS

Edited by:
Zhong-Lin Lu,
New York University, United States

Reviewed by:
Givago Silva Souza,
Federal University of Pará, Brazil
Kowa Koida,
Toyohashi University of Technology, Japan

***Correspondence:**
Ippei Negishi
negishi@neptune.kanazawa-it.ac.jp
Keizo Shinomori
shinomori.keizo@kochi-tech.ac.jp

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *15 February 2021*
Accepted: *25 May 2021*
Published: *28 June 2021*

Citation:
Negishi I and Shinomori K (2021) Suppression of Luminance Contrast Sensitivity by Weak Color Presentation. Front. Neurosci. 15:668116. doi: 10.3389/fnins.2021.668116

INTRODUCTION

Numerous studies, including some by the present authors, have the investigated the interaction between chromatic and luminance information in the human visual system. The most general example of such an interaction is the change in color perception resulting from changes in the luminance of the surrounding background, referred to as blackness induction (Shinomori et al., 1994, 1997). Some psychophysical studies have reported interactions between chromatic and luminance information on luminance (Switkes et al., 1988; Kingdom, 2003;

Kingdom and Kasrai, 2006; Miquilini et al., 2017; Sousa et al., 2020), orientation (Clifford et al., 2003; Kingdom et al., 2010) and object shape (Clery et al., 2013) perception, while others have provided evidence that there is no effect (Victor et al., 1998; Xiao and Wade, 2010). Because of differences in the visual conditions and tasks used in those studies, it is difficult to conclude whether this interaction exists or not. Switkes et al. (1988) reported experimental evidence of an interaction between chromatic and luminance information on the basis of suppression of luminance information by chromatic information. Subsequent articles confirmed those results (Kingdom, 2003; Kingdom and Kasrai, 2006; Kingdom et al., 2010; Miquilini et al., 2017; Sousa et al., 2020); thus, we expected that some required conditions for suppression would exist. In former studies the saturation of color stimuli had not been considered, so in this study we specifically manipulated the saturation of the color in visual stimuli. In these studies, chromatic and luminance stimuli were presented at a spatially identical or proximate position, so the subjects observed the combined signal of color and luminance, and most of the subjects could not perceive chromatic and luminance information independently. Therefore, we presented chromatic stimuli in the surround of the visual field because we tried to measure the effect of chromatic information on luminance information rather than the antagonistic interaction between luminance and chromatic information.

There are two principal pathways in the human visual cortices (Goodale and Milner, 1992). Many studies have demonstrated that the visual cortices in the ventral pathway respond to chromatic stimuli (Mullen et al., 2007; Wade et al., 2008) and those in the ventral pathway processes features of objects, thus the ventral pathway is often referred to as the "what" pathway. With respect to chromatic signals the ventral pathway has been thought of as the main processing site. The dorsal pathway mostly processes information about environmental space and three-dimensional perception, and hence is often referred to as the "where" pathway. The information processed in the dorsal pathway seems to be mediated primarily by luminance signals. However, the results of some previous studies indicate that chromatic information influences the dorsal pathway under specific conditions, such as color defined motion stimuli (Maunsell, 1992; Tootell et al., 1995; Wandell et al., 1999), suggesting that the dorsal and ventral pathways are connected via the ventral occipital fasciculus (Yeatman et al., 2014; Xing et al., 2015; Takemura et al., 2016). This would suggest that luminance information receives the effect of chromatic information before the divergence of the dorsal and ventral pathways. Primary or early visual cortices, especially human V1, are known to respond to chromatic stimuli (Engel et al., 1997; Mullen et al., 2007); luminance-selective, color-selective and both luminance- and color-selective neurons were found in macaque V1 (Livingstone and Hubel, 1984; Johnson et al., 2001), human visual cortices (V1, V2, and V3), the ventral pathways (hV4), and the dorsal pathways (V3A/B) in observing visual stimuli consisting of luminance and chromatic components (Bartels and Zeki, 2000; Brouwer and Heeger, 2009). Through this experiment, we aimed to clarify the stimulus conditions for the interaction between luminance and chromatic information, and to identify the visual cortex where this interaction occurs.

It is broadly accepted and supported by the results of numerous psychophysical (Brefczynski and DeYoe, 1999; Somers et al., 1999; Turatto and Galfano, 2000; Morrone et al., 2002; Fuller and Carrasco, 2006) and physiological (Bannert and Bartels, 2013) studies that human visual perception is influenced by attention. The amplitude of blood-oxygen-level dependent (BOLD) responses in early human visual cortices is also modulated by subject attention (Brefczynski and DeYoe, 1999; Somers et al., 1999; Turatto and Galfano, 2000; Morrone et al., 2002; Fuller and Carrasco, 2006). Additionally, brain activity patterns can receive the effect of the attention from higher-level cortices (Bannert and Bartels, 2013). As we thought attention might also affect the interaction between luminance and color (although the clear evidence of this point has not been shown), we measured the effect of the chromatic information when subjects were not required to attend to some specific feature in their field of vision, as the tasks in previous experiments (Switkes et al., 1988) did not require active attention.

We firstly conducted a physiological experiment to measure brain activity using fMRI. However, since we obtained unexpected results, we subsequently performed a psychophysical experiment to confirm that the fMRI results were in accordance with psychophysical results obtained for similar visual stimuli (see "Psychophysical Experiment" section for details). The results of the fMRI experiment provided evidence of the suppression of luminance information by chromatic information, and confirmed that the colored patches themselves did not give rise to differences in brain activity for different Chroma values. Interestingly, when the saturation of the patches was lower there was an interaction such that the suppression was stronger rather than weaker. The results of the fMRI experiment were supported by the results of the psychophysical experiment measuring luminance contrast discrimination ability.

fMRI EXPERIMENT

Methods

Subjects

Thirteen subjects (twelve male and one female, mean age: 22.7 $\pm$ 2.1 years) participated in the fMRI experiment. All subjects had normal or corrected to normal visual acuity better than 1.4 min of visual angle. Their color vision was classified as color normal by traditional color vision tests: Ishihara-plate (International 38 plates edition); Standard Pseudo-isochromatic Plates (SPP), and Panel D-15.

Visual Stimuli

Visual stimuli were presented on a rear projection screen using an LCD projector (DLA-X70R-B, Victor). A special visual stimulus presentation system (ViSaGe, Cambridge Research Systems, Inc.) was used to generate the visual stimuli. The subjects observed the screen by means of a mirror attached to the head coil. The resolution of the projector was 1,280 (horizontal) $\times$ 1,024 (vertical) pixels and the stimuli were presented in 30-bit color

mode. The angular size of the visual stimuli was 31.5° × 25.2° at a 69 cm viewing distance. Chromaticity coordinates and luminance of all colors in the stimuli were controlled and calibrated on the presentation screen, using the special visual stimulation presentation system and its Gamma and chromaticity coordinates calibration data. The measurements were performed directly on the screen by means of a colorimeter (CS-200, Konica-Minolta, Inc.) and spectral radiometer (CS-1000, Konica-Minolta, Inc.) confirming that the screen presentation error was less than 0.005 for CIE (Commission Internationale de l'Éclairage) 1931 xy chromaticity coordinates, and less than 5% for luminance.

Visual stimuli consisted of patch and scrambled patterns, as shown in **Figure 1A**. The patch pattern consisted of ten colored patches, the centers of which were placed on a 12.4° diameter circle with equal arc angle (36.0°) on an achromatic background (top panel of **Figure 1A**). The diameter of each patch was 3.0° and each patch had a 15 min. wide black fringe. The same background pattern, consisting of 600 achromatic ellipses of fixed size (3.3° × 1.8°) with randomized orientation and position, was used for all stimuli. In this manner, we could present luminance edges to evoke brain activity while avoiding a possible anisotropic effect (Loffler and Orbach, 2001; Krukowski and Stone, 2005). The luminance of the ellipses was assigned randomly as one of eight levels from 0.23 to 30.24 cd/m^2 and the mean luminance was 15.4 cd/m^2. The multiple levels of luminance could make more edges than simple two-leveled luminance, and reduce luminance induction effects (Shinomori et al., 1994, 1997). A white fixation rectangle was presented at the center of the screen. Each patch had one of ten different color hues to avoid hue specific activations (Engel et al., 1997; Shinomori and Werner, 2012; Kuriki et al., 2015) and the colors are presented in **Figure 1B**. The hues were calibrated to 5R, 5YR, 5Y, 5GY, 5G, 5BG, 5B, 5PB, 5P, and 5RP in the Munsell color system under a D65 lightning condition. The Munsell Value (which corresponds to lightness) of the patches was fixed to 5/ (approximately 16.0 cd/m^2). The Chroma value (which corresponds to saturation) was manipulated as an experimental parameter and was set to one of /6, /4, /2, and /0. The Chroma /0 condition represents achromatic patches, meaning the color of the ten patches was an identical gray color, so no effect of chromatic information was expected in that condition. The subjects could distinguish each hue in the magnetic resonance imaging (MRI) scanner in the /6, /4, and /2 Chroma conditions. The position of each patch was randomly chosen from among three conditions: the uppermost of the patches was in either the 12 o'clock position, rotated 12° clockwise, or rotated 12° counterclockwise. The scrambled pattern was generated by randomizing the location of the pixels. All pixels on the screen, including background pixels (but excluding the white fixation rectangle), were randomized. The fixation point was set at the center of the screen.

In the control experiment described later, the achromatic patterns (ellipses) were removed from the background, so that the background was uniformly black (0.05 cd/m^2). In the scrambled pattern in the control experiment, only pixels in the locations of the patches were pseudorandomized and the rest of the screen was kept continuously black.

Procedure

A block design was employed for the experimental sequence. One block consisted of stimulation and resting phases, and each phase lasted 15 s. In the stimulation phase the patch and scrambled patterns were altered at 1 Hz to sustain the activation of the visual cortex (**Figure 1A**). The arrangements of the hues of the ten patches were consistent in each block and randomized among blocks. The subjects had no task in the stimulation phase; they were only required to observe the visual stimuli. The purpose of the resting phase was to reset brain activity.

An easy cognitive task, detecting a change of a central target between "+" and "o" occurring at 50% probability, was performed in the first 5 s of the resting phase so as to sustain and confirm awareness of the subject. If the subject erred (including no response) more than three times in one run, the fMRI data in that run was not included in further analyses. The correct answer rate of the cognitive task was 98.7% (SEM: 0.4%). The last 10 s of the resting phase was completely rest time and nothing was presented on the screen.

We set each block as a pair of one stimulation and one resting phase. One run consisted of one dummy block and twelve subsequent blocks performed sequentially, and each Chroma condition (/6, /4, /2, and /0) was presented three times in pseudorandom order. Each subject performed 18 runs and the subjects could rest between runs for an unlimited period of time in or out of the scanner. The experiments were conducted over several days.

MRI Parameters

MRI data were collected on a Siemens 3T scanner (Verio) and a 32 ch head coil. The subjects who needed visual acuity correction wore goggles with appropriate corrective lenses. Structural MRI images (T1-weighted images) were collected by an MPRAGE sequence with TR = 2,250 ms, TE = 3.85 ms, FA = 9°, and 1.0 mm × 1.0 mm × 1.0 mm voxel size. Functional images were collected by an EPI sequence with TR = 2,500 ms, TE = 30 ms, FA = 80°, and 3.0 mm × 3.0 mm × 3.0 mm voxel size. The images consisted of 38 slices (64 × 64 voxel) with 0.45 mm gaps.

MRI Data Processing

We used BrainVoyager QX (ver. 2.8) for the processing and analysis of the MRI images. We realigned functional images to remove head movement during experiments and aligned them to the structural image of each subject with a six-parameter affine transformation. We applied a temporal high pass filter to functional images to reduce noise. No spatial filters were applied to images because spatial resolution was considered more important than the S/N ratio, and the number of blocks was sufficient to secure reasonable S/N ratios. BOLD responses were averaged for each Chroma condition and each visual cortex. Time courses of the BOLD response obtained from 54 blocks (three blocks per run × 18 runs) were averaged (excluding the data of one run that was removed for two subjects due to a low awareness level as defined by attention to the cognitive task).

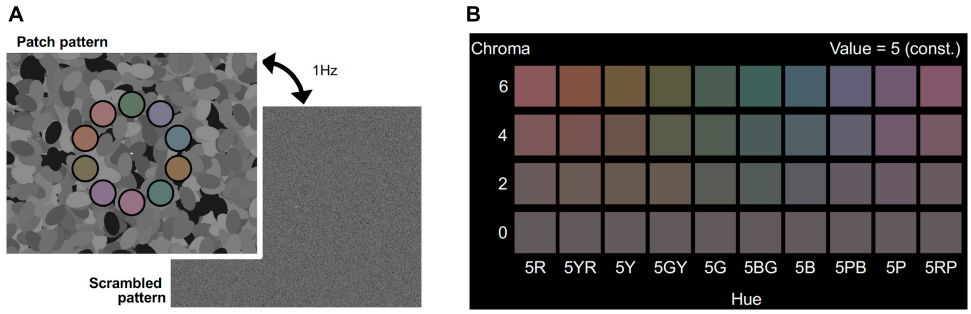


FIGURE 1 | (A) Visual stimuli for the fMRI experiment. Visual stimuli alternating at 1 Hz between the patch pattern (top) and the scrambled pattern (bottom) was presented in the stimulation phase of the fMRI experiment. Colors of the 10 patches had the same Chroma (/6, /4, /2, or /0) and Value (5/) of Munsell color system. Achromatic ellipses in the background had one of eight luminance levels (from 0.3 to 30.4 cd/m^2). Scrambled patterns were generated by randomly shuffling the coordinates of pixels in the patch patterns. The fixation point was set at the center of the screen. **(B)** Set of colors for patches in the four Chroma conditions. Chroma value of each visual stimulus was either /6 (Top raw), /4 (second raw), /2 (third raw), or /0 (bottom raw). All patches had the same gray color in Chroma /0 condition.

We defined five ROIs using the traveling wave method (Wandell et al., 2007) and calculated the change in the BOLD response in those ROIs (**Figure 2**). We selected a two-Gaussian hemodynamic response function (HRF) model (Rajapakse et al., 1998) and applied it to the discrete time course of each BOLD response for each visual cortex and each subject. The model response ($\boldsymbol{ModelR}_{ch}$) at time $\boldsymbol{t}_n$ is expressed by Eq. 1.

$$\begin{aligned} ModelR_{ch}(t_n) &= \int_{t_n-15}^{t_n} \left[a_{1,Ch} H(\tau) \times \exp\left\{ \frac{-(\tau - T_1)^2}{2\alpha_1^2} \right\} \right. \\ &\quad \left. + a_{2,Ch} H(\tau) \times \exp\left\{ \frac{-(\tau - T_2)^2}{2\alpha_2^2} \right\} \right] d\tau \\ &\quad + \left(b_{Ch} t_n + d_{Ch} \right) \end{aligned} \quad (1)$$

Where $\boldsymbol{a}_1$ (positive value) and $\boldsymbol{a}_2$ (negative value) are the intensity parameters of two Gaussian pulses; $\boldsymbol{T}_1$, $\boldsymbol{T}_2$, α_1, and α_2 are the four temporal parameters of the pulses; $\boldsymbol{H(\tau)}$ is the Heaviside step function; $\boldsymbol{b}_{ch}\boldsymbol{t}_n + \boldsymbol{d}_{ch}$ is the linear component, and the discrete time $\boldsymbol{t}_n$ ($\boldsymbol{n}$ = 0–12) is the series time in seconds ($\boldsymbol{t}_n$ = 0.0, 2.5, 5.0, 7.5, 10.0, 12.5, 15.0, 17.5, 20.0, 22.5, 25.0, 27.5, and 30.0). In the model fit, $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{b}$, and $\boldsymbol{d}$ were varied for each data fit to minimize Error Value ($\boldsymbol{EV}$) defined by Eq. 2 where the four temporal parameters ($\boldsymbol{T}_1$, $\boldsymbol{T}_2$, α_1, and α_2) were identical for each subject and each cortex through all four Chroma conditions ($\boldsymbol{Ch}$ = /0, /2, /4, and /6):

$$EV = \sum_{Ch\ =0,/2,/4,/6} \left[\sum_{n\ =0}^{12} \left[\frac{\left\{ R_{Ch}(t_n) - ModelR_{Ch}(t_n) \right\}}{ModelR_{Ch}(t_n)} \right]^2 \right] \quad (2)$$

Where $\boldsymbol{R}_{Ch}$ is the measured BOLD response, $\boldsymbol{ModelR}_{Ch}$ is the best-fit model response of Eq. 1, and $\boldsymbol{t}_n$ is the discrete time. We employed the peak point of the fitted curve as the response (% signal change) under each condition, interpreted as responses to visual stimuli. The best, worst and intermediate (around the average and the median) examples of application are shown in **Figure 2**. The left and right hemisphere responses were combined for analysis. **Figure 3** is a histogram of $\boldsymbol{EV}$ in all fittings for each subject and each visual cortex (total $\boldsymbol{N}$ = 65). The distribution of the error in the model fits was small except for the V3A/B data of some subjects. $\boldsymbol{EV}$s were large in V3A/B for some subjects because the BOLD response of V3A/B was the weakest among the visual fields.

Visual Cortex Segmentation

We measured retinotopic mappings using the traveling wave method (Wandell et al., 2007), defining V1, V2, V3, V3A/B, and V4 separately for each subject. We attempted to keep the region of lateral occipital area 1 (LO1) separated so as to clarify the border with V3 and V3A/B. However, since the anterior border of LO1 was not clear, we did not include LO1 in the BOLD signal analysis. We used a flickering checkerboard pattern (11.9° × 11.9° of visual angle) of minimum black (0.05 cd/m^2) and maximum white (360.3 cd/m^2) on the screen, a wedge-shaped aperture moving at a polar angle (45°), and a ring shape of eccentricity from 0 to 11.9°. Cycle time was 24 s. Functional images were collected by means of EPI sequence with TR = 2,000 ms, TE = 30 ms, FA = 80°, and 3.0 mm × 3.0 mm × 3.0 mm voxel size. In addition to BrainVoyager QX (ver. 2.8), we used FreeSurfer (ver. 5.3.0) for segmentation of white and gray matter. A typical example of visual field segmentation is displayed in **Supplementary Figure 1**.

Results

We compared brain activity measured while the subjects were observing alternating visual stimuli (**Figure 1A**) in the four Chroma conditions of the patches. Brain activity of each subject in each condition was expressed by the normalized BOLD response; each BOLD response was normalized by the average of all four Chroma conditions in all runs (216 blocks). The means of the averaged BOLD response (% signal change) under all four Chroma conditions for all subjects were 3.14% (V1), 2.22% (V2), 1.80% (V3), 0.706% (V3A/B), and 1.44% (hV4). We compared the normalized BOLD response between the Chroma conditions (**Figure 4**). Brain activity in the Chroma /0 condition was the largest in all identified visual cortices, and those in the Chroma /6, /4, and /2 conditions showed a decreasing order of strength

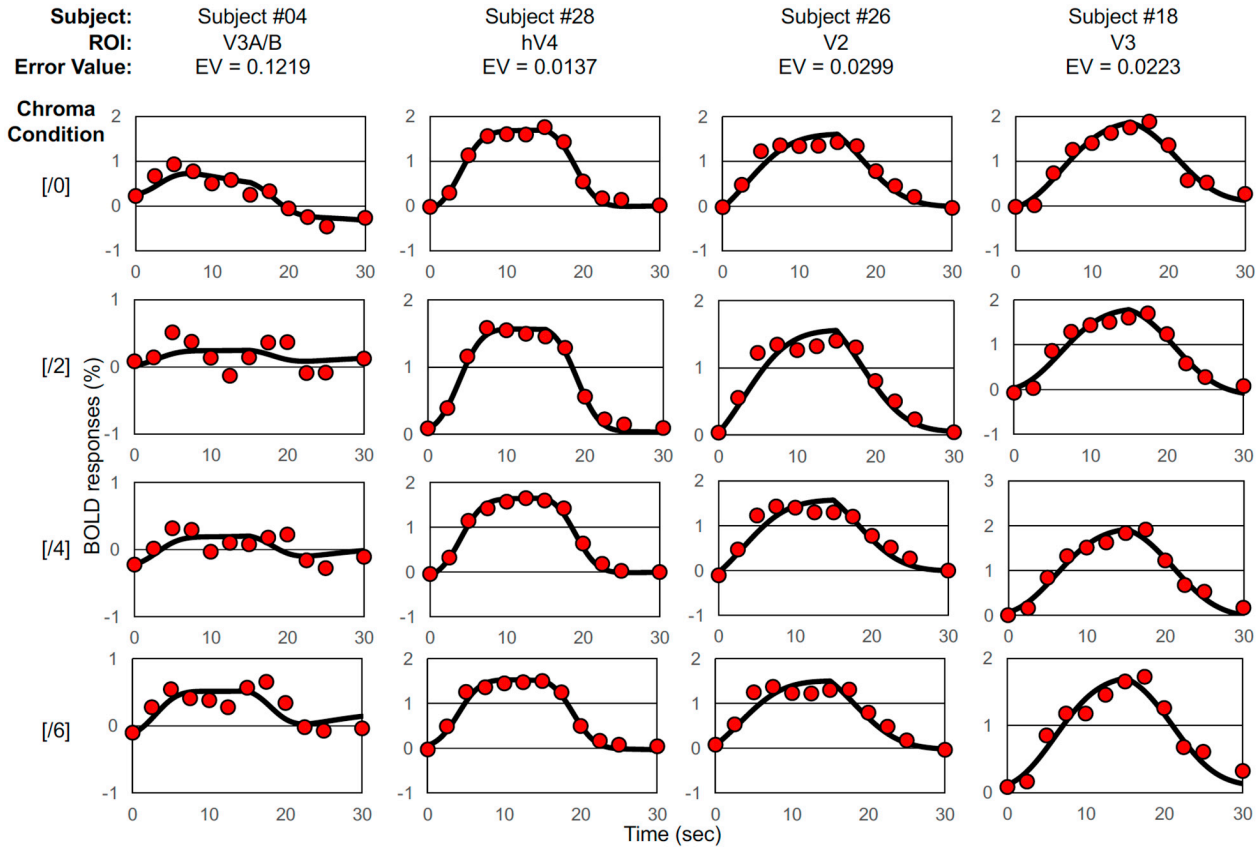


FIGURE 2 | Examples of BOLD signal data and model fits. Selected BOLD signal data (denoted by red circles) and model fittings generated by the two-Gaussian pulse model (Eq. 1, solid curves) under Chroma conditions from /0 (top raw) to /6 (bottom raw), the worst-case fitting (EV = 0.1219) (first column), the best-case fitting (EV = 0.0137) (second column), the case of fitting where EV was around the average (EV = 0.0299) (third column), and the case of fitting where EV was the median (EV = 0.0223) (forth column).

(from /6 to /4 to /2) in all visual cortices analyzed (V1, V2, V3, V3A/B, and hV4). A one-way repeated measures (within subjects) analysis of variance (ANOVA) was performed for the data before normalization, and showed a significant main effect of the Chroma condition in all measured visual cortices (**Table 1**). Activity in the Chroma /2 condition was significantly smaller than that in the Chroma /0 condition, as shown by Tukey's HSD test (V1: $p = 0.007$; V2: $p = 0.024$; V3: $p = 0.006$; V3A/B: $p = 0.004$; and hV4: $p = 0.045$) for all cortices, and in Chroma /6 condition for V3A/B ($p = 0.044$). These results suggest that the Chroma of the patches modulated brain activity of the visual cortices, but the order of Chroma values (/0, /2, /4, and /6) showed no correlation with the order of the change in the signal strength of brain activity.

We thought that the brain activity observed in the experiment should correspond to that under achromatic pattern backgrounds, since from the viewpoint of retinotopy, the brain regions corresponding to the visual field of the colored patches had to be relatively small and must be predictable from the segmentation data of the visual cortices. Moreover, the effect of the Chroma condition was also significant in the dorsal pathway (V3A/B), despite the fact that the primary stream for chromatic information processing is known to be the ventral pathway (hV4). Thus, we assumed that brain activity caused by chromatic stimuli did not reflect the strength of chromatic information (in terms of saturation), but rather reflected the strength of achromatic (luminance) information. This means that the patches suppressed the brain activity driven by the achromatic pattern background and that the suppression magnitude was larger under lower-Chroma conditions. Thus, we conducted a control experiment to confirm that the differences in brain activity in the results of the main experiment were not a result of the difference in brain activity driven directly by colored patches. If the results of the main experiment were the results of the activity evoked directly by the patches, the results of the control experiment would present the same pattern as those of the main experiment. Thirteen individuals (nine male and four female, mean age: 23.5 $\pm$ 4.3 years) participated in the control experiment and eight of them had already participated in the main experiment.

However, as can be seen in **Figure 5**, the results of the control experiment were different from those of the main experiment in terms of dependency on the Chroma value. There were no significant differences among the Chroma conditions (**Table 1**).

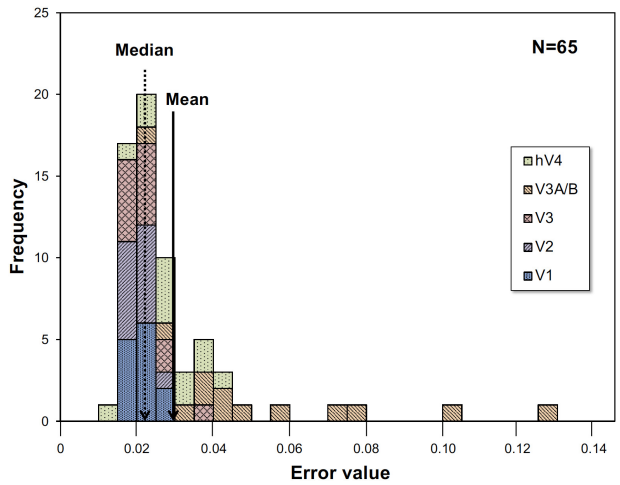


FIGURE 3 | Histogram of error value (EV). Histogram for all thirteen subjects and five visual cortices. EVs were large in V3A/B for some subjects.

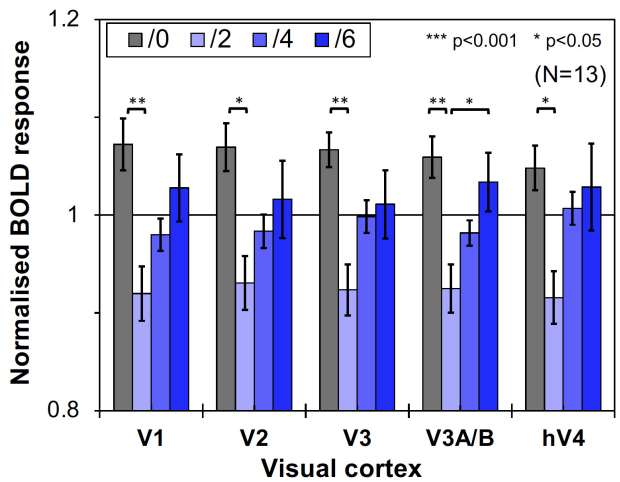


FIGURE 4 | Brain activity of visual cortices under the four Chroma conditions (/0, /2, /4, and /6). The vertical axis units are percent amount of signal change defined as incremental rate of BOLD response subtracted from the mean of the incremental rate under all four Chroma conditions in the stimulation phase. Error bars denote SEM. Asterisks denote statistical significance by Tukey's honestly significant difference (HSD) test (**$p < 0.01$; *$p < 0.05$).

The means of the normalized BOLD responses for the Chroma conditions in the control experiment were 1.69% (V1), 1.38% (V2), 1.18% (V3), 0.973% (V3A/B), and 1.02% (hV4). Among the results, brain activity for the patches could not explain the greater depression of the responses under low-Chroma conditions in the background pattern in the main experiment. This lends support to our suppression hypothesis as an explanation of the results of the main experiment. Conversely, the absence of the achromatic pattern background and the simple order of the signal change from Chroma /0 to /6 suggest that the responses measured in the control experiment were evoked by the patches directly, although the differences among Chroma conditions were not statistically significant and the response of V3 could not be explained adequately. Additionally, V1, V2, V3, and hV4 brain activity in the control experiment was significantly smaller in magnitude than that in the main experiment while there is no significant difference in V3A/B (**Figure 6**).

Furthermore, we compared the activities of the areas that retinotopically corresponded to the areas inside the patches (less than 4.5° eccentricity) and on the patches (5.0°–6.5° eccentricity). We used only the V1 BOLD data since the accuracy of the retinotopy in the regions was considered and the magnitude of the BOLD responses in V1 was the maximum (**Figure 6**). As shown in **Figure 7**, the activities caused by the Chroma conditions in both areas shared the same common tendency we described in the main experiment (**Table 1**). This supports our hypothesis that the color patches affected the luminance information. We also confirmed that the rate of the activities inside the patches relative to those on the patch were smaller in the control experiment than in the main experiment (**Figure 8**), suggesting reasonably good quality in the segmentation of areas inside- and on- the patches.

PSYCHOPHYSICAL EXPERIMENT

We conducted a psychophysical experiment to investigate the correlation between the fMRI imaging data and psychophysical data expressing the relationship between color presentation and luminance-dependent perception. The suppression hypothesis was also tested by a psychophysical experiment in which we conducted a luminance contrast discrimination task in order to estimate the strength of the luminance signal when colored patches were presented. Initially we measured perceived luminance contrast in conditions with or without the patches, but we failed to observe significant difference between them because it was difficult for the subjects to respond to luminance contrast of chromatic images. Thus, we employed the slope of fitting function as an index of luminance perception.

We used an achromatic grating with 10 overlapping colored patches (shown in the second and fifth panels of **Figure 9**) and measured the response rate of choosing the grating with higher luminance contrast between two gratings presented successively in a temporal two alternative forced choice (2AFC) method (**Figure 9**). The response rate data as a function of luminance contrast were fitted to a cumulative Gaussian distribution function (psychometric function) and discrimination ability (represented by the deviation of the data distribution) was obtained. From the psychophysical data we obtained the relationship between the Chroma of the patches and the luminance signal which was estimated from discrimination ability.

Methods

Subjects

Eleven subjects (five male and six female, mean age: 24.1 ± 5.4 years) participated in the psychophysical experiment. Six subjects participated in the fMRI experiments and this experiment. Visual acuity and color vision were evaluated in all subjects, as in the

TABLE 1 | ANOVA results of fMRI experiments for each visual cortex.

		Main experiment				Control experiment			
Source	**Df**	**Sum Sq**	**Mean Sq**	***F* value**	***p***	**Sum Sq**	**Mean Sq**	***F* value**	***p***
V1									
Chroma condition	3	0.167	0.056	4.392	0.001**	0.007	0.002	0.064	0.979
Subject	12	22.547	1.879			21.141	1.762		
Residuals	36	0.457	0.013			1.316	0.037		
V2									
Chroma condition	3	0.132	0.044	3.154	0.037*	0.020	0.007	0.177	0.912
Subject	12	13.325	1.353			13.894	1.158		
Residuals	36	0.504	0.014			1.389	0.039		
V3									
Chroma condition	3	0.136	0.045	4.166	0.012*	0.071	0.024	0.896	0.452
Subject	12	7.507	0.626			6.729	0.561		
Residuals	36	0.390	0.011			0.944	0.026		
V3A/B									
Chroma condition	3	0.140	0.047	5.113	0.005**	0.082	0.027	1.179	0.331
Subject	12	3.771	0.314			12.442	1.037		
Residuals	36	0.329	0.009			0.836	0.023		
hV4									
Chroma condition	3	0.134	0.045	2.933	0.046*	0.236	0.079	1.783	0.168
Subject	12	9.103	0.759			10.982	0.915		
Residuals	36	0.548	0.015			1.587	0.044		
V1 (inside the patch)									
Chroma condition	3	0.293	0.097	5.009	0.005**	0.085	0.028	0.767	0.520
Subject	12	14.950	1.246			9.791	0.816		
Residuals	36	0.702	0.020			1.333	0.037		
V1 (on the patch)									
Chroma condition	3	0.160	0.054	5.521	0.003**	0.061	0.020	0.564	0.642
Subject	12	39.230	3.269			28.136	2.345		
Residuals	36	0.350	0.010			1.246	0.036		

$^{**}p < 0.001$; $^{*}p < 0.05$.

fMRI experiments. Note that we excluded the data of one subject because their mean standard deviation under the four Chroma conditions was substantially higher (0.35) than that for other subjects (ranging from 0.05 to 0.12).

Visual Stimuli

In the psychophysical experiment, visual stimuli were generated by ViSaGe and presented on a CRT monitor (FlexScan E57T, EIZO) in a dark room (not in the fMRI scanner room). The monitor was calibrated in the same manner as the fMRI experiment. The subjects were fixed at a viewing distance of 54.4 cm by means of a chin rest, and the size of the screen was 31.5° × 23.6°. Though we first tried to use the background pattern in the fMRI experiment, the task was so easy that the slope of their psychometric functions were too steep for comparison between chromatic conditions. Therefore, we decided to employ another visual stimulus in the psychophysical experiment rather than maintaining consistency of visual stimuli appearance between the two experiments (see section "Differences of Visual Stimuli Between the Experiments" for details). In the temporal 2AFC method two types of visual stimuli, test and reference stimuli, were presented sequentially in random order (**Figure 9**). The stimuli consisted of an achromatic grating (0.67 cycles per degree) enveloped by a decremental exponential-function (3.75° for 1/e decline) and 10 colored patches with a black fringe overlapping the grating (**Supplementary Figure 2**). The Gabor function was not used for the envelope of the grating because the decline rate of the peak luminance produced by the Gaussian envelope was too high and the gratings at the low luminance contrast part (ex. 2.5%) were overlapping with some of the patches (as can be seen in **Supplementary Figure 2**). The number of visible grating peaks could have influenced subject judgment of luminance contrast. The mean luminance of the grating and the luminance of the patches were set to 19.8 cd/m^2 in both test and reference stimuli. The direction and phase of the stripes in the grating were also fixed. In the reference stimuli, the spatially-maximum Michelson contrast (negative contrast at the center) of the luminance grating was fixed at 50%. In the test stimuli, the maximum contrast was varied from 34 to 66%. We also manipulated the Chroma of the patches (from /0 to /6) and other details of the patches (size, position, arrangement of colors, and color properties) to be the same as

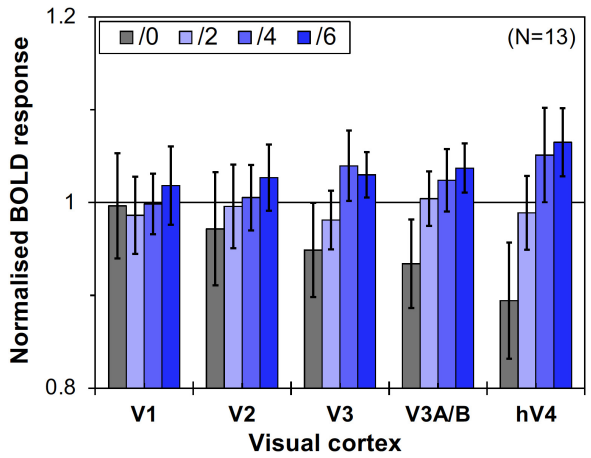


FIGURE 5 | Brain activity of visual cortices under the four Chroma conditions (/0, /2, /4, and /6) in the control experiment. Other notations are the same as in **Figure 4**.

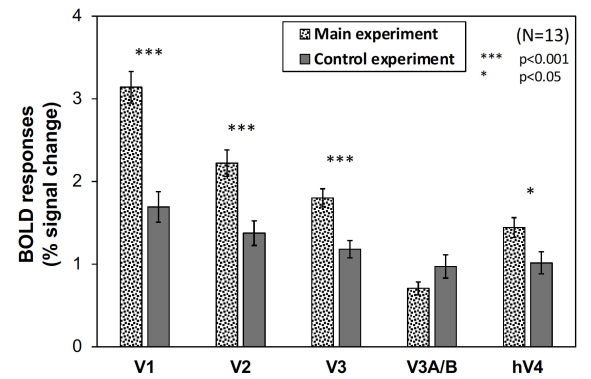


FIGURE 6 | The magnitude (% signal change) of BOLD responses in main and control experiments. The values were averaged across all Chroma conditions and subjects. Error bars denote SEM among the subjects. Significant differences in V1 ($p < 0.001$), V2 ($p < 0.001$), V3 ($p < 0.001$), and hV4 ($p < 0.05$) were confirmed by *t*-test.

those in the fMRI experiments. The Chroma condition of the test and the reference stimuli were the same in the trial but were changed pseudorandomly between trials. The mask pattern in the psychophysical experiment consisted of achromatic squares ($1.5° \times 1.5°$) of which the luminance was chosen randomly from among eight levels ranging from 0.32 to 30.37 cd/m^2 and the mean luminance was 15.38 cd/m^2.

Procedure

In the trial, test and reference stimuli were presented sequentially in random order (**Figure 9**). The duration of both stimuli was 500 ms, and a mask stimulus was presented for 300 ms after each stimulus. ISI between test and reference stimuli was randomly set between 1,000 and 2,000 ms. After the appearance of the two stimuli, the subject reported which stimulus had higher achromatic (luminance) contrast in the grating. Except during the time when the test, reference, and mask stimuli were being presented, a bright fixation point was presented on a uniformly gray background (19.8 cd/m^2) (**Figure 9**). There were 36 conditions in total (nine conditions for contrast of the test stimuli by four chromatic conditions) and all conditions were presented in each session. Each subject performed 12 experimental sessions. Throughout the experiment the subjects received no feedback about their responses, and were not informed which (first or second) stimulus was the test stimulus in each trial. The data for the initial two sessions were discarded so as to avoid the strong learning effect in the earliest sessions.

Data Analysis

The data was fit with a cumulative Gaussian distribution function (GDF) as the psychometric function (Eq. 3).

$$P(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) dx \qquad (3)$$

Where x is the maximum luminance contrast (Michelson contrast) of the grating in the test stimulus, μ is the contrast at 50% probability (chance level), and σ is the standard deviation of the distribution function. In fits to the probability data using this function, the optimized parameters μ (mean of the data) and σ were obtained. A larger σ results in a gentler slope of the cumulative GDF and indicates higher contrast discrimination threshold at defined criterion (i.e., 75%) as shown in **Figure 10**. In signal detection theory this larger σ gives rise to lower and wider signal distribution functions and corresponds to a lower S/N ratio of the discrimination task (**Figure 10**). Under the constant noise assumption the lower S/N ratio indicates lower signal strength for the discrimination task. Thus, we employed σ values as an index for luminance contrast sensitivity.

Results

Figure 11 shows examples of the data and psychometric functions for one subject in the four Chroma conditions. **Figure 12** shows the mean of the standard deviation, σ under the four Chroma conditions. A one-way repeated measures (within subjects) ANOVA showed a significant main effect of Chroma condition (**Table 2**), and the standard deviation under the Chroma /2 condition was significantly larger than those under the other Chroma conditions (by Tukey's HSD test: Chroma /0: $p = 0.046$, Chroma /6: $p = 0.021$, and Chroma /4: $p = 0.007$). This result indicates that luminance contrast discrimination ability was worse in the Chroma /2 condition, and this result is consistent with the result of fMRI experiment that the suppression of luminance information was most prominent in Chroma /2 patches.

GENERAL DISCUSSION

The relationship between luminance, mostly treated in a form defined by luminance in previous studies, and color has been studied extensively in various visual stimuli (Moutoussis, 2015). The results of previous psychophysical studies using gratings show that presentation of color contrast has a masking effect on luminance contrast and vice versa (Mullen and Losada,

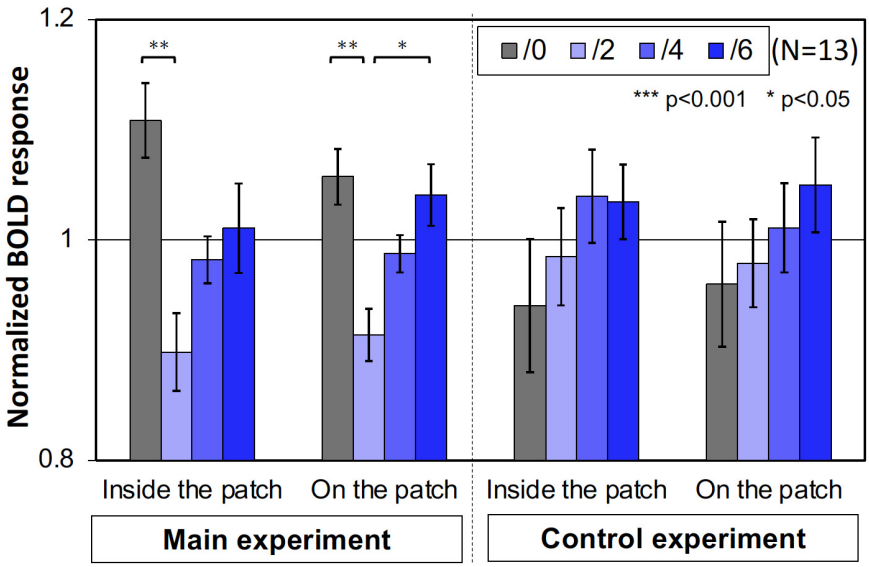


FIGURE 7 | The normalized brain activity of the areas corresponded to inside and on the patches. A one-way ANOVA for the data before normalization, showed a significant effect on both areas only in main experiment (**Table 1**). Error bars denote SEM. Asterisks denote statistical significance by Tukey's HSD test (**$p < 0.01$; *$p < 0.05$).

1994; Moutoussis, 2015). Switkes et al. demonstrated that presentation of luminance information (i.e., mask grating) suppressed responses mediating chromatic information (i.e., test grating), and that suppression is stronger at a certain luminance contrast (about 2%) than for higher or lower contrasts. However, their report does not indicate the suppression of luminance information by chromatic information at low color contrast; rather, the results show monotonic increment of suppression (i.e., higher threshold) as the color contrast increased, indicating a simple masking effect (Switkes et al., 1988).

Here, we first demonstrated the suppression of brain activity by the presentation of chromatic stimuli. We also demonstrated suppression due to dependency of saturation on stimulus colors in both brain activity and luminance contrast sensitivity; the suppression of the luminance information mostly occurred with the presentation of color with low saturation (low Chroma), and little or no suppression was observed with the presentation of no color (Chroma /0) or highly-saturated colors (Chroma /6) (**Figure 13**). We note that most of previous fMRI studies of color vision focused on responses mediating chromatic information in visual objects. In our study an achromatic background was dominant in the visual stimuli and we succeed in identifying the effect of color presentation on brain activity through examination of responses evoked by achromatic stimulation.

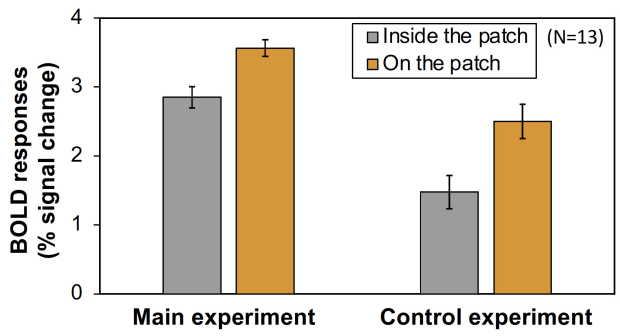


FIGURE 8 | The comparison of the magnitude (% signal change) of V1 BOLD responses inside- and on- the patch in main and control experiments. The values were averaged across all Chroma conditions and subjects. Error bars denote SEM among the subjects. The Inside/On ratio was significantly larger in main experiment (Main experiment: 0.842, Control experiment: 0.609, $p = 0.011$ by t-test).

Luminance Suppression and Color Enhancement Model

Brain activity relating to luminance information was only suppressed when the colored patches were attached to an achromatic background, and the amount of suppression did not simply correspond to the color strength (Chroma), recognized perceptually as intensity of chromatic information. Instead, suppression occurred only when the visual stimuli contained low saturation (low chromatic information), and did not occur when color (chromaticness) of visual stimuli was not recognized or when chromaticness of stimuli was sufficiently strong. These suppression-related phenomena can be accounted for by the hypothesis of the color enhancement model using bimodality; if the strength of the chromatic information is relatively weak, the chromatic information will be relatively enhanced to achieve a better balance between chromatic and luminance information (**Figure 13**). In terms of benefit, weak colors are perceived more clearly under suppression, although the S/N ratio of luminance information deteriorates as a result.

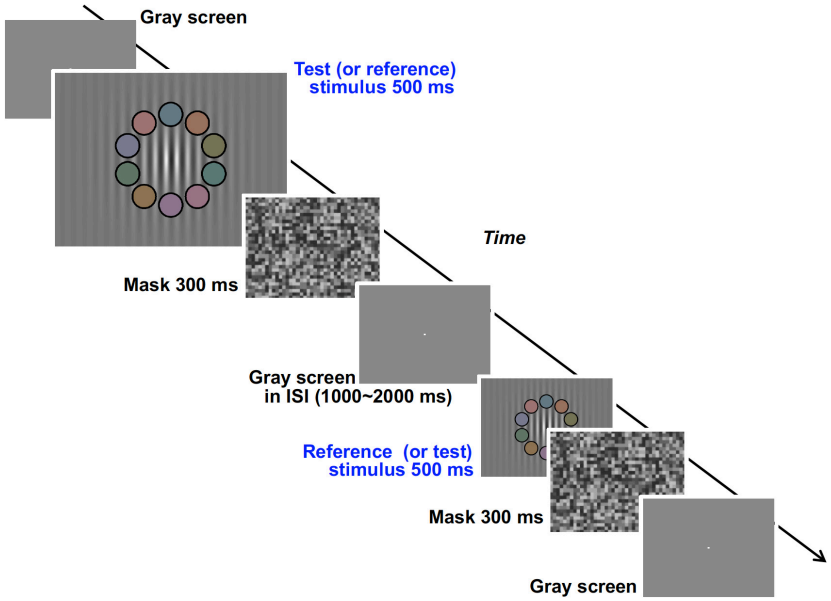


FIGURE 9 | Visual stimuli and time course of the experimental procedure in the psychophysical experiment.

In the psychophysical experiment, the signal strength of luminance information was measured indirectly from the rate of the "higher contrast" response curve with signal detection theory (details in section "Methods" and **Figure 10**). Under the assumption that the noise level for one psychophysical task is almost constant within one subject, the lower S/N ratio indicates that lower signal strength was used for the discrimination task. The results of the psychophysical experiment also indicate that the S/N ratio of luminance information was lower in the Chroma /2 condition than in other conditions. Conversely, without the achromatic background, averaged brain activity tends to correspond to the Chroma value, especially in hV4. Human visual cortices contain color selective cells, and the activity of those cells corresponds to the strength of chromatic information (the Chroma value in our experiments). Since hV4 contains a substantial number of those cells and strongly responds to chromatic visual stimuli (Livingstone and Hubel, 1988; Zeki et al., 1991; Bartels and Zeki, 2000; Wade et al., 2008; Brouwer and Heeger, 2009), it is likely that this result indicates activity of color preference cells.

Mechanism

Chromatic information is mainly processed in the ventral and lateral visual pathways (Wade et al., 2008; Hansen and Gegenfurtner, 2017), although some reports indicate that chromatic information is also utilized in the dorsal pathway (Tootell et al., 1995; Wandell et al., 1999; Takemura et al., 2016). The suppression occurred over the primary, dorsal, ventral, and lateral visual pathways. Thus, it is expected that suppression initially occurs in the early stage of the visual information process and will be propagated to higher stages not as a chromatic signal but as a luminance signal. Our results suggest that suppression occurred latest in V1. This corresponds to the results of a recent study using both Visual Evoked Potential (VEP) measurement and psychophysical tests, which indicate that the inhibitory signal for brightness and color interaction arises in a recurrent inhibitory network in V1 (Xing et al., 2015); cells responding to both luminance and color stimuli in macaque V1 have been found (Livingstone and Hubel, 1984; Johnson et al., 2001). Thus, one possible explanation for the suppression mechanism is that these cells respond to luminance during observation of achromatic stimuli, but respond to color during observation of chromatic stimuli (particularly when the color signal is weak) causing a reduction of aggregate luminance response. Another possible explanation is that chromatic information directly suppresses the response of luminance selectivity cells. In either case, there must be a determinant signal for suppression in V1 and/or higher visual cortices.

We manipulated the Munsell Chroma value of the patches and controlled the saturation of presented colors as the variable. However, the mechanism determining the strength of the suppression is still not clear. A neuron corresponding to a particular value of saturation has not been found, at least in early visual cortices; it has been considered that chromatic information, including saturation (strength of color), is coded by the input from single- and/or double- opponent cells consisting of L - M and S - (L + M) chromatically opponent systems in V1 (Lee, 2011; Shapley and Hawken, 2011). Moreover, brain activity patterns were different among hues of stimuli (Engel et al., 1997; Mullen et al., 2007; Kuriki et al., 2015). Thus, the presence and strength of suppression might be the result of computational processes for saturation using the responses of those hue-processing systems.

The suppression could also be accounted by the change of attentional states when the difference between the fMRI and psychophysical experiments is considered. In our fMRI

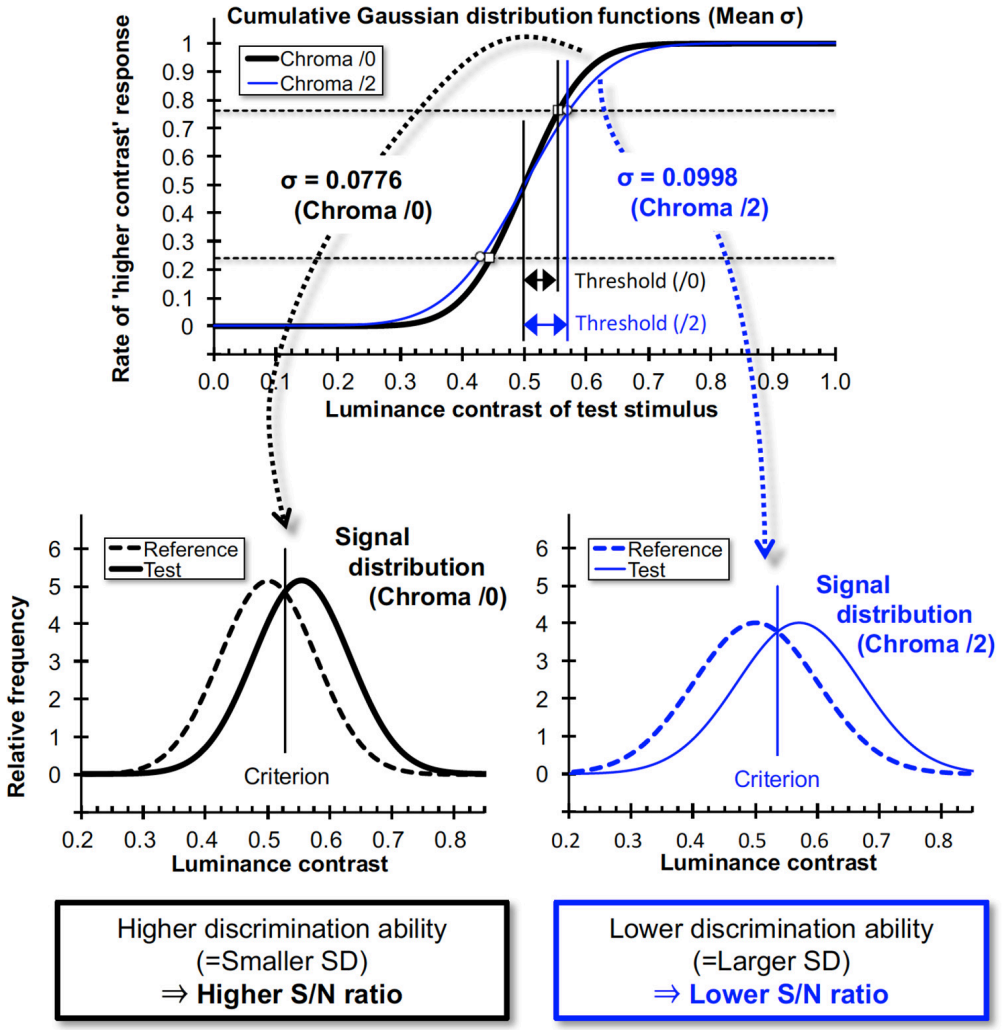


FIGURE 10 | Relationship between cumulative Gaussian distribution functions for data fits and signal distributions in signal detection theory. When the slope of the "higher contrast" response rate is high, the neural Signal/Noise ratio will be high, and vice versa. σ values here are means of all subjects for each Chroma value.

experiments the subjects were not required to attend to some specific feature in their field of vision actively. However, the colored patches in our experiments might raise bottom-up attention that modulated the BOLD responses in visual cortices. Thus, the attention effect could also have a role to attenuate the luminance signal relative to the chromatic signal, although it is hard to explain how this attenuation would be dominant only at low saturation (Chroma /2) at which the color of the patches were weak and when the subjects were not asked to observe nor to perform any tasks on these color patches. The results of a previous study (Switkes et al., 1988) suggest that the interaction between luminance and chromatic information, which was observed as a change of thresholds, occurred when active attention was not required. It may be a possible explanation is that the patches of the low saturated colors did not evoke a strong attention effect causing the interaction between luminance and chromatic information to be observed strongly in this condition, but the patches at higher saturated colors evoked strong attention effect that caused the interaction not to be observed. This hypothesis can correspond to the psychophysical results (**Figure 12**), however, it is still difficult to explain the fMRI data showing the gradual change of the suppression (**Figure 4**). In the case of the psychophysical experiment, the subjects had to judge the higher luminance contrast in a temporal two alternative forced choice task and were required to pay attention to the grating actively. It might indicate that color patches captured some amount of attentional resources from the grating unconsciously and the suppression might be reduced in the two higher saturation conditions.

Differences of Visual Stimuli Between the Experiments

Different visual stimulus patterns were used in the fMRI and psychophysical experiments because of the purpose of the experiments. In the fMRI experiment, the background pattern

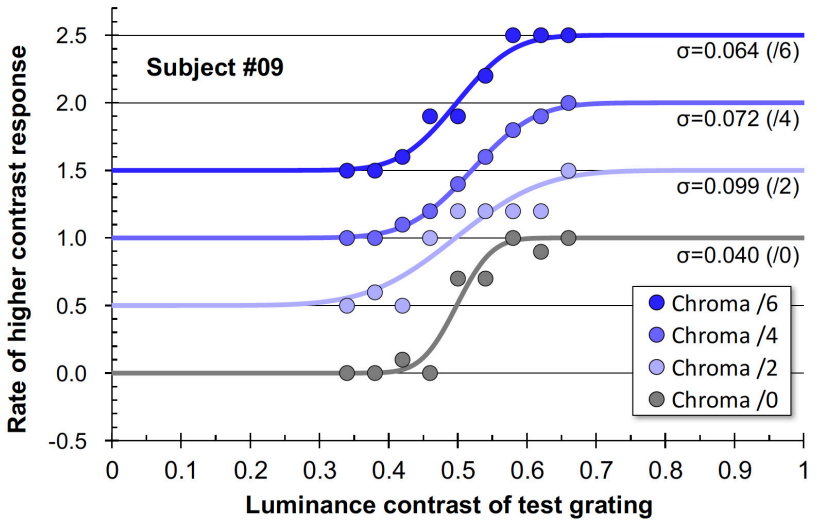


FIGURE 11 | Example data and psychometric functions for one subject under four Chroma conditions. The ordinate was defined as the data for Chroma /0 condition. Other data were vertically shifted in multiples of 0.5 to clarify. The standard deviations (**σ**) for all conditions are shown in the panel.

consisted of gray ellipses to stimulate neurons in the visual cortices that are tuned to multiple orientations and spatial frequencies. On the other hand, in the psychophysical experiment the vertical gratings with color patches were compared in a 2AFC method to select the higher luminance contrast. There were two reasons to employ different visual stimuli in the psychophysical experiment in exchange for consistency of appearance between experiments.

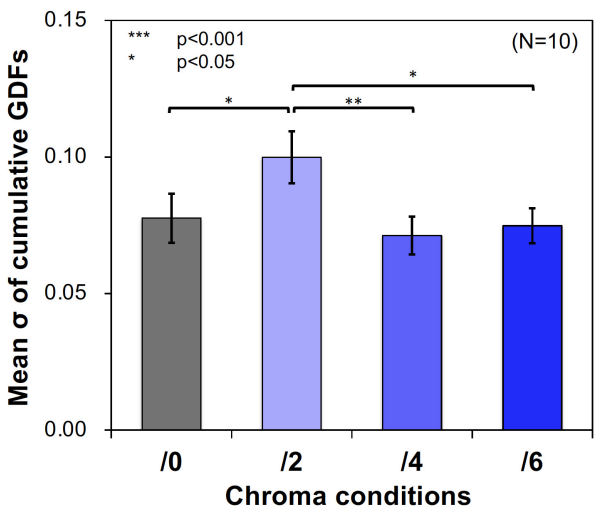


FIGURE 12 | Mean standard deviation (**σ**) of data distributions. Standard deviations were calculated using cumulative Gaussian distribution functions (GDFs) fitted to response rate data of luminance contrast discrimination under the four Chroma conditions (/0, /2, /4, and /6) in the psychophysical experiment. Error bars denote SEM. Asterisks denote statistical significance by Tukey's HSD test (**$p < 0.01$; *$p < 0.05$).

The first reason was that we thought the psychometric functions we obtained with the ellipse background pattern used in the fMRI experiment were too steep due to the existence of many luminance edges. The subjects could detect the contrast easily by simply attending to the edge of the highest contrast. Thus, the contrast at the sharp luminance edge had to be minimized to obtain a less steep psychometric function. Additionally, attending to a certain luminance edge may induce a difference in attentional state between the fMRI experiment and the psychophysical experiment. We therefore tried to design the experiment in such a way that the subject would not to pay attention to only a certain luminance edge of the visual stimulus.

The other reason was that in the psychophysical experiment, we tried to measure the luminance contrast sensitivity using psychometric functions in order to obtain the magnitude of the luminance signal. Although a psychometric function can be obtained by a complex visual stimulus (Ueda et al., 2018; Li et al., 2021), in our case we were afraid we would not be able to obtain the magnitude of the luminance signal from the function. Therefore, we thought it would be better to make the visual stimulus simple because we did not have a specific model to estimate the magnitude of the luminance signal from the psychometric functions obtained from the visual stimulus with this ellipse background pattern. For example, we were afraid of possible interactions in the contrast detection among the

TABLE 2 | ANOVA results of psychophysical experiment.

Source	Df	Sum Sq	Mean Sq	*F* value	*p*
Chroma condition	3	0.005	0.002	5.180	0.006**
Subject (DV)	9	0.015	0.002		
Residuals	27	0.009	0.000		

**$p < 0.01$.

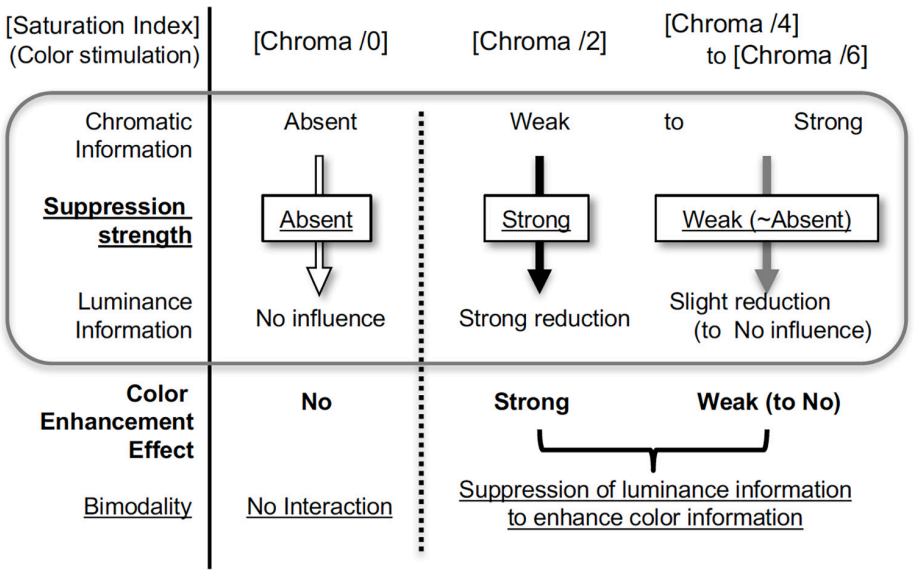


FIGURE 13 | Scheme of color enhancement model.

multiple orientations and spatial frequencies and that this ellipse background pattern would make the results unpredictable. Thus, we employed the grating pattern to minimize these interactions. Therefore, we can consider the σ values (Eq. 3) as the index for luminance contrast sensitivity of one orientation and one spatial frequency, and they can be used to predict the magnitude of the luminance signal. We consider the results obtained by the fMRI and psychophysical experiments to reflect the magnitude of the luminance responses between chromatic conditions; however, these results might be influenced by additional phenomena due to the differences in the visual stimuli.

Arguments

It is also possible to explain the psychophysical data of this study by the gamut expansion effect (Brown and MacLeod, 1997), in which strength of color is expanded or compressed from zero (neutral color) to the most saturated color in the visual stimulus. However, the gamut expansion effect on color perception can also be accounted by the aforementioned suppression hypothesis. When visual stimuli include highly saturated (high-Chroma) colors, suppression, i.e., the enhancement of perceived saturation, is weak or absent, so that perceived saturation is almost the same as presented saturation. On the other hand, when the visual stimuli are composed of desaturated (low-Chroma) colors, suppression works strongly and perceived saturation is higher than presented saturation.

Kim and Mullen (2016) reported that humans tend to use the chromatic edge as well as the luminance edge to recognize natural scenes, and they suggest there is a weak masking effect of luminance by color, and the results of our experiments also suggests that chromatic information suppresses luminance information.

The results of our psychophysical experiment were consistent with the results of previous studies in which the effect of chromatic information on luminance contrast sensitivity was investigated (Miquilini et al., 2017; Sousa et al., 2020) in that the presence of color reduces luminance contrast sensitivity. Nevertheless, the correlative relationship between contrast sensitivity and saturation of color was the opposite in our result. It can be explained by two differences between our experiments and theirs. The first is the way to measure contrast sensitivity; the detection threshold of contrast was measured in the previous studies, while our experiment measured contrast discrimination performance above the detection threshold. The second is the spatial arrangement of the visual stimuli; color and luminance information were spatially overlapped in the previous studies and separated in our experiment. Thus, we could expect that the effect of chromatic information to luminance contrast measured in our experiments depends complexly on color appearance, which may involve a possible attention effect, whereas the strength of the chromatic signal could be a dominant factor in the results of previous studies.

There are some limitations of this investigation of suppression. Firstly, our experiment couldn't reveal the hue factor of suppression, especially in terms of cone types mediating hues (Shinomori and Werner, 2012; DeLawyer et al., 2018), because the patches in our visual stimuli were distributed across hues. However, psychophysical data show that suppression occurs for all hues and the tendencies are the same across all hues (Kingdom et al., 2010). Secondly, when the Chroma value of the patches was decreased (lower than /2 and below a certain Chroma threshold) it seemed that suppression disappeared or decreased. The mechanism and behavior of suppression around the threshold point is still unclear, so detailed experimentation is essential to pursue these questions. Furthermore, investigation is needed regarding possible spatial structure, such as an effective spatial range of the suppression, since it is reasonable to expect suppression only in the region around the color stimulus

(Xing et al., 2015) to avoid decreasing luminance sensitivity over the entire visual field. If such a range exists its precise spatial structure should be determined because that spatial structure could be of use in investigations of the suppression mechanism.

CONCLUSION

The human visual system generally processes chromatic and luminance signals separately. Nevertheless, interaction between the two signal types have been suggested. We performed fMRI measurements of brain activity during exposure to visual stimuli consisting of chromatic and luminance components. Brain activity driven by luminance components was suppressed by chromatic components and the suppression was stronger when the presented chromatic stimuli were less saturated, although suppression was absent when there was no chromatic component. The psychophysical measurements of luminance contrast discrimination also support luminance signal reduction. These results directly imply that the interaction enhances chromatic information by supporting the use of weak color among visual stimuli, and explain the phenomenon that weaker-color images appear to have lower achromatic contrast.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Kochi University of Technology Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

IN and KS designed and carried out the experiments, analyzed the data, and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by JSPS KAKENHI Grant Numbers 24300085, 24650109, and 18H03323 to KS and also by Kochi University of Technology (The Focused Research Laboratory Support Grant).

ACKNOWLEDGMENTS

The authors appreciate to Tanner DeLawyer for his thoughtful comments and to Research Center for Brain Communication in Kochi University of Technology for technical supports in fMRI measurements.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.668116/full#supplementary-material

REFERENCES

Bannert, M. M., and Bartels, A. (2013). Decoding the yellow of a gray banana. *Curr. Biol.* 23, 2268–2272. doi: 10.1016/j.cub.2013.09.016

Bartels, A., and Zeki, S. (2000). The architecture of the colour centre in the human visual brain: new results and a review. *Eur. J. Neurosci.* 12, 172–193. doi: 10.1046/j.1460-9568.2000.00905.x

Brefczynski, J. A., and DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nat. Neurosci.* 2, 370–374. doi: 10.1038/7280

Brouwer, G. J., and Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex. *J. Neurosci.* 29, 13992–14003. doi: 10.1523/jneurosci.3577-09.2009

Brown, R. O., and MacLeod, D. I. (1997). Color appearance depends on the variance of surround colors. *Curr. Biol.* 7, 844–849. doi: 10.1016/s0960-9822(06)00372-1

Clery, S., Bloj, M., and Harris, J. M. (2013). Interactions between luminance and color signals: effects on shape. *J. Vis.* 13:16. doi: 10.1167/13.5.16

Clifford, C. W. G., Pearson, J., Forte, J. D., and Spehar, B. (2003). Colour and luminance selectivity of spatial and temporal interactions in orientation perception. *Vision Res.* 43, 2885–2893. doi: 10.1016/j.visres.2003.08.005

DeLawyer, T., Tayon, M., Yu, C. L., and Buck, S. L. (2018). Contrast-dependent red-green balance shifts depend on S-cone activity. *J. Opt. Soc. Am. A Opt. Image. Sci. Vis.* 35, B114–B121.

Engel, S., Zhang, X., and Wandell, B. (1997). Colour tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature* 388, 68–71. doi: 10.1038/40398

Fuller, S., and Carrasco, M. (2006). Exogenous attention and color perception: performance and appearance of saturation and hue. *Vision Res.* 46, 4032–4047. doi: 10.1016/j.visres.2006.07.014

Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25. doi: 10.1016/0166-2236(92)90344-8

Hansen, T., and Gegenfurtner, K. R. (2017). Color contributes to object-contour perception in natural scenes. *J. Vis.* 17:14. doi: 10.1167/17.3.14

Johnson, E. N., Hawken, M. J., and Shapley, R. (2001). The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nat. Neurosci.* 4, 409–416. doi: 10.1038/86061

Kim, Y. J., and Mullen, K. T. (2016). Effect of overlaid luminance contrast on perceived color contrast: shadows enhance, borders suppress. *J. Vis* 16:15. doi: 10.1167/16.11.15

Kingdom, F. A. (2003). Color brings relief to human vision. *Nat. Neurosci.* 6, 641–644. doi: 10.1038/nn1060

Kingdom, F. A., Bell, J., Gheorghiu, E., and Malkoc, G. (2010). Chromatic variations suppress suprathreshold brightness variations. *J. Vis.* 10:13. doi: 10.1167/10.10.13

Kingdom, F. A., and Kasrai, R. (2006). Colour unmasks dark target in complex displays. *Vision Res.* 46, 814–822. doi: 10.1016/j.visres.2005.08.018

Krukowski, A. E., and Stone, L. S. (2005). Expansion of direction space around the cardinal axes revealed by smooth pursuit eye movement. *Neuron* 45, 315–323. doi: 10.1016/j-neuron.2005.01.005

Kuriki, I., Sun, P., Ueno, K., Tanaka, K., and Cheng, K. (2015). Hue selectivity in human visual cortex revealed by functional magnetic resonance imaging. *Cereb. Cortex* 25, 4869–4884. doi: 10.1093/cercor/bhv198

Lee, B. B. (2011). Visual pathways and psychophysical channels in the primate. *J. Physiol.* 589(Pt 1), 41–47. doi: 10.1113/jphysiol.2010.192658

Li, F., Wang, L., Jia, L., Lu, J., Wu, Y., Wang, C., et al. (2021). The varying coherences of implied motion modulates the subjective time perception. *Front. Psychol.* 6:1407. doi: 10.3389/fpsyg.2021.602872

Livingstone, M., and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240, 740–749. doi: 10.1126/science.3283936

Livingstone, M. S., and Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *J. Neurosci.* 4, 309–356. doi: 10.1523/jneurosci.04-01-00309.1984

Loffler, G., and Orbach, H. S. (2001). Anisotropy in judging the absolute direction of motion. *Vision. Res.* 41, 3677–3692. doi: 10.1016/s0042-6989(01)00209-7

Maunsell, J. H. (1992). Functional visual streams. *Curr. Opin. Neurobiol.* 2, 506–510. doi: 10.1016/0959-4388(92)90188-q

Miquilini, L., Walker, N. A., Odigie, E. A., Guimarães, D. L., Salomão, R. C., Lacerda, E. M. C. B., et al. (2017). Influence of spatial and chromatic noise on luminance discrimination. *Sci. Rep.* 7:16944. doi: 10.1038/s41598-017-16817-0

Morrone, M. C., Denti, V., and Spinelli, D. (2002). Color and luminance contrasts attract independent attention. *Curr. Biol.* 12, 1134–1137. doi: 10.1016/s0960-9822(02)00921-1

Moutoussis, K. (2015). The physiology and psychophysics of the color-form relationship: a review. *Front. Psychol.* 6:1407. doi: 10.3389/fpsyg.2015.01407

Mullen, K. T., Dumoulin, S. O., McMahon, K. L., de Zubicaray, G. I., and Hess, R. F. (2007). Selectivity of human retinotopic visual cortex to S-cone-opponent, L/M-cone-opponent and achromatic stimulation. *Eur. J. Neurosci.* 25, 491–502. doi: 10.1111/j.1460-9568.2007.05302.x

Mullen, K. T., and Losada, M. A. (1994). Evidence for separate pathways for color and luminance detection mechanisms. *J. Opt. Soc. Am. A* 11, 3136–3151. doi: 10.1364/JOSAA.11.003136

Rajapakse, J. C., Kruggel, F., Maisog, J. M., and von Cramon, D. Y. (1998). Modeling hemodynamic response for analysis of functional MRI time-series. *Hum. Brain. Mapp.* 6, 283–300. doi: 10.1002/(sici)1097-0193(1998)6:4<283::aid-hbm7>3.0.co;2-#

Shapley, R., and Hawken, M. J. (2011). Color in the cortex: single- and double-opponent cells. *Vision. Res.* 51, 701–717. doi: 10.1016/j.visres.2011.02.012

Shinomori, K., Nakano, Y., and Uchikawa, K. (1994). Influence of the illuminance and spectral composition of surround fields on spatially induced blackness. *J. Opt. Soc. Am. A. Opt. Image. Sci. Vis* 11, 2383–2388. doi: 10.1364/josaa.11.002383

Shinomori, K., Schefrin, B. E., and Werner, J. S. (1997). Spectral mechanisms of spatially-induced blackness: Data and quantitative model. *J. Opt. Soc. Am. A. Opt. Image. Sci. Vis* 14, 372–387. doi: 10.1364/josaa.14.000372

Shinomori, K., and Werner, J. S. (2012). Aging of human short-wave cone pathways. *Proc. Natl. Acad. Sci. U.S.A* 109, 13422–13427. doi: 10.1073/pnas.1119770109

Somers, D. C., Dale, A. M., Seiffert, A. E., and Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A* 96, 1663–1668. doi: 10.1038/7280

Sousa, B. R. S., Loureiro, T. M. G., Goulart, P. R. K., Cortes, M. I. T., Costa, M. F., Bonci, D. M. O., et al. (2020). Specificity of the chromatic noise influence on the luminance contrast discrimination to the color vision phenotype. *Sci. Rep.* 10:17897. doi: 10.1038/s41598-020-74875-3

Switkes, E., Bradley, A., and Valois, K. K. D. (1988). Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings. *J. Opt. Soc. Am. A* 5, 1149–1162. doi: 10.1364/josaa.5.001149

Takemura, H., Rokem, A., Winawer, J., Yeatman, J. D., Wandell, B. A., and Pestilli, F. (2016). A major human white matter pathway between dorsal and ventral visual cortex. *Cereb. Cortex* 26, 2205–2214. doi: 10.1093/cercor/bhv064

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J. Neurosci.* 15, 3215–3230. doi: 10.1523/jneurosci.15-04-03215.1995

Turatto, M., and Galfano, G. (2000). Color, form and luminance capture attention in visual search. *Vision. Res.* 40, 1639–1643. doi: 10.1016/s0042-6989(00)00061-4

Ueda, H., Yamamoto, K., and Watanabe, K. (2018). Contribution of global and local biological motion information to speed perception and discrimination. *J. Vis.* 18:2. doi: 10.1167/18.3.2

Victor, J. D., Purpura, K. P., and Conte, M. M. (1998). Chromatic and luminance interactions in spatial contrast signals. *Vis. Neurosci.* 15, 607–624. doi: 10.1017/s0952523898154032

Wade, A., Augath, M., Logothetis, N., and Wandell, B. (2008). fMRI measurements of color in macaque and human. *J. Vis.* 8:6. doi: 10.1167/8.10.6

Wandell, B. A., Dumoulin, S. O., and Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron* 56, 366–383. doi: 10.1016/j.neuron.2007.10.012

Wandell, B. A., Poirson, A. B., Newsome, W. T., Baseler, H. A., Boynton, G. M., Huk, A., et al. (1999). Color signals in human motion-selective cortex. *Neuron* 24, 901–909. doi: 10.1016/s0896-6273(00)81037-5

Xiao, B., and Wade, A. R. (2010). Measurements of long-range suppression in human opponent S-cone and achromatic luminance channels. *J. Vis.* 10:10. doi: 10.1167/10.13.10

Xing, D., Ouni, A., Chen, S., Sahmoud, H., Gordon, J., and Shapley, R. (2015). Brightness–color interactions in human early visual cortex. *J. Neurosci.* 35, 2226–2232. doi: 10.1523/JNEUROSCI.3740-14.2015

Yeatman, J. D., Weiner, K. S., Pestilli, F., Rokem, A., Mezer, A., and Wandell, B. A. (2014). The vertical occipital fasciculus: a century of controversy resolved by in vivo measurements. *Proc. Natl. Acad. Sci. U.S.A* 111, E5214–E5223. doi: 10.1073/pnas.1418503111

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., and Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *J. Neurosci* 11, 641–649. doi: 10.1523/jneurosci.11-03-00641.1991

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 05 July 2021
doi: 10.3389/fnins.2021.671121

Simulating Visibility and Reading Performance in Low Vision

Ying-Zi Xiong[1], Quan Lei[1,2], Aurélie Calabrèse[3] and Gordon E. Legge[1]*

[1] Department of Psychology, University of Minnesota, Minneapolis, MN, United States, [2] Department of Psychology, Wichita State University, Wichita, KS, United States, [3] Inria, Université Côte d'Azur, Sophia Antipolis, France

Purpose: Low vision reduces text visibility and causes difficulties in reading. A valid low-vision simulation could be used to evaluate the accessibility of digital text for readers with low vision. We examined the validity of a digital simulation for replicating the text visibility and reading performance of low-vision individuals.

Methods: Low-vision visibility was modeled with contrast sensitivity functions (CSFs) with parameters to represent reduced acuity and contrast sensitivity. Digital filtering incorporating these CSFs were applied to digital versions of the Lighthouse Letter Acuity Chart and the Pelli-Robson Contrast Sensitivity Chart. Reading performance (reading acuity, critical print size, and maximum reading speed) was assessed with filtered versions of the MNREAD reading acuity Chart. Thirty-six normally sighted young adults completed chart testing under normal and simulated low-vision conditions. Fifty-eight low-vision subjects (thirty with macular pathology and twenty-eight with non-macular pathology) and fifteen normally sighted older subjects completed chart testing with their habitual viewing. We hypothesized that the performance of the normally sighted young adults under simulated low-vision conditions would match the corresponding performance of actual low-vision subjects.

Results: When simulating low-vision conditions with visual acuity better than 1.50 logMAR (Snellen 20/630) and contrast sensitivity better than 0.15 log unit, the simulation adequately reduced the acuity and contrast sensitivity in normally sighted young subjects to the desired low-vision levels. When performing the MNREAD test with simulated low vision, the normally sighted young adults had faster maximum reading speed than both the Non-macular and Macular groups, by an average of 0.07 and 0.12 log word per minute, respectively. However, they adequately replicated the reading acuity as well as the critical print size, up to 2.00 logMAR of both low-vision groups.

Conclusion: A low-vision simulation based on clinical measures of visual acuity and contrast sensitivity can provide good estimates of reading performance and the accessibility of digital text for a broad range of low-vision conditions.

Keywords: reading, low vision, text visibility, visual acuity, contrast sensitivity

Edited by:
Fang Hou,
Wenzhou Medical University, China

Reviewed by:
Pete R Jones,
City University of London, United Kingdom
Allen Ming Yan Cheong,
Hong Kong Polytechnic University, China

****Correspondence:***
Ying-Zi Xiong
yingzi@umn.edu

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *23 February 2021*
Accepted: *09 June 2021*
Published: *05 July 2021*

Citation:
Xiong Y-Z, Lei Q, Calabrèse A and Legge GE (2021) Simulating Visibility and Reading Performance in Low Vision. Front. Neurosci. 15:671121. doi: 10.3389/fnins.2021.671121

INTRODUCTION

Low vision refers to any vision impairment that cannot be corrected by glasses or contact lenses. For readers with low vision, text legibility is limited by acuity and contrast sensitivity. In practical terms, reduced acuity and contrast sensitivity limit the ability to see graphics and text on web pages and in other digital formats. Other factors affecting vision, such as field loss, light level and glare, often add to the difficulties in low-vision function (Fletcher et al., 1999; Turano et al., 2004; Kiser et al., 2005). While it is not always sufficient for successful low-vision functioning, the visibility of key features is usually a necessary condition for low-vision functioning. The goal of our project is to validate a simulation of the loss of visibility due to reduced acuity and reduced contrast sensitivity. The simulation is based on image filtering that uses transformations of the normal contrast sensitivity function (CSF) to represent reduced visibility associated with low vision. We evaluated the validity of the simulation by testing normally sighted subjects on filtered images of text to determine if measures of acuity, contrast sensitivity and reading performance match the performance of people with actual low vision. A valid simulation of low-vision visibility could be useful to eye-care clinicians, display designers, website creators, and family members in evaluating the accessibility of digital rendering of text or graphics for people with low vision.

Low-vision simulations, such as diffusive filters, optical defocus and digital blur, have been utilized for research or education purposes (Peli, 1990; Dickinson and Rabbitt, 1991; Bowers and Reid, 1997; Thompson et al., 2017; Jones and Ometto, 2018; Jones et al., 2020). A desirable property of an digital simulation is that it can be parameterized by measurable properties of vision status such as acuity and contrast sensitivity (Peli, 1990; Thompson et al., 2017).

The CSF is a detailed measurement of an individual's acuity limit and contrast sensitivity across a range of spatial frequencies (Campbell and Robson, 1968), which determines the visibility of any pattern. Compared to people with normal vision, people with low vision often have reduced contrast sensitivity and a decreased range of visible spatial frequencies (Ross et al., 1984; Sokol et al., 1985; Chylack et al., 1993). Peli described a methodology using low-vision CSF filters to process images to represent the reduction in sensitivity of low-vision eyes (Peli, 1990). A key assumption of the method is that target features in the original image that are not visible or recognizable with specific levels of low vision are not visible or recognizable to normally sighted subjects viewing the filtered image.

It is difficult in practice to directly measure CSFs for people with low vision, although recent development of a quick CSF measurement facilitates such measurement (Lesmes et al., 2010; Elfadaly et al., 2020). Another approach is to derive low-vision CSFs from a typical CSF for normal vision. Chung and Legge (2016) proposed that low vision CSFs can be approximated by horizontal and/or vertical scaling of a normal vision CSF template, with the horizontal scaling representing the loss in high spatial frequency resolution, and the vertical scaling representing the loss in peak contrast sensitivity (Chung and Legge, 2016).

Recent studies have further shown that the horizontal and vertical scaling factors for deriving the low-vision CSF can be estimated by clinical measures of visual acuity and contrast sensitivity (Thurman et al., 2016; Thompson et al., 2017). Specifically, clinical testing tools such as letter acuity charts [e.g., the Early Treatment of Diabetic Retinopathy (ETDRS) chart] and letter contrast sensitivity charts (e.g., the Pelli-Robson Chart), were designed to provide convenient measures of individual visual acuity and contrast sensitivity. These measures provide reasonable estimations of the high spatial frequency resolution and the peak contrast sensitivity of the individual's CSF curve (Thurman et al., 2016; Thompson et al., 2017). Using the filtering method proposed by Peli (1990), Thompson and colleagues (Thompson et al., 2017) parameterized their low-vision filters using these clinical measures in an attempt to simulate visibility experienced by individuals with reduced acuity and contrast sensitivity. Their simulation was validated by a letter recognition task, showing that the measured acuity for filtered letters closely matched their intended visibility as specified by the filter parameters. Despite the potential usefulness of the method, it is unknown whether the method can also be used to simulate the impact of reduced visibility on more complex tasks such as reading.

A primary goal of the current study was to examine the validity of the CSF filtering method for predicting visual performance in a task beyond simple visibility. We simulated the reading performance of people with low vision. Following Peli (1990) and Thompson et al. (2017), we embedded an estimate of the reader's CSF in the simulation filter. The implementation included two key steps: (1) clinical acuity and contrast sensitivity measured by letter charts were used to estimate the scaling factors used to derive the low-vision CSF; and (2) the low-vision CSF thus derived was used to filter the input image to generate the simulation.

To summarize, the current study was aimed to extend previous work by using clinical measures of acuity and contrast sensitivity to parameterize the simulation method and to systematically validate the method by examining the impact of simulated low vision on both simple tasks such as letter recognition and complex tasks such as reading. Specifically, we asked two main questions: (1) Do normally sighted subjects tested with filtered images of the letter charts show reduced acuity and contrast sensitivity close to the simulated low-vision levels? And (2) Do the reduced acuity and contrast sensitivity have the same impact on reading as real low vision? To this end, we compared the reading performance of normally sighted subjects, tested with simulated reduction of acuity and contrast sensitivity, with the performance of low-vision subjects with the equivalent acuity and contrast sensitivity. We also examined whether two other factors beyond acuity and contrast sensitivity, namely age and central vision status, need to be considered in the simulation. It has been well studied that people with central field loss due to macular diseases have greater difficulty in reading (Legge et al., 1992), therefore we included low-vision groups with non-macular and macular diseases, to compare the validity of our simulation for low vision with or without central vision disturbance. We included a group of normally sighted older

subjects, to examine the need for age adjustment when simulating older low-vision individuals.

MATERIALS AND METHODS

Subjects

One hundred and nine subjects participated in this study. All subjects were native English speakers with no known visual reading disabilities. Normal cognitive status was verified by the Mini-Mental State Examination (score > 24). All subjects were tested with their most up-to-date reading glasses, if any.

Thirty-six of the subjects were normally sighted young adults (YN, 20.5 ± 3.6 years) recruited from the University of Minnesota. Fifteen of the subjects were normally sighted older adults (ON, 68.0 ± 5.0 years) recruited from the Retiree Volunteer Center at the University of Minnesota. Fifty-eight of the subjects (64.8 ± 18.0 years) were adults with low vision whose data were included from two published studies (Cheong et al., 2008; Calabrèse et al., 2018). The low-vision data were separated into macular disease (Mac, n = 30) and non-macular disease (Non-Mac, n = 28) groups based on whether the diagnoses primarily affected the macular area (see **Supplementary Appendix 2** for individual diagnoses). This study was approved by the University of Minnesota Institutional Review Board and followed the Declaration of Helsinki. Consent forms were acquired from all subjects prior to their participation.

Apparatus and Stimuli

Digital versions of the Lighthouse Letter Acuity Chart, Pelli-Robson Contrast Sensitivity Chart and MNREAD Chart were adapted from the original printed charts (Ferris et al., 1982; Pelli et al., 1988; Mansfield and Legge, 2007), using Psychtoolbox 3.0 software (Pelli, 1997) with Matlab R2016a. In the digital acuity test, a group of five letters was presented on the screen each time, equivalent to a single line on the printed chart. In the digital contrast sensitivity test, a group of three letters was presented on the screen each time, equivalent to a single contrast level on the printed chart. The MNREAD sentences were created by a MNREAD sentence generator (Mansfield et al., 2019). Each MNREAD chart had 21 sentences with decreasing sizes in 0.1 log unit steps from 1.7 logMAR to −0.3 logMAR (equivalent to a range of x-heights from 4.18 to 0.04 degree). Each sentence was formatted on three equal-length lines like the printed MNREAD chart. Only one sentence was presented on the screen at one time.

A large LCD monitor was used (dimensions = 59.6 × 33.4 cm) to ensure the presentation of large size letters (Cinema Display, Apple, Inc.). The refresh rate was 60 Hz and the resolution was 2,560 × 1,440. Stimuli were displayed with 14-bit grayscale resolution using Bits++ (Cambridge Research Systems Ltd., United Kingdom). The output luminance of the monitor at each gray level was measured using a photometer (PR655 Spectroradiometer, Photo Research Inc.), and a look-up table was created to present letters at each contrast level. The white background had a fixed luminance of 298.5 cd/m^2. For the Lighthouse Letter Acuity Chart and MNREAD Chart, the high-contrast black letters had a fixed luminance of 1.5 cd/m^2. For the Pelli-Robson Contrast Sensitivity Chart, the luminance of the sixteen three-letter groups ranged from 1.5 to 296.8 cd/m^2.

The viewing distance was 100 cm, with the exception that the small print sizes (<0 logMAR) on the Lighthouse Letter Acuity charts and MNREAD charts were tested at 160 cm to ensure adequate resolution. To change the viewing distance, the test was paused and subjects were moved back from 100 to 160 cm.

CSF Filters

In **Figure 1A**, the black curve illustrates a normal CSF template, with the y-axis representing contrast sensitivity and x-axis representing spatial frequency. The CSF was constructed based on Barten's simplified CSF formula (Equation 1, Barten, 1999, 2003). In Equation 1, $S_{\mathrm{NV}}(f)$ is the contrast sensitivity at spatial frequency f, equivalent to the inverse of the corresponding Michelson contrast at threshold. There are two free parameters: the luminance (L) of the image and the angular area ($X_0{}^2$) of the picture area. The luminance was fixed as the mean luminance of the screen (150 cd/m^2), and the image area was fixed as the angular area of the screen (33 × 19 deg^2).

$$S_{NV}(f) = \frac{5200e^{-0.0016f^2(1+100/L)^{0.08}}}{\sqrt{\left(1+\frac{144}{X_0^2}+0.64f^2\right)\times\left(\frac{63}{L^{0.83}}+\frac{1}{1-e^{-0.02f^2}}\right)}} \quad (1)$$

Low-vision CSF curves were created by shifting the normal CSF curve horizontally along the spatial frequency axis by factor a and vertically along the contrast sensitivity axis by factor b (Equation 2; Chung and Legge, 2016). The scaling corresponds to horizontal and vertical translations of the normal template in the log-log coordinates of **Figure 1A**. The red curves in **Figure 1A** provide three low-vision CSF examples of different combinations of horizontal and vertical scaling. Note that in some conditions the shifted low-vision sensitivities at lower spatial frequencies would exceed that of the normal vision (**Figure 1A**, gray dashed curves). To avoid this problem the low-vision CSF was clamped at the value of the normal CSF.

$$S_{LV}(f) = \min\left(bS_{NV}\left(\frac{f}{a}\right),\ S_{NV}(f)\right) \quad (2)$$

The CSF filter is constructed by computing the attenuation in spatial frequency components of the input image due to contrast sensitivity loss across the low-vision CSF relative to the normal CSF. It is defined as the ratio between a low-vision CSF and the normal CSF (Equation 3, **Figure 1B**).

$$F(f) = \frac{S_{LV}(f)}{S_{NV}(f)} \quad (3)$$

The CSF filters can be applied to digital texts and pictures to simulate pattern visibility to the corresponding low-vision eyes. Specifically, the amplitude of the Fourier transform of the input image at each spatial frequency is multiplied by the corresponding value of the filter function to achieve spatial-frequency specific attenuation, and then an inverse Fourier transform is applied to create the filtered image. **Figure 1C** shows

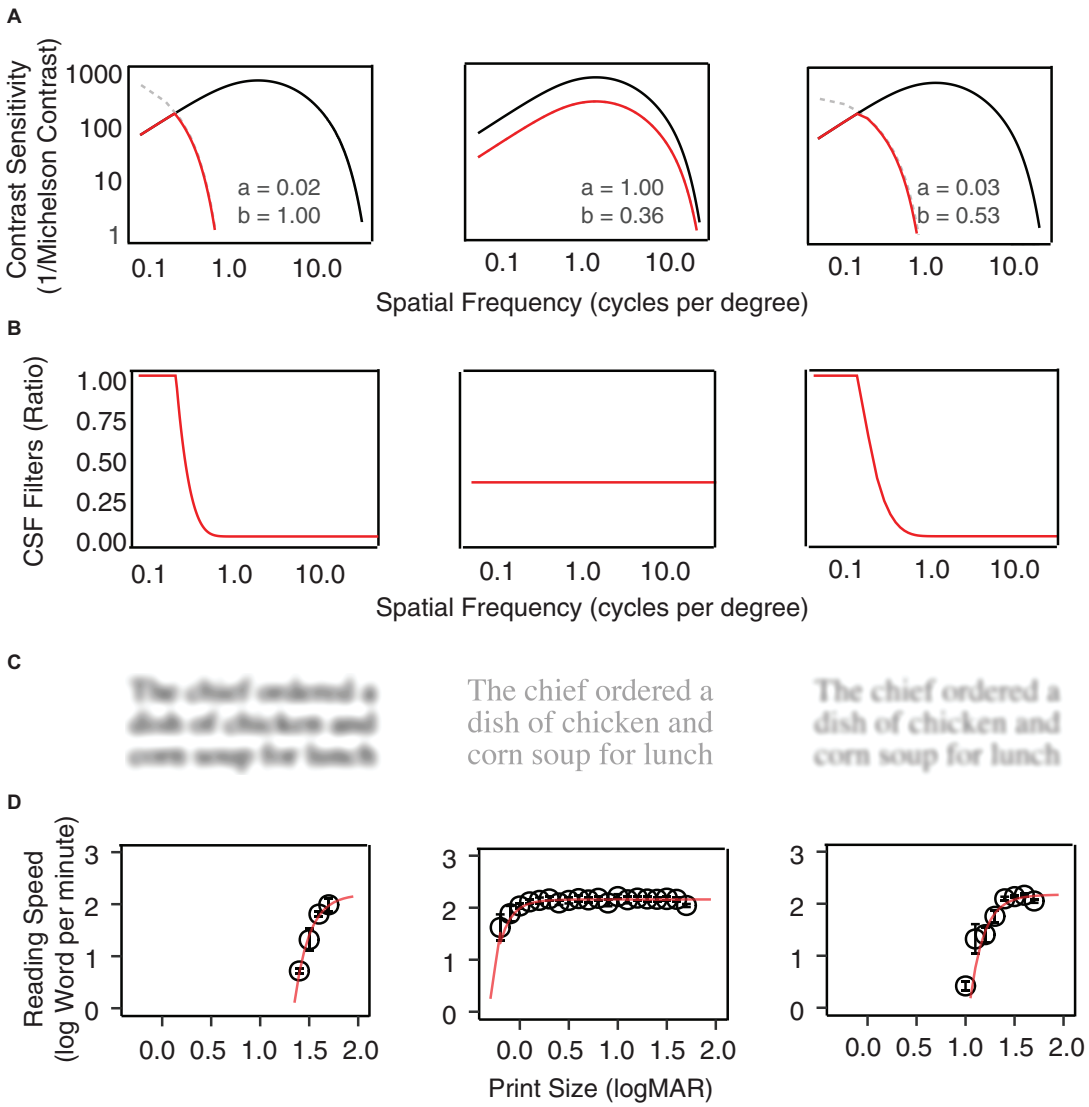


FIGURE 1 | Examples of CSF filters. **(A)** Normal vision CSFs are represented with black curves, and the low vision CSFs are represented by red curves with horizontal and vertical scaling of the normal CSF. From left to right, the plots show examples of horizontal scaling, vertical scaling, and horizontal-plus-vertical scaling conditions. The scaling factors are listed in each plot. **(B)** The CSF filters are defined by the ratio between the low vision and normal vision CSF in **(A)**. **(C)** A MNREAD sentence filtered by the three CSF filters. **(D)** Reading speed (log word per minute) as a function of print size under the three conditions.

examples of a MNREAD sentence after filtering by three CSF filter conditions. **Figure 1D** shows the average reading curves under each of the three simulated conditions.

As an aside, we comment on a methodological difference in the implementation of the CSF filtering between the current study and the previous studies of Peli (1990) and Thompson et al. (2017). Specifically, in the two previous studies, a visual image was decomposed into a discrete set of frequency bands. A contrast threshold was then derived from the low-vision CSF for each frequency band and applied to the corresponding sub-image to completely eliminate the image contents with sub-threshold contrasts. This non-linear filtering approach is particularly suitable for the simulation of the appearance of complex images where local contrast plays a vital role in pattern perception. However, One issue with this approach is the noticeable artifacts (i.e., banding or ringing effects) generated in the filtered images due to the use of non-linear hard-thresholding. Although a solution has been proposed by Thompson et al. (2017) to minimize the artifacts, they are still visible and can be distracting in deciphering letters in text.

In the current study, we adopted an alternative linear approach, using a single-channel filter based on the CSF, that does not involve decomposing the entire frequency range into a discrete set of frequency bands and no explicit thresholding is performed. Our previous work has preliminarily validated this approach for simulating low-vision visibility (Lei et al., 2016). In this approach, the ratios of low-vision and normal-vision contrast sensitivities at all spatial frequencies spanning the CSFs were

calculated as the filter to represent the loss of contrast sensitivities in low vision relative to normal vision. The filter was then used to linearly scale the spatial frequency contents of an image, such that each frequency component was attenuated by an amount that is commensurate with the relative loss in contrast sensitivity of a low-vision observer at that frequency. The linear approach results in filtered images of text that are virtually free of artifacts. Linear filtering is also simpler to implement with fewer parameter settings than needed for the sub-band thresholding implemented in the non-linear method.

Simulated Low-Vision Conditions

A close association can be established between the scaling factors (a and b) used in the simulation of low vision and the corresponding visual acuity (VA) and contrast sensitivity (CS) values we intend to simulate. Briefly, VA provides an estimation of the high frequency cut-off of the corresponding CSF, and CS provides an estimation of the peak contrast sensitivity of the corresponding CSF. For people with low vision, the reductions in their VA and CS compared to the normal baselines can therefore be associated with the horizontal and vertical scaling factors. For purposes of our modeling, the normal baseline acuity was −0.24 logMAR, corresponding to the high-frequency cutoff of the normal CSF, and the normal baseline value for CS was 2.13 log units, corresponding to the mean Pelli-Robson score of our YN subjects (see footnotes in **Table 1**). **Supplementary Appendix 1** describes the transformations relating the scaling factors *a* and *b* to measured values of VA and CS. The parameterization procedure is similar in logic to that of Thompson et al. (2017) but different in implementation due to the adoption of a different functional form for the CSF.

Forty hypothetical low-vision conditions were simulated using different combinations of horizontal and vertical scaling. The scaling factors *a* and *b* used in the forty hypothetical low-vision conditions are listed in **Table 1** and illustrated in **Figure 2A**.

Twenty-five of the low-vision conditions (**Figure 2A**, filled circles; **Table 1**, Filter 1–25) were determined based on the empirical relationship between VA and CS (adapted from Xiong et al., 2020). Specifically, across a large sample of subjects (N = 1,040) including those with normal ocular health and various ocular pathologies, the reductions in VA and CS compared to normal baselines were significantly correlated following a linear relationship (the regression line and confidence intervals are presented in **Figure 2B**). We first determined five hypothetical low-vision conditions corresponding to 0.6, 0.9, 1.2, 1.5, and 1.8 logMAR reductions compared to normal VA. For each level of VA reduction, five levels of CS reductions were determined by steps of 0.25 log unit, with the middle level centered approximately at the regression line (**Figure 2B**, black dots). The remaining fifteen low-vision conditions (**Figure 2A**, open circles; **Table 1**, Filter 26–40) were retrospectively included to supplement the boundary conditions.

Procedure

All tests were conducted under binocular viewing. Each YN subject was tested with a baseline condition where no filtering was applied, and between 10 and 16 simulated low-vision conditions. VA, CS, and reading performance were measured under each condition, using digital versions of the Lighthouse Letter Acuity charts, Pelli-Robson Contrast Sensitivity charts and the MNREAD charts, respectively. The ON, Mac, and Non-Mac groups also completed the three tests, under the no filtering condition only.

All the testing and scoring followed the standard protocols for the tests. VA was scored on a letter-by-letter basis with each letter worth 0.02 logMAR (Ferris et al., 1982), and CS scored as the log value of the lowest contrast at which subjects can correctly report at least 2 letters in a triplet (Pelli et al., 1988).

Reading speed in log word per minute (log wpm) was obtained at each tested print size. Reading speed as a function of print size was fitted with a function (Equation 4) by non-linear mixed-effects (NLME) modeling, in which subject variations were modeled as random effects (Cheung et al., 2008).

$$Reading\ Speed = mrs \times (1 - e^{(-e^{lrc} \times (Print\ Size - xint))}) \quad (4)$$

where mrs is the plateau of the reading curve, lrc is the slope of the reading curve, and xint is the intercept of the reading curve with x-axis. Three standard reading indices were derived from each fitted curve:

- Maximum reading speed: the fastest reading speed subjects can achieve. Calculated as the asymptote of the fitted exponential curve.
- Critical print size: the smallest print size yielding the maximum reading speed. Calculated as the print size corresponding to a reading speed of 90% of the maximum reading speed.
- Reading acuity: the smallest print size that can just be read. Reading acuity = smallest print size attempted + number of errors × 0.01.

Table 2 provides a summary of the VA, CS, and reading indices for each group.

Statistical Analysis

The statistical analyses were performed using the R package (R Core Team, 2018). When examining the validity of simulating reduced VA and CS, two Linear Mixed Effects (LME) models (Pinheiro and Bates, 2000) were conducted on the VA or CS values, with value types (expected and measured) and filter conditions as fixed factors and subject as a random effect. In addition, we used the test-retest reliabilities (95% coefficient of repeatability) of VA (0.20 logMAR) and CS (0.30 log unit) for low vision as the criterion of clinically significant difference (Kiser et al., 2005).

Three LME models were fit to describe the impact of simulated VA and CS reduction on reading performance (maximum reading speed, critical print size, reading acuity), respectively. Specifically, the models treated the reading indices as the dependent variable, expected VA and CS as the fixed effect factors, and subject and filter as random effects. For maximum reading speed, an additional Non-linear Mixed Effect (NLME) model

(Bates et al., 2015) was fit to further quantify the impact of simulated CS reduction.

Lastly, when examining the validity of the simulation in predicting reading performance in low-vision subjects and older control subjects, again LME models were fit to compare the predicted and actual reading performance (maximum reading speed, critical print size and reading acuity). The models treated the reading indices as the dependent variable, condition (predicted vs. actual values) and group (Non-Mac, Mac, and ON) as fixed effects, and subject as a random effect.

For all the LME analyses described above, the significance of each fixed factor was examined by the ANOVA

TABLE 1 | Simulated low-vision conditions: scaling factors (a and b), expected VA and CS Reductions, expected VA and CS values, and measured VA and CS values (mean [standard deviation]) in the normally sighted young group.

Filter	Horizontal scaling *a*	Vertical scaling *b*	Expected VA reduction	Expected CS reduction	Expected VA	Expected CS	Measured VA	Measured CS
0*	1.000	1.000	0.00	0.00	−0.24†	2.13‡	−0.09 [0.02]	2.13 [0.01]
1	0.288	0.288	0.60	−0.54	0.36	1.59	0.28 [0.01]	1.58 [0.02]
2	0.157	0.157	0.90	−0.80	0.66	1.33	0.62 [0.01]	1.33 [0.02]
3	0.086	0.086	1.20	−1.06	0.97	1.07	0.97 [0.01]	0.99 [0.02]
4	0.048	0.048	1.51	−1.31	1.27	0.82	1.31 [0.02]	0.64 [0.02]
5	0.027	0.027	1.81	−1.54	1.57	0.59	1.72 [0.01]	0.10 [0.02]
6	0.250	1.000	0.60	−0.07	0.36	2.05	0.25 [0.02]	2.11 [0.02]
7	0.134	0.534	0.90	−0.31	0.66	1.81	0.56 [0.01]	1.80 [0.03]
8	0.072	0.288	1.20	−0.57	0.97	1.55	0.93 [0.02]	1.44 [0.02]
9	0.039	0.157	1.51	−0.83	1.27	1.29	1.27 [0.03]	0.95 [0.07]
10	0.022	0.086	1.81	−1.09	1.57	1.03	1.67 [0.02]	0.35 [0.04]
11	0.267	0.534	0.60	−0.27	0.36	1.86	0.26 [0.01]	1.91 [0.04]
12	0.144	0.288	0.90	−0.54	0.66	1.59	0.57 [0.01]	1.60 [0.03]
13	0.078	0.157	1.20	−0.80	0.97	1.33	0.96 [0.02]	1.23 [0.03]
14	0.043	0.086	1.51	−1.06	1.27	1.07	1.30 [0.02]	0.72 [0.03]
15	0.024	0.048	1.81	−1.31	1.57	0.82	1.68 [0.01]	0.27 [0.03]
16	0.314	0.157	0.60	−0.80	0.36	1.33	0.30 [0.02]	1.26 [0.07]
17	0.172	0.086	0.90	−1.06	0.66	1.07	0.65 [0.03]	1.01 [0.06]
18	0.096	0.048	1.20	−1.31	0.97	0.82	1.01 [0.02]	0.69 [0.05]
19	0.055	0.027	1.51	−1.54	1.27	0.59	1.40 [0.04]	0.36 [0.03]
20	0.032	0.016	1.81	−1.76	1.57	0.42	1.77 [0.01]	0.00 [0.00]
21	0.345	0.086	0.60	−1.06	0.36	1.10	0.35 [0.02]	1.07 [0.04]
22	0.193	0.048	0.90	−1.31	0.66	0.85	0.69 [0.01]	0.79 [0.04]
23	0.110	0.027	1.20	−1.54	0.97	0.62	1.01 [0.03]	0.64 [0.05]
24	0.064	0.016	1.51	−1.76	1.27	0.40	1.38 [0.05]	0.30 [0.04]
25	0.038	0.010	1.81	−1.96	1.57	0.14	1.57 [0.01]	0.15 [0.00]
26	0.439	0.027	0.60	−1.54	0.36	0.58	0.50 [0.04]	0.62 [0.07]
27	0.256	0.016	0.90	−1.76	0.66	0.36	0.85 [0.04]	0.47 [0.05]
28	0.154	0.010	1.20	−1.96	0.97	0.16	1.31 [0.07]	0.29 [0.07]
29	0.033	0.534	1.51	−0.55	1.27	1.61	1.32 [0.03]	1.50 [0.04]
30	0.018	0.288	1.81	−0.81	1.57	1.35	1.63 [0.02]	0.77 [0.03]
31	0.125	1.000	0.90	−0.16	0.66	1.96	0.64 [0.02]	2.06 [0.04]
32	0.063	1.000	1.20	−0.29	0.97	1.85	0.97 [0.02]	1.93 [0.02]
33	0.031	1.000	1.51	−0.45	1.27	1.67	1.33 [0.01]	1.52 [0.08]
34	0.016	1.000	1.81	−0.63	1.57	1.49	1.55 [0.01]	0.69 [0.09]
35	1.000	0.355	0.05	−0.45	−0.19	1.65	−0.12 [0.01]	1.70 [0.03]
36	1.000	0.178	0.09	−0.75	−0.15	1.35	−0.07 [0.02]	1.43 [0.03]
37	1.000	0.089	0.14	−1.04	−0.10	1.06	−0.02 [0.03]	1.10 [0.03]
38	1.000	0.045	0.20	−1.34	−0.04	0.76	0.07 [0.03]	0.88 [0.03]
39	1.000	0.022	0.27	−1.63	0.03	0.47	0.17 [0.04]	0.58 [0.03]
40	1.000	0.011	0.36	−1.90	0.13	0.20	0.45 [0.09]	0.36 [0.04]

**Filter 0 represents no filtering condition.*

†The baseline acuity value was directly obtained from the cut-off spatial frequency (51.9 cpd) of the normal CSF function, using Equation 11 in ***Supplementary Appendix 2****.*

‡The baseline Pelli-Robson contrast sensitivity value obtained from the peak of the normal CSF function is 2.5 log units, which exceeds the test capacity of the Pelli-Robson chart and therefore was normalized to the baseline value of our subject pool.

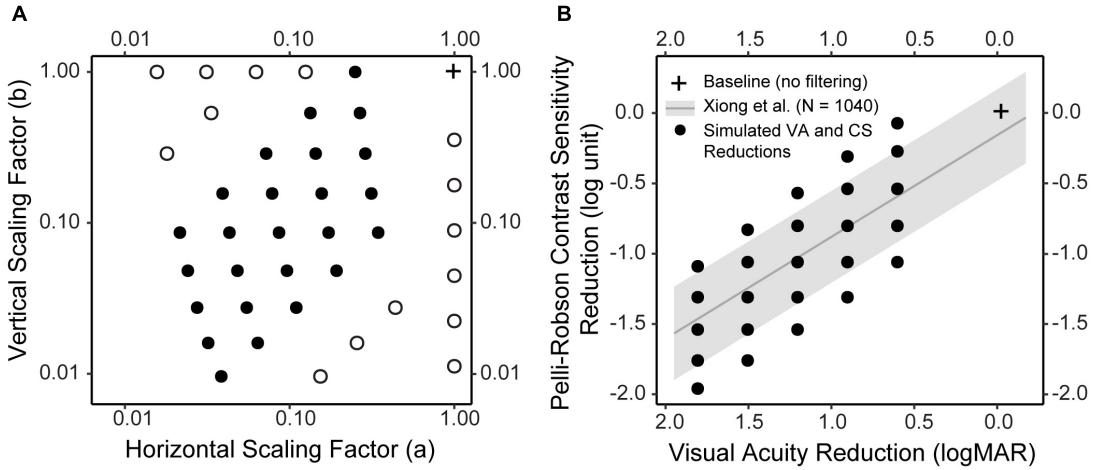


FIGURE 2 | Simulated low vision conditions. **(A)** The horizontal and vertical scaling factors of the forty simulated low vision conditions. Twenty-five conditions (solid dot) were determined so that the reductions in CS and VA would closely resemble the empirical distribution (see **B** for details). Fifteen conditions (empty dot) were boundary conditions. **(B)** The empirical distribution of CS reduction as a function of VA reduction, both compared to Age-adjusted normal baselines, in a large sample of subjects (N = 1,040) with various vision conditions (adapted from Xiong et al., 2020). The gray line represents the regression line, and the gray ribbon represents the 95% confidence interval around the regression line. The black solid dots represent the CS and VA reductions in the 25 simulations. Note that in **B** the VA axis was reversed, i.e., from large to small, to visualize the correspondence between **A** and **B**. In both figures, the cross represents the baseline condition without filtering.

TABLE 2 | Age, vision status, and reading performance of the young normal (YN), older normal (ON), non-macular (Non-Mac) and macular (Mac) groups (mean [standard deviation]).

Groups	Age (years)	VA (logMAR)	CS (log unit)	Maximum reading speed (wpm)	Critical print size (logMAR)	Reading acuity (logMAR)
YN*	20.5 [3.6]	−0.09 [0.11]	2.13 [0.07]	2.20 [0.06]	0.21 [0.10]	−0.14 [0.07]
ON	68.0 [5.0]	0.00 [0.11]	2.02 [0.11]	2.17 [0.05]	0.26 [0.12]	−0.05 [0.10]
Non-Mac	56.7 [16.1]	0.82 [0.40]	0.96 [0.52]	2.06 [0.16]	1.35 [0.82]	0.71 [0.41]
Mac	72.9 [16.7]	0.64 [0.31]	1.16 [0.33]	2.07 [0.21]	1.49 [0.90]	0.71 [0.40]

**For the YN group, the listed VA, CS and reading indices were under baseline condition with no filtering.*

function in the "lme4" package. *Post hoc* analysis was performed with Bonferroni corrections ("emmeans" package, Piepho, 2004). p-values smaller than 0.05 were considered statistically significant.

RESULTS

Validity: Simulating Reduced Acuities and Contrast Sensitivities in Normally Sighted Young Subjects

First, we asked if the CSF filter with certain parameters would actually yield the expected VA and CS scores for normally sighted young subjects. For example, for parameter values a = 0.29 and b = 0.29 (**Table 1**, Filter 1), would the filtered versions of the acuity and contrast sensitivity charts yield the expected test scores of 0.36 logMAR and 1.59 log unit? This validation is important to confirm that two key assumptions underlying our simulation are valid—first, that the use of a horizontally and vertically shifted normal CSF template is a good approximation of a low-vision reader's CSF, and second, that clinical measures of visual acuity and contrast sensitivity can be used to calculate the horizontal and vertical shifts.

Figures 3A,B are scatter plots of the measured VA and CS vs. the expected values from the simulation for the YN group. The dots represent the group average for each low vision simulation, and the solid line represents the equality line.

We used the test-retest reliability of VA and CS (0.20 logMAR and 0.30 log unit, respectively) as criteria for clinically significant differences (Kiser et al., 2005). The difference between the measured and expected VA ranged from −0.11 to 0.35 logMAR, with a median of 0.05 logMAR. LME analysis on VA showed significant main effect of value type [expected and measured; $F(1, 852) = 141.31$, $p < 0.001$], filter conditions [$F(40, 774) = 2098.15$, $p < 0.001$], and an interaction between them [$F(40, 852) = 19.04$, $p < 0.001$]. *Post hoc* analysis with Bonferroni corrections showed that only two conditions showed a significant difference ($p < 0.05$) larger than 0.20 logMAR (**Figure 3A**, red triangles).

The difference between the measured and expected CS ranged from −0.81 to 0.16 log unit, with a median of −0.03 log unit. LME analysis on CS showed significant main effect of value type [expected and measured; $F(1, 829) = 278.05$, $p < 0.001$], filter conditions [$F(40, 485) = 885.76$, $p < 0.001$], and an interaction between them [$F(40, 829) = 38.73$, $p < 0.001$]. *Post hoc* analysis

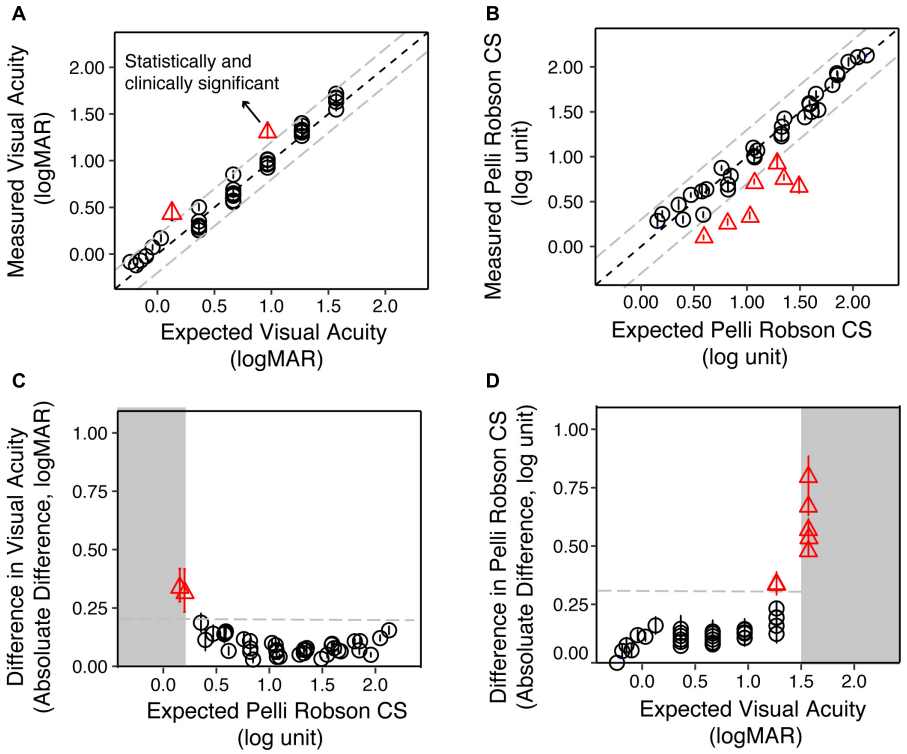


FIGURE 3 | The validity of simulating reduced contrast sensitivity and visual acuity in the Young Normal (YN) group. **(A)** The measured VA as a function of the expected VA. The circles represent the group average of each condition, and the error bars represent the standard errors. The red triangles represent conditions where the differences between the measured and expected values were statistically ($p < 0.05$) and clinically significant (>0.20 logMAR). **(B)** The measured CS as a function of the expected CS. The red triangles represented conditions where the differences between the measured and expected values were statistically ($p < 0.05$) and clinically significant (>0.30 log unit). **(C)** The difference between measured and expected VA as a function of the expected CS. When the expected CS was close to or worse than 0.15 log unit, the difference between measured and expected VA was the largest. **(D)** The difference between measured and expected CS value as a function of the expected VA. When the expected VA values were close to or exceeded 1.50 logMAR, the difference between measured and expected CS were the largest. In all four plots, the gray dashed lines represent the clinically significant difference in low vision (0.20 logMAR for VA and 0.30 log unit for CS), respectively.

with Bonferroni corrections showed that seven conditions had significant differences ($p < 0.05$) that were larger than 0.30 log unit (**Figure 3B**, red triangles).

We asked whether these deviant points are associated with an interaction between poor acuity or contrast sensitivity. We plotted the difference between the expected and measured VA as a function of the expected CS (**Figure 3C**), and found that the two conditions that reached clinical significance both had the lowest expected CS (0.15 and 0.20 log unit). Similarly, the seven conditions that reached clinical significance in CS all had expected VA close to or larger than 1.5 logMAR (**Figure 3D**). LME modeling confirmed this mutual impact between VA and CS. The difference between measured and expected VA increase as CS worsens [$F(1, 43) = 14.07$, $p < 0.001$] and vice versa [$F(1, 39) = 34.40$, $p < 0.001$].

Validity: Simulation of Reading Performance

Do the simulated acuity and contrast sensitivity reductions show similar impacts on reading as real low vision? If this is the case, the real and simulated low-vision subjects with equivalent VA and CS should have similar reading performance.

Figure 1D illustrated how reading speed changed with print size when text images were filtered with three sample filters. To quantify the impact of the simulated VA and CS on reading, we built LME models across all simulation conditions for the YN subjects on the three reading indices, with the expected VA and CS as predictors. Maximum reading speed was only significantly affected by CS [$F(1, 32) = 30.83$, $p < 0.001$], and it was mostly unaffected until the simulated CS was very low (**Figure 4A**). We quantified the impact of CS on the maximum reading speed by an exponential function, which showed that when CS dropped to 0.69 log unit, the maximum reading speed only decreased by 10%. VA and CS were both significant predictors for critical print size and reading acuity (all $p < 0.001$). VA alone explained 88% and 90% of the variations in the critical print size and reading acuity, respectively (**Figures 4B,C**). With the addition of CS, they explained 95% of the variance in the critical print size and 96% in the reading acuity. The parameters of the three regression models are provided in **Table 3**.

We then tested if the models we derived from simulation can reasonably predict the reading performance of subjects

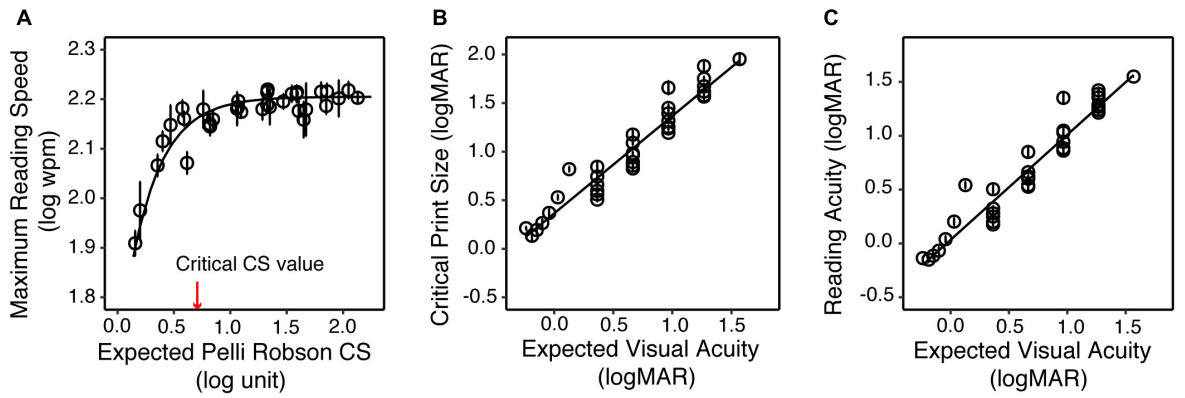


FIGURE 4 | The impact of simulated acuity and contrast sensitivity reduction on the reading performance of the YN group. **(A)** Maximum reading speed as a function of the expected CS. The red arrow represents the critical CS for a maximum reading speed of at least 90% of the maximum reading speed for normal vision (asymptote). **(B)** Critical print size as a function of the expected VA. **(C)** Reading acuity as a function of the expected VA. In all three plots the circles represent the group average of each simulated low vision condition, and the error bars represent standard errors.

TABLE 3 | Regression models on maximum reading speed, critical print size, and reading acuity with VA and CS as predictors.

	VA	CS	R^2	Regression model
Maximum reading speed	$F(1, 35) = 0.42$, $p = 0.52$	$F(1, 32) = 30.83$, $p < 0.001$	0.25	Maximum reading speed=2.20 × (1− exp(−3.56 × (CS+0.40)))
Critical print size	$F(1, 38) = 1333.47$, $p < 0.001$	$F(1, 34) = 88.78$, $p < 0.001$	0.95	Critical print size=0.65 + 0.95 × VA - 0.21 × CS
Reading acuity	$F(1, 35) = 1111.36$, $p < 0.001$	$F(1, 37) = 73.14$, $p < 0.001$	0.96	Reading acuity=0.30 + 0.95 × VA - 0.21 × CS

with actual low vision. Specifically, the reading indices of actual low-vision subjects were obtained from the curves fitted to their reading data (individual data are provided in **Supplementary Appendix 2**), and the predicted reading indices were obtained by entering their VA and CS into the regression models in **Table 3**. **Figure 5** shows scatterplots of the actual versus predicted reading indices for each low-vision subject.

The predicted maximum reading speed (**Figure 5A**) was faster than the actual value in both Non-Mac and Mac groups, by an average of 0.07 log wpm (equivalent to 17%, $p = 0.058$) and 0.12 log wpm (equivalent to 32%, $p < 0.001$), respectively. The predicted critical print size (**Figure 5B**) was in close agreement with the actual values in the Non-Mac group ($p = 0.27$), but it was significantly smaller than the actual values in the Mac group ($p < 0.001$). Large deviations in the critical print size mostly appeared when the actual critical print size values were larger than 2.0 logMAR. These deviations may be due to the fact that under severe low-vision simulation conditions, the tested print size (−0.3 to 1.7 logMAR) was not sufficient to reflect the plateau of the reading curve, therefore the fitted curve may have yielded an unreliable estimation of critical print size. Within the 2.0 logMAR limit, the predicted critical print sizes were not significantly different from the actual values in both groups ($p = 0.13$ for the Non-Mac group, and $p = 0.15$ for the Mac group). The predicted reading acuity (**Figure 5C**) was in close agreement with the actual values in the Mac group ($p = 0.13$), and was slightly larger than the actual value in the Non-Mac group by an average of 0.16 logMAR ($p < 0.001$).

Consideration of Age as an Additional Parameter for Low-Vision Simulation

Many low-vision conditions are age related, as shown by the age and pathology distributions of our low-vision sample. Therefore, we considered whether age should be included as an additional parameter to fine tune the simulation. To answer this question, we included a group of older subjects with normal ocular health (ON group).

Compared to the YN group, the ON group had a significantly larger value of VA by 0.09 logMAR ($p = 0.017$), and were lower in CS by 0.11 log units ($p = 0.002$). When comparing their reading performance with the YN group under the unfiltered condition, the ON group showed lower maximum reading speed, and higher logMAR values of critical print size and reading acuity, but only reading acuity reached significance (by 0.09 logMAR, $p = 0.005$). However, the reduced VA and CS in the ON group were sufficient to explain the age-related change of the reading acuity. When individual subject's VA and CS values were entered into the regression model in **Table 3**, the predicted reading acuity was not significantly different from the actual reading acuity of the ON group ($p = 0.12$, **Figure 5C**). This means that we found no additional age effect on reading, once acuity and contrast sensitivity are taken into account.

DISCUSSION

Digital images of test letters and text were constructed based on the CSF filtering principle, to simulate low vision with various combinations of acuity and contrast sensitivity

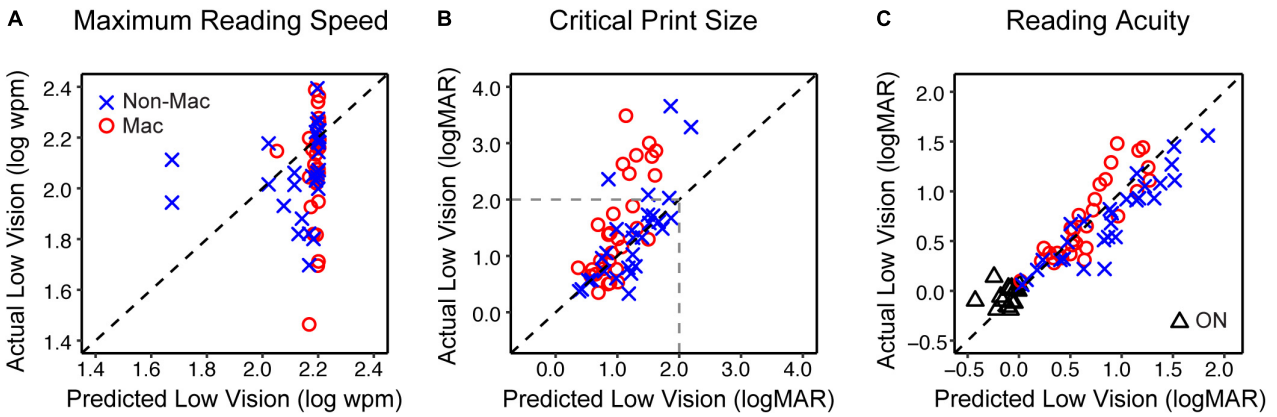


FIGURE 5 | The validity of simulating the reading performance of low vision individuals. **(A)** The comparison of the maximum reading speed between actual low vision and predicted low vision from the simulation. The blue crosses and red circles represent the Non-Mac and Mac groups, respectively. **(B)** Comparison of the critical print size for actual and predicted low vision from the simulation. Note that critical print sizes larger than 2.0 logMAR (dashed lines) may be unreliable due to limitations on the tested print size range. **(C)** Comparison of reading acuity for actual and simulated low vision from the simulation. Comparisons between actual and simulated reading acuity of the Older Normal (ON) group are also shown (black triangles). Note that for the ON group this comparison was only conducted on reading acuity.

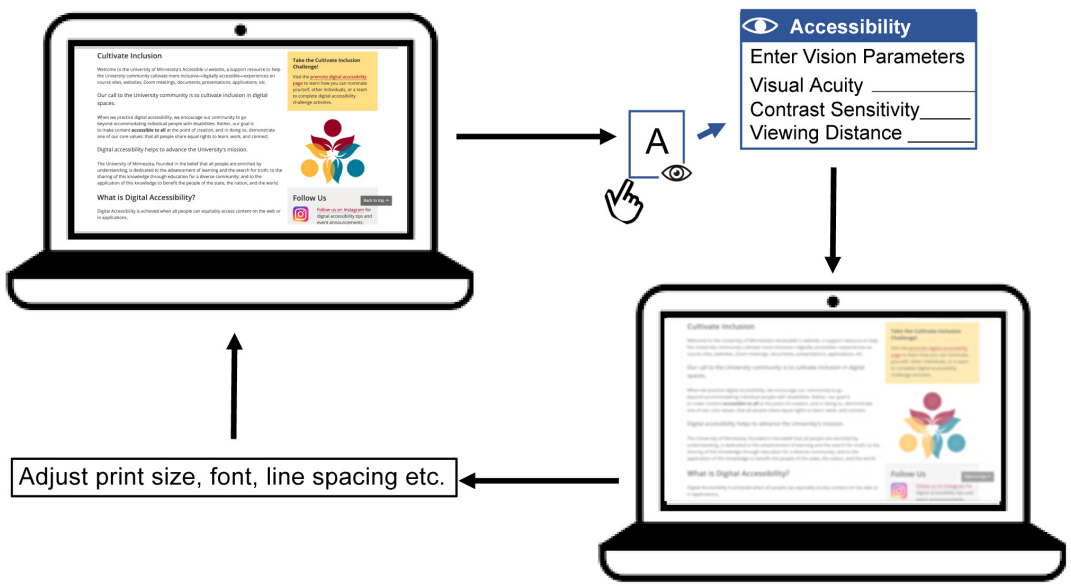


FIGURE 6 | A flowchart illustrating the conceptual design of an accessibility checker for web pages. A web page with text, pictures and links is presented on a laptop screen. The accessibility checker functions as a plug-in for the web browser, which can be initiated by the user. Once initiated, the accessibility checker requests user input including acuity, contrast sensitivity (if available) and preferred viewing distance. The accessibility checker then transforms the web page to demonstrate visibility for the corresponding low-vision condition. This simulation will show an eye-care clinician, family member or web designer how visible the web content is expected to be for the person with low vision. Lastly, the user can adjust the display properties such as overall zoom or print size, font, line spacing etc. to make the page more accessible. The user can repeat this process until the web page appears accessible under the corresponding low vision condition.

reduction. We examined the validity of this simulation by attempting to replicate low-vision performance by testing normally sighted subjects with test-chart letters and text reading. Regarding visibility, we found that our simulation reproduced the desired visual acuity and contrast sensitivity we intended to simulate in normally sighted young subjects. Regarding reading, we found that the simulation overestimated the maximum reading speed but provided a good estimate of critical print size and reading acuity, for real low-vision individuals with corresponding acuity and Pelli-Robson contrast sensitivity.

There has been increasing interest in estimating low-vision CSFs from clinical measures of acuity and contrast sensitivity (Chung and Legge, 2016; Thurman et al., 2016; Thompson et al., 2017). Our simulation of low-vision visibility rests on a simple model in which images

are filtered by shifted versions of the normal CSF. The shifts along the log spatial frequency and log contrast sensitivity axes are related to clinical measures of letter acuity and contrast sensitivity by equations described in **Supplementary Appendix 1**.

The first step in validating the simulation was to verify that the filtered images of letter charts would produce the expected values of reduced acuity and contrast sensitivity when viewed by normally sighted subjects. We simulated forty low-vision conditions based on the empirical distributions of VA and CS across a large sample of subjects with normal vision and vision pathologies (Xiong et al., 2020). We found that when simulating low-vision conditions with acuities better than 1.5 logMAR (approximately 20/630) and contrast sensitivities above 0.15 log unit, the measured acuities and contrast sensitivities closely matched the expected values. For low-vision conditions outside these boundaries, the simulation had poorer performance. It is noteworthy that Pelli-Robson letter size (2 × 2 inches) subtends 2.91 degrees (1.54 logMAR) at a typical viewing distance of 1m used for low vision. This print size is difficult to recognize for individuals whose acuity is 1.5 logMAR or worse, even if they have good contrast sensitivity. This may explain the upper bound of logMAR acuity for our simulation, and indicate that in clinical practice the viewing distance of Pelli-Robson letters should be reduced for patients whose acuity is worse than 1.5 logMAR (Njeru et al., 2021).

We then asked if this simulation generalizes to more interesting real-world stimuli such as digital text or graphics. We compared the reading performance of real low-vision subjects with normally sighted subjects who read under simulated low-vision conditions with equivalent VA and CS. The simulation overestimated the maximum reading speed. The lack of correlation between maximum reading speed and acuity and the weak correlation between maximum reading speed and contrast sensitivity are consistent with previous findings in normal and low-vision reading (Legge et al., 1987; Rubin and Legge, 1989). It is likely that visual factors other than acuity and contrast sensitivity, e.g., visual field loss or unstable reading eye movements (Fletcher et al., 1999; Crossland et al., 2004; Calabrèse et al., 2014), may have detrimental effects on maximum reading speed.

The current simulation only considered acuity and contrast sensitivity reductions associated with low vision, but other visual or non-visual factors might be incorporated to improve the simulation. We examined whether an age adjustment should be included to account for the slight decline of reading performance in older age (Owsley, 2016; Calabrèse et al., 2016). We also compared the validity of the current approach in simulating low vision with or without central vision disturbance.

We found that the simulation based on acuity and contrast sensitivity reductions was sufficient to account for the decrements in reading in our older subjects. However, differences were shown between the Mac and Non-Mac groups, with the overestimations of the simulation for the reading indices being more prominent in the Mac group. These differences are consistent with the adverse impact of central field loss on reading that has been reported in earlier studies. Legge et al. (1985) found that low-vision subjects with central field loss showed slower peak reading speeds than acuity-matched subjects with remaining central vision. Crossland et al. (2004) found that subjects with macular diseases showed impaired fixation stability when reading texts. Therefore, although the visibility-based simulation can adequately replicate the reading performance of the majority of our low-vision subjects, including central field status as a third factor might improve the validity of the simulation.

Digital simulation makes it possible to visualize the information available to a person with low vision. Our study examined the feasibility of utilizing clinically measured acuity and contrast sensitivity to estimate low-vision CSFs across a wide range of low-vision conditions for the purpose of simulating the visibility of image features. What practical value might such a simulation have? Such digital simulation could serve as an "accessibility checker" in the development of architecture and reading related products, and assist people with low vision in choosing optimal reading configurations. Lei et al. (2018) and Thompson et al. (2021) have used similar CSF-based filtering methods to predict the visibility of architectural hazards for people with specified levels of reduced acuity and contrast sensitivity. In the context of reading displays for low vision, it may be possible to construct a web-based accessibility checker for low vision to predict whether a particular combination of print size, font and viewing distance would be legible for someone with specified acuity and contrast sensitivity. Such simulation could also be valuable for educational purposes, to help people with normal vision better understand constraints on visual performance due to low vision. **Figure 6** is a flowchart illustrating how an "accessibility checker" might be used for web pages, potentially as a plug-in for a web browser. When initiated by the user, who might be an eye-care clinician, web designer or family member of a low vision reader, the accessibility checker would take as input a potential user's acuity, contrast sensitivity and desired viewing distance. If contrast sensitivity is not known, an estimate can be made based on the linear relationship between acuity and contrast sensitivity (Xiong et al., 2020). The accessibility checker will then transform the appearance of the web page to simulate visibility of the screen features for the corresponding low-vision condition. Lastly, the user can adjust the overall zoom or properties (e.g., print size, font, line spacing etc.) of the web page to make it more accessible to the potential low-vision reader. Clinicians can also use this accessibility checker to examine whether the icons and texts of operating systems are accessible for a particular low vision patient. These examples refer to the appearance a single page or website might appear in a static view. In our reading test, the processing time of sentence transformation ranged from 0.5 to 0.7 s. This processing time is likely to be acceptable for evaluating the appearance of static text or other static web content. To simulate the appearance of dynamic content, such as web videos, faster processing would be necessary.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Minnesota Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Y-ZX and GEL designed the research. Y-ZX and QL performed the research. Y-ZX analyzed the data. Y-ZX, GEL, and QL wrote and revised the manuscript. AC provided editorial suggestions. All authors approved the final version.

FUNDING

This work was supported by grants from the National Institutes of Health (R01 EY002934 to GEL; 1K99EY030145-01A1 to Y-ZX). Y-ZX was also supported by a fellowship from Envision Research Institute during the preparation of the manuscript.

ACKNOWLEDGMENTS

We thank Jeff Boucher and Chen-Yue Qiao for their assistance in data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.671121/full#supplementary-material

REFERENCES

Barten, P. G. J. (1999). "Chapter 3 Model for the spatial contrast sensitivity of the eye," in *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*, (Bellingham WA: SPIE Press). doi: 10.1117/3.353254

Barten, P. G. J. (2003). "Formula for the contrast sensitivity of the human eye," in *Proceedings of the SPIE 5294, Image Quality and System Performance*, eds Y. Miyake, and D. R. Rasmussen (Bellingham WA: SPIE). doi: 10.1117/12.537476

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bowers, A. R., and Reid, V. M. (1997). Eye movements and reading with simulated visual impairment. *Ophthalmic Physiol. Opt.* 17, 392–402. doi: 10.1111/j.1475-1313.1997.tb00071.x

Calabrèse, A., Bernard, J. B., Faure, G., Hoffart, L., and Castet, E. (2014). Eye movements and reading speed in macular disease: the shrinking perceptual span hypothesis requires and is supported by a mediation analysis. *Invest. Ophthalmol. Vis. Sci.* 55, 3638–3645. doi: 10.1167/iovs.13-13408

Calabrèse, A., Cheong, A. M. Y., Cheung, S. H., He, Y. C., Kwon, M. Y., Mansfield, J. S., et al. (2016). Baseline MNREAD measures for normally sighted subjects from childhood to old age. *Invest. Ophthalmol. Vis. Sci.* 57, 3836–3843. doi: 10.1167/iovs.16-19580

Calabrèse, A., To, L., He, Y., Berkholtz, E., Rafian, P., and Legge, G. E. (2018). Comparing performance on the MNREAD iPad app with the MNREAD acuity chart. *J. Vis.* 18:8. doi: 10.1167/18.1.8

Campbell, F. W., and Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *J. Physiol.* 197, 551–566. doi: 10.1113/jphysiol.1968.sp008574

Cheong, A. M. Y., Legge, G. E., Lawrence, M., Cheung, S. H., and Ruff, M. (2008). Relationship between visual span and reading performance in age-related macular degeneration. *Vis. Res.* 48, 577–588. doi: 10.1016/j.visres.2007.11.022

Cheung, S. H., Kallie, C. S., Legge, G. E., and Cheong, A. M. Y. (2008). Nonlinear mixed-effects modeling of MNREAD data. *Invest. Ophthalmol. Vis. Sci.* 49, 828–835. doi: 10.1167/iovs.07-0555

Chung, S. T. L., and Legge, G. E. (2016). Comparing the shape of contrast sensitivity functions for normal and low vision. *Invest. Ophthalmol. Vis. Sci.* 57, 198–207. doi: 10.1167/iovs.15-18084

Chylack, L. T., Jakubicz, G., Rosner, B., Khu, P., Libman, J., Wolfe, J. K., et al. (1993). Contrast sensitivity and visual acuity in patients with early cataracts. *J. Cataract Refract. Surg.* 19, 399–404. doi: 10.1016/s0886-3350(13)80313-6

Crossland, M. D., Culham, L. E., and Rubin, G. S. (2004). Fixation stability and reading speed in patients with newly developed macular diseases. *Ophthal. Physiol. Opt.* 24, 327–333. doi: 10.1111/j.1475-1313.2004.00213.x

Dickinson, C. M., and Rabbitt, P. M. A. (1991). Simulated visual impairment: effects on text comprehension and reading speed. *Clin. Vis. Sci.* 6, 301–308.

Elfadaly, D., Abdelrazik, S. H., Thomas, P., Dekker, T., Dahlmann-Noor, A., and Jones, P. R. (2020). Can psychophysics be fun? Exploring the feasibility of a gamified contrast sensitivity function measure in amblyopic children aged 4 – 9 years. *Front. Med.* 7:469. doi: 10.3389/fmed.2020.00469

Ferris, F. L., Kassoff, A., Bresnick, G. H., and Bailey, I. L. (1982). New visual acuity charts for clinical research. *Am. J. Ophthalmol.* 94, 91–96. doi: 10.1016/0002-9394(82)90197-0

Fletcher, D. C., Schuchard, R. A., and Watson, G. (1999). Relative locations of macular scotomas near the PRL: effect on low vision reading. *J. Rehabil. Res. Dev.* 36, 356–364.

Jones, P. R., and Ometto, G. (2018). "Degraded reality: using VR/AR to simulate visual impairments," in *Proceedings of 2018 IEEE Workshop on Augmented and Virtual Realities for Good (VAR4Good)*, Reutlingen, 1–4. doi: 10.1109/VAR4GOOD.2018.8576885

Jones, P. R., Somoskeöy, T., Chow-Wing-Bom, H., and Crabb, D. P. (2020). Seeing other perspectives: evaluating the use of virtual and augmented reality to simulate visual impairments (OpenVisSim). *NPJ Digit. Med.* 3:32. doi: 10.1038/s41746-020-0242-6

Kiser, A. K., Mladenovich, D., Eshrachi, F., Bourdeau, D., and Dagnelie, G. (2005). Reliability and consistency of visual acuity and contrast sensitivity measures in adcanced eye disease. *Optom. Vis. Sci.* 82, 946–954. doi: 10.1097/01.opx.0000187863.12609.7b

Legge, G. E., Ross, J. A., Isenberg, L. M., and LaMay, J. M. (1992). Psychophysics of reading: clinical predictors of low-vision reading speed. *Invest. Ophthalmol. Vis. Sci.* 33, 677–687.

Legge, G. E., Rubin, G. S., and Luebker, A. (1987). Psychophysics of reading – V. The role of contrast in normal vision. *Vis. Res.* 27, 1165–1177. doi: 10.1016/0042-6989(87)90028-9

Legge, G. E., Rubin, G. S., Pelli, D. G., and Schleske, M. M. (1985). Psychophysics of reading: low vision. *Vis. Res.* 25, 253–266. doi: 10.1016/0042-6989(85)90118-x

Lei, Q., Carpenter, B., Kersten, D., and Legge, G. E. (2018). Visibility of steps and ramps in natural lighting: effects of simulated loss of acuity and contrast sensitivity. *Invest. Ophthalmol. Vis. Sci.* 59, 3421–3421.

Lei, Q., Kersten, D., Thompson, W., and Legge, G. E. (2016). Simulating reduced acuity in low vision: validation of two models. *Invest. Ophthalmol. Vis. Sci.* 57, 634–634.

Lesmes, L. A., Lu, Z., Baek, J., and Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J. Vis.* 10, 17.1–21. doi: 10.1167/10.3.17

Mansfield, J. S., Atilgan, N., Lewis, A. M., and Legge, G. E. (2019). Extending the MNREAD sentence corpus: computer-generated sentences for measuring visual performance in reading. *Vis. Res.* 158, 11–18. doi: 10.1016/j.visres.2019.01.010

Mansfield, J. S., and Legge, G. E. (2007). "The MNREAD acuity chart," in *Psychophysics of Reading in Normal and Low Vision*, ed. G. E. Legge (Mahwah, NJ: Lawrence Erlbaum), 167–191.

Njeru, S. M., Osman, M., and Brown, A. M. (2021). The effect of test distance on visual contrast sensitivity measured using the Pelli-Robson chart. *Transl. Vis. Sci. Technol.* 10:32. doi: 10.1167/tvst.10.2.32

Owsley, C. (2016). Vision and aging. *Annu. Rev. Vis. Sci.* 2, 255–271. doi: 10.1146/annurev-vision-111815-114550

Peli, E. (1990). Contrast in complex images. *J. Opt. Soc. Am. A* 7, 2032–2040. doi: 10.1364/JOSAA.7.002032

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897x00366

Pelli, D. G., Robson, J. G., and Wilkins, A. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clin. Vis. Sci.* 2, 187–199. doi: 10.1097/00041327-200112000-00014

Piepho, H. P. (2004). An algorithm for a letter-based representation of all pairwise comparisons. *J. Comput. Graph. Stat.* 13, 456–466. doi: 10.1198/1061860043515

Pinheiro, J., and Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. New York, NY: Springer.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ross, J. E., Bron, A. J., and Clarke, D. D. (1984). Contrast sensitivity and visual disability in chronic simple glaucoma. *Br. J. Ophthalmol.* 68, 821–827. doi: 10.1136/bjo.68.11.821

Rubin, R. S., and Legge, G. E. (1989). Psychophysics of reading: VI. The role of contrast in low vision. *Vis. Res.* 29, 79–91. doi: 10.1016/0042-6989(89)90175-2

Sokol, S., Moskowitz, A., Skarf, B., Evans, R., Molitch, M., and Senior, B. (1985). Contrast sensitivity in diabetics with and without back- ground retinopathy. *Arch. Ophthalmol.* 103, 51–54. doi: 10.1001/archopht.1985.01050010055018

Thompson, W. B., Legge, G. E., Kersten, D. J., Shakespeare, R. A., and Lei, Q. (2017). Simulating visibility under reduced acuity and contrast sensitivity. *J. Opt. Soc. Am. A* 24, 583–593. doi: 10.1364/JOSAA.34.000583

Thompson, W. B., Shakespeare, R. A., Liu, S., Kersten, D. J., Creem-Regehr, S., and Legge, G. E. (2021). Evaluating the visibility of architectural features for people with low vision: a quantitative approach. *Leukos*. doi: 10.1080/15502724.2021.1890115

Thurman, S. M., Davey, P. G., McGray, K. L., Paronian, V., and Seitz, A. R. (2016). Predicting individual contrast sensitivity functions from acuity and letter contrast sensitivity measurements. *J. Vis.* 16:15. doi: 10.1167/16.15.15

Turano, K. A., Broman, A. T., Bandeen-Roche, K., Munoz, B., Rubin, G. S., and West, S. (2004). Association of visual field loss and mobility performance in older adults: Salisbury Eye Evaluation Study. *Optom. Vis. Sci.* 81, 298–307. doi: 10.1097/01.opx.0000134903.13651.8e

Xiong, Y. Z., Kwon, M. Y., Kiser, A. K., Virgili, G., Giacomelli, G., and Legge, G. E. (2020). Relationship between acuity and contrast sensitivity - differences due to eye disease. *Invest. Ophthalmol. Vis. Sci.* 61:40. doi: 10.1167/iovs.61.6.40

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AC declared a past co-authorship with several of the authors GEL, AC, and Y-ZX to the handling Editor.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership