Gradiate: A radial sweep approach to measuring detailed contrast sensitivity functions from eye movements

Scott W. J. Mooney

Burke Neurological Institute, White Plains, NY, USA
Blythedale Children's Hospital, Valhalla, NY, USA

✉

Nazia M. Alam

Burke Neurological Institute, White Plains, NY, USA
Blythedale Children's Hospital, Valhalla, NY, USA

✉

N. Jeremy Hill

Burke Neurological Institute, White Plains, NY, USA
Blythedale Children's Hospital, Valhalla, NY, USA

✉

Glen T. Prusky

Burke Neurological Institute, White Plains, NY, USA
Blythedale Children's Hospital, Valhalla, NY, USA
Weill Cornell Medical College, New York, NY, USA

✉

The contrast sensitivity function (CSF) is an informative measure of visual health, but the practical difficulty of measuring it has impeded detailed analyses of its relationship to different visual disorders. Furthermore, most existing tasks cannot be used in populations with cognitive impairment. We analyzed detailed CSFs measured with a nonverbal procedure called "Gradiate," which efficiently infers visibility from eye movements and manipulates stimulus appearance in real time. Sixty observers of varying age (38 with refractive error) were presented with moving stimuli. Stimulus spatial frequency and contrast advanced along 15 radial sweeps through CSF space in response to stimulus-congruent eye movements. A point on the CSF was recorded when tracking ceased. Gradiate CSFs were reliable and in high agreement with independent low-contrast acuity thresholds. Overall CSF variation was largely captured by two orthogonal factors ("radius" and "slope") or two orthogonal shape factors when size was normalized ("aspect ratio" and "curvature"). CSF radius was highly predictive of LogMAR acuity, as were aspect ratio and curvature together, but only radius was predictive of observer age. Our findings suggest that Gradiate holds promise for assessing spatial vision in both verbal and nonverbal populations and indicate that variation between detailed CSFs can reveal useful information about visual health.

Introduction

Impairments in the optical and neural mechanisms that govern spatial vision can have far-reaching negative consequences across many domains of visual and behavioral performance. The dimension of spatial vision most commonly measured in basic assessment—visual acuity—is suitable for diagnosing refractive error and prescribing corrective lenses but has limitations when assessing visual impairments that affect performance at lower contrast levels (Owsley, 2003). Contrast sensitivity is impaired by many optical disorders, including cataracts (Hess & Woo, 1978), glaucoma (Ross, Bron, & Clarke, 1984; Stamper, 1984), optic neuritis (Zimmern, Campbell, & Wilkinson, 1979), diabetic retinopathy (Dosso et al., 1996; Hyvärinen, Laurinen, & Rovamo, 1983), and uncorrected myopia (Collins & Carney, 1990), as well as neurological disorders like multiple sclerosis (Regan, Silver, & Murray, 1977; Regan, Raymond, Ginsburg, & Murray, 1981) and schizophrenia (Cimmer et al., 2006). Contrast sensitivity measurements are also superior to acuity at predicting outcomes for numerous visually driven abilities, such as face recognition (West et al., 2002), driving (Freeman, Muñoz, Turano, & West, 2005), and walking (Geruschat, Turano, & Stahl, 1998).

The contrast sensitivity function (CSF) is a more comprehensive assessment of spatial vision than any single contrast sensitivity or acuity measure but is often impractical to perform in clinical settings and less readily interpretable than one-number thresholds. The CSF curve is generally interpolated from contrast sensitivity thresholds at a small number of spatial frequencies measured with a forced-choice staircase method. The number of trials required for this procedure is often burdensome, and studies have thus

typically measured curves with 6 or fewer points, without repeat measurements, with few observers, and/or with faster but less robust methods. The Pelli-Robson chart (Pelli, Robson, & Wilkins, 1988) uses fixed-size letters of decreasing contrast to measure a threshold around the middle of the CSF (three to five cycles per degree [cpd]), which can provide an estimate of the contrast sensitivity function when combined with the usual acuity chart (Bradley, Hook, & Haeseker, 1991; Elliott, Sanderson, & Conkey, 1990; Wender, 2007). Computerized versions of letter contrast sensitivity tasks have also been developed (Chandrakumar, Colpa, Reginald, Goltz, & Wong, 2013). Others estimate a low-contrast acuity (LCA) threshold (Brown & Lovie-Kitchin, 1989; Pesudovs, Marsack, Donnelly, Thibos, & Applegate, 2004; Regan, 1988), citing the disproportionate number of visual disorders that affect the middle of the CSF. Alternative methods, such as qCSF (Dorr et al., 2017; Lesmes, Lu, Baek, & Albright, 2010), employ Bayesian statistics to compute a parameterized CSF based on an empirical model (e.g., a four-parameter truncated log parabola; Lesmes et al., 2010). This minimizes redundancy across trials and reduces testing duration to a few minutes but adds an additional assumption that CSF curves take approximately the same form in different populations. (See Pelli and Bex [2013] for a review of CSF methodologies.)

A common but limiting feature of current CSF assessment procedures is that stimulus visibility is inferred from verbal or button-press responses across a sequence of discrete trials. Since observers must comprehend instructions, consistently provide volitional feedback, and maintain focus throughout the task, individuals with cognitive deficits cannot be reliably measured. Cortical visual impairment (CVI), for example, is often accompanied by cognitive, communicative, and/or attentional impairments that preclude even basic acuity chart assessments (Good, Jan, DeSa, Barkovich, & Groenveld, 1994; Huo, Burden, Hoyt, & Good, 1999), which makes the diagnosis and quantification of CVI challenging. CVI also predominantly occurs in children, who find it difficult to perform psychophysical tasks even when cognitively healthy (Witton, Talcott, & Henning, 2017).

Numerous nonverbal alternatives to classical tasks exist for clinical use in assessing vision but often at the cost of accuracy or efficiency. Visual evoked potentials, for example, are automatic and objective (Leat, Yadav, & Irving, 2009) but are also less sensitive than behavioral tasks (de Faria, Katsumi, Arai, & Hirose, 1998) and involve lengthy setup and testing sessions. Gaze-based methods have also been developed that replace direct behavioral reports with the observer's eye movements. These tasks operate using certain assumptions about how stimulus visibility drives attention and oculomotor behavior, and they vary widely in complexity and efficacy. Well-known examples in spatial vision assessment include preferential looking paradigms such as Teller cards (Teller, McDonald, Preston, Sebris, & Dobson, 1986), which exploit the human tendency to spend more time fixating on more visible patterns, and optokinetic drum methods (Thomas, Seiler, Sadda, Coffey, & Aramant, 2004), in which observers are placed inside a rotating cylinder and pattern visibility is inferred from the presence of the optokinetic nystagmus reflex (OKN).

The most promising gaze-based procedures for clinical application are portable tracking tasks that can be used to present a large number of distinct stimuli to immobilized patients. These procedures may use eye trackers to automate preferential looking tasks (Jones, Kalwarowsky, Atkinson, Braddick, & Nardini, 2014) but can also capture smooth pursuit/OKN eye movements that occur in sync with moving targets, such as scrolling full-screen noise (Dakin & Turnbull, 2016). Pursuits that match the velocity of the target are interpreted as strong evidence of target visibility, based on previous work demonstrating the close (although imperfect) correspondence between ocular pursuit and perception (Gegenfurtner, 2016; Schütz, Braun, & Gegenfurtner, 2011; Spering & Carrasco, 2015; Spering & Montagnini, 2011). Smooth pursuits are extremely difficult for humans to generate in the absence of a moving target, and false-positive rates in these tasks are consequently much lower than tasks based solely on fixations and/or saccades (such as preferential looking). Tracking evidence can be gathered in discrete bursts and used to drive a sequence of staircasing trials with individual stimuli, a la traditional psychophysics, but can also be interpreted as a continuous data stream. Bonnen, Burge, Yates, Pillow, and Cormack (2015), for example, used a cursor-based tracking task to estimate visual sensitivity to a noisy luminance signal on a frame-by-frame basis. They aptly called this type of task "continuous psychophysics," in which each task frame is an informative "mini-trial."

We have previously described a contrast sensitivity assessment procedure called "Curveball" that combines the approaches described above—eye tracking, smooth pursuits, wandering targets, and frame-by-frame analysis—with real-time, continuous stimulus manipulation (Mooney et al., 2018). In Curveball, observers tracked a patch of filtered noise with fixed spatial frequency around the display. The noise target continuously decreased in contrast on each frame it was smoothly pursued until the observer's contrast threshold was reached, at which point tracking necessarily ceased. We demonstrated that Curveball was reliable and valid (as compared to a conventional forced-choice staircase task) but also reported several shortcomings. In particular, a high false-negative rate led to the exclusion of numerous healthy participants (almost 20%) who did not track the target with enough

smoothness or consistency to reach their threshold before the trial timed out. Furthermore, when tested in children with CVI, the task was often not engaging enough to hold their attention, took too long to complete (even at 5 min), and had issues distinguishing false negatives from nonstarter trials that fell completely outside the patient's CSF.

Here, we describe a new CSF measurement procedure called "Gradiate." Gradiate shares multiple features with Curveball: Both procedures infer stimulus visibility from eye movements and manipulate stimulus appearance in real time, which makes the tasks intuitive, rapid, nonverbal, and resistant to false positives. Unlike Curveball, however, Gradiate does not interpolate a CSF from contrast thresholds measured at fixed spatial frequencies. Instead, it is explicitly designed to generate detailed CSFs. When tracked, stimuli change in spatial frequency *and* contrast at discrete intervals along sweep trajectories that approach the CSF curve from multiple angles. All Gradiate sweeps begin at a common origin and radiate outward along vectors in log-log CSF space until they reach the limit of function, at which point stimulus invisibility prevents any further tracking by the observer and a threshold is recorded (Figure 1). This polar approach to CSF estimation has advantages over the conventional Cartesian approach that varies contrast only. First, each sweep is initiated at a spatial frequency that is highly visible, which ensures that most observers are able to make progress on all sweeps regardless of their spatial vision loss. Second, each sweep approaches the CSF curve from an approximately perpendicular direction, which minimizes the length of the curve interval that the sweep could intersect, given some unavoidable uncertainty in the precise effective spatial frequency and contrast of each stimulus. (This is not always the case for vertical sweeps, as a small amount of uncertainty in stimulus spatial frequency produces a disproportionately large contrast sensitivity error at the steeper parts of the CSF.) At the same time, Gradiate makes fewer assumptions about the shape of the CSF than parameterized approaches like qCSF (Lesmes et al., 2010), which could significantly benefit efforts to characterize the CSFs of diverse populations with poorly understood disorders such as CVI.

We addressed Curveball's problem with excessive false negatives by implementing a new algorithm for inferring stimulus visibility. The task now exploits fixations and saccades in addition to smooth pursuits, which allows observers who exhibit a high saccade-to-pursuit ratio ("saccadic trackers") to make progress rather than timing out. Gradiate also differs from previous tasks in other aspects: (a) Stimulus contrast does not fade continuously in response to observer tracking (as in Curveball) but rather by discrete steps, which permits changes in spatial frequency and indicates progress to the observer more frequently; (b) multiple (five) moving targets are presented in each trial to minimize
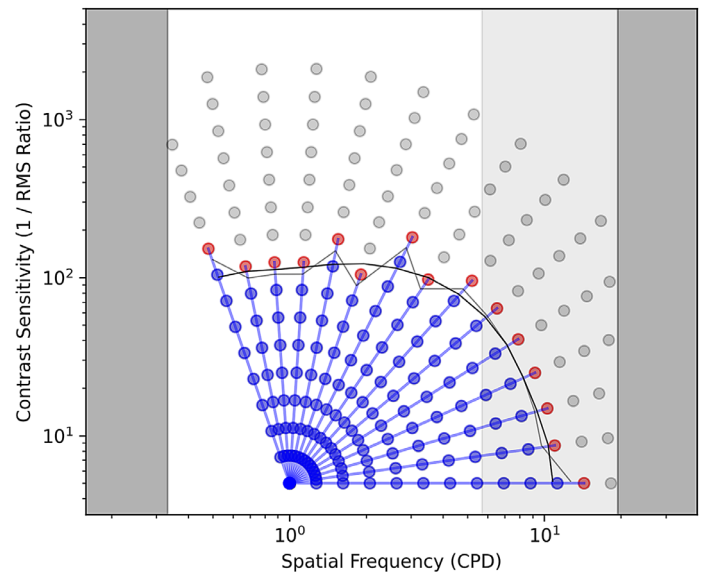

Figure 1. Example of a detailed CSF measured with Gradiate, with spatial frequency on the horizontal axis and contrast sensitivity on the vertical axis (log-log scale). Each dot represents one stimulus in a single radial sweep from a common origin, which observers make progress along through successful target tracking. The blue, red, and gray dots represent stimuli that this observer successfully tracked, failed to track, and never reached, respectively. Thresholds halfway between the final success and first failure are joined to form the CSF (gray line), with the black curve depicting a smoothed version. The dark gray regions to either side of the plot indicate spatial frequencies that could not be displayed due to stimulus size (at the low end) and screen resolution (at the high end), which truncated the leftmost and five rightmost sweeps. The light gray region indicates spatial frequencies that, when multiplied by the stimulus speed, put individual pixel fluctuations above the screen's temporal Nyquist limit; to render these stimuli, additional anisotropic spatial filtering was required to remove Fourier components close to the orientation of stimulus motion. This CSF was generated in about 2 min.

the downtime that occurs whenever the observer must wait for new targets to appear; and (c) reinforcement is given to the observer on a moment-by-moment and trial-by-trial basis through auditory feedback and score presentation, which facilitates task engagement and reduces distractions during testing.

We tested whether Gradiate is reliable in neurologically intact observers despite its numerous differences to Curveball—particularly given its novel "radial threshold" approach to CSF estimation—and examined how Gradiate CSFs relate to a simple yet common cause of visual dysfunction: impaired acuity due to refractive error. Many individuals with cerebral impairment may have undiagnosed refractive error (e.g., in cases of impairment from birth) or outdated preinjury results. Other individuals may not tolerate

corrective eyewear even if correctly diagnosed. It is therefore critical that we determine how refractive error impacts the shape and size of the CSF before attempting to utilize Gradiate (or CSF assessment in general) in the diagnosis of more complex optical or cerebral visual disorders, particularly in clinical populations. We measured 15-point CSFs in 60 healthy observers of varying age, 38 of whom had corrective eyewear. Measurements were collected binocularly with no head restraints under standard room illumination, which we consider to be the minimum requirements for a task to be practicable in studies of hospitalized populations with brain injury and/or severe cognitive impairment. We then used factor analyses to generate an empirical model of the Gradiate CSF and examined how its normative form is impacted by refractive error (as measured with a LogMAR eye chart) and age. To confirm that the polar Gradiate CSFs capture conventional Cartesian measures, we collected LCA thresholds (a commonly used clinical get-around when full CSF measurement is impractical) and compared them to scores interpolated from the radial sweeps. Finally, six observers (five with refractive error) also had their CSFs assessed with a four-alternative forced-choice (4AFC) task to examine the relationship between the contrast sensitivity curves generated by conventional and tracking-based methodologies.

Methods

Observers

Sixty healthy observers (34 female) participated. Observer age ranged from 11 to 74 years, with a mean of 34.95 and standard deviation of 13.85. The four authors were among the observers; all other observers were naive to the aims of the experiment and were recruited as volunteers from employees at the Burke Neurological Institute and their families and friends. Thirty-eight observers had corrective eyewear and were tested both with and without it. All observers provided informed consent under an approved institutional review board protocol and were not financially compensated. Experimental data were secured and managed with the REDCap database (Harris et al., 2009).

Apparatus

A 27-in. widescreen LCD Dell (Round Rock, TX, USA) Optiplex all-in-one computer was mounted on a wheeled stand with an articulated arm and outfitted with a Tobii 4C eye tracker (operating distance of 50–95 cm; sampling rate of 90 Hz) with a professional-level license (Tobii Technology, Stockholm, Sweden). Eye-tracker data were accessed with the Tobii Stream Engine library, which computes the gaze point on the display using one or both eyes (as detected in real time) and applies a small amount of smoothing to the data stream. The exact parameters of this smoothing are not accessible or modifiable with the Stream Engine library but resemble a simple sliding average over approximately 10–15 frames. Real-time control of stimulus behavior was programmed in Python using the Shady graphics toolbox (Hill, Mooney, Ryklin, & Prusky, 2019), which was also used to calibrate screen gamma. Minimum and maximum values of screen luminance were measured under the experimental illumination conditions with an ILT1700 radiometer (International Light Technologies, Peabody, MA, USA) as 10.0 cd/m^2 and 221.1 cd/m^2, respectively. Head movements were not restrained, and observers were given no specific instruction to keep their head still during the task. The Tobii 4C tolerates head movements well, and no observer experienced difficulty with the task due to their idiosyncratic preference for head versus eye tracking. The software was configured to blank out the screen and display a warning message whenever the eye tracker detected that the observer's eyes were closer than 520 mm or further than 720 mm from the screen.

Gradiate task

Observers were presented with a set of filtered noise patches—each defined by a combination of spatial frequency and root-mean-square (RMS) contrast—that moved around the screen on a uniform gray background of normalized luminance 0.5 (Figure 2A). An algorithm continuously inferred the visibility of each target in real time by updating an evidence counter based on the relationship between the observer's gaze and the changing position of that target. Positive evidence required the observer to exhibit tracking that fell within a position-based tolerance radius and a trajectory-based tolerance radius simultaneously, which was determined continuously at 60 Hz (i.e., every video frame). Position-based tracking error was calculated by summing the distances between target position and gaze position over the last eight frames. The error threshold was set to the radius of the stimulus (3°) plus an additional 2° and is depicted by the red circle around the central stimulus in Figure 2A. Trajectory-based tracking error was calculated by summing gaze-target distances over the same eight frames *after* subtracting the most recent gaze-target deviation vector from each of those eight gaze samples. This component of the algorithm had a much stricter error threshold of 0.4°; it was therefore only satisfied when an observer's gaze closely matched the trajectory of the target but was not
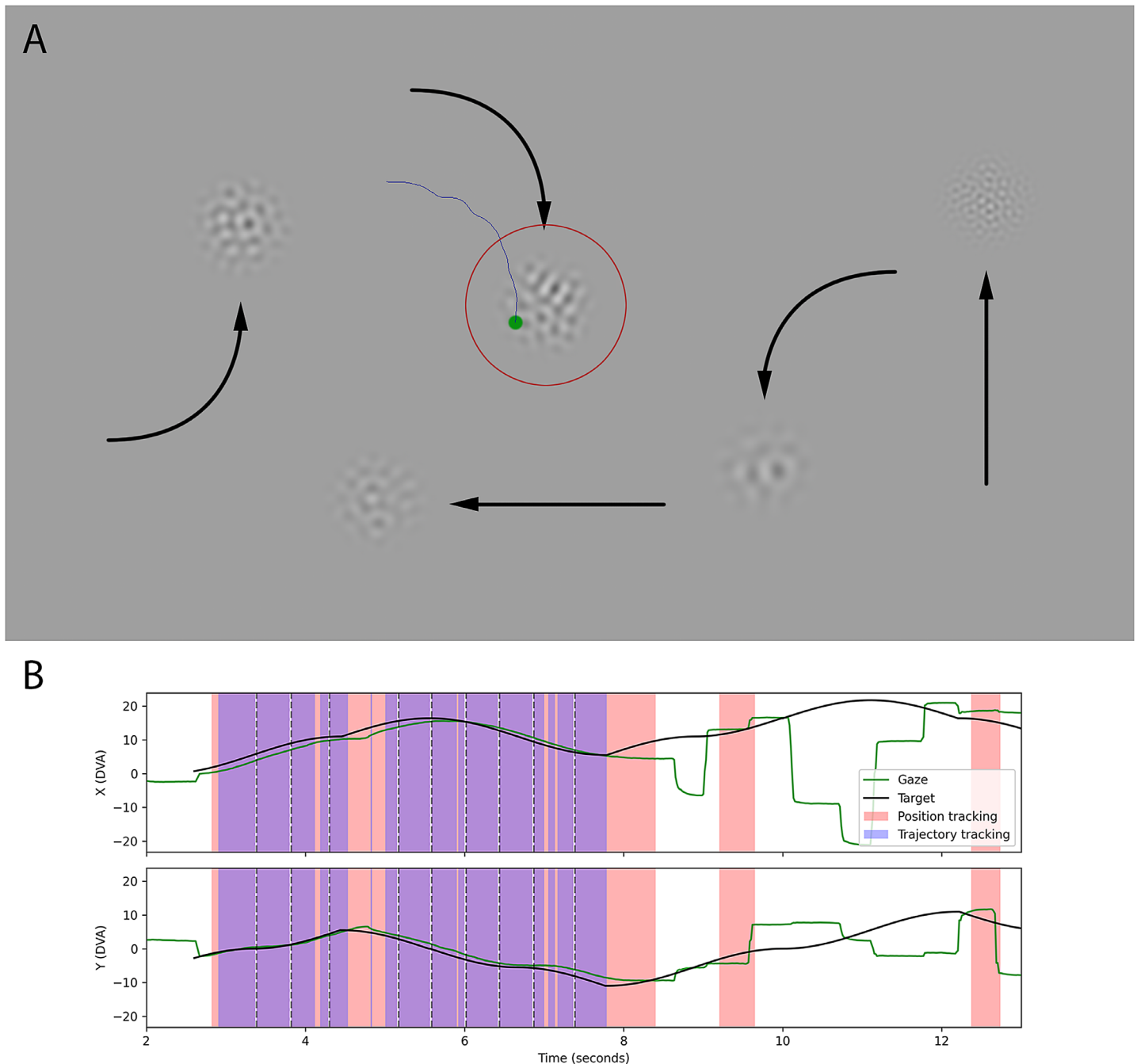
Figure 2. (A) Screenshot from one trial of the Gradiate task, which measures thresholds along five radial sweeps in one presentation (see Figure 1). Background and mean stimulus luminance are 0.5 normalized display luminance. Stimuli move around the screen randomly, as shown by the example arrows, making smooth turns as often as possible while avoiding collisions and synchronous movements. The Gradiate algorithm collects evidence about the visibility of each stimulus from the observer's eye movements and advances that stimulus by changing its spatial frequency and/or contrast as soon as visibility is inferred. The trial ends, and final thresholds are computed, after a period of no successful tracking of any stimulus. In this frame, the observer is currently tracking the central stimulus, as their gaze (green dot) both falls within its positional tolerance radius (red circle) and closely matches its recent trajectory (blue line vs. black arrow). See also Supplementary Movie S1 for a video of one complete trial. (B) Gaze (green) and target (black) X and Y traces in degrees of visual angle for one stimulus in another example trial. Timeline segments highlighted in red and blue indicate periods when position-based tracking and trajectory-based tracking were detected, respectively. The tracking types appear purple when overlapping. The vertical black dotted lines indicate when the stimulus was tracked enough to progress along its sweep. The dotted lines always have white backgrounds because the gaze sample buffers used to detect tracking are cleared whenever the stimulus changes appearance. In this example, the observer responds to a sharp change in target movement direction at 4.5 s and loses the ability to track the target at 7.5 s after the last dotted line. Brief periods of coincidental position-only tracking at 9.5 and 12.5 s are ignored while the observer makes saccades across the now perceptually empty display.

sensitive to any systematic deviation in position. This combination of a generous position-based tracking detector and strict trajectory-based tracking detector was designed to reduce reliance on perfect eye-tracker calibration, which cannot always be obtained in clinical populations. Eye trackers that are miscalibrated by up to ∼5° in any direction can be used without precluding either component of the tracking detection algorithm, provided that the eye tracker is calibrated to the correct display size. Blinks and other interruptions to the eye-tracker data stream were not classified or handled explicitly; tracking analysis required an additional eight frames of buffered samples to resume after any such data stream interruption occurred. Example gaze versus target position traces in the horizontal (X) and vertical (Y) screen dimensions are depicted in Figure 2B, with periods of position-based and trajectory-based tracking highlighted in red and blue, respectively.

Gaze samples that satisfied both conditions were classified as smooth tracking and progressed the evidence counter of that stimulus by 5 units per frame. Samples that satisfied only the position-based tracking condition caused the evidence counter to remain unchanged, which allowed observers who tend to mix short bursts of smooth pursuit with frequent catch-up saccades ("saccadic trackers") to make progress. Samples that satisfied neither condition caused the evidence counter to decay at a rate of 1 unit per frame (to a minimum of zero evidence). When enough evidence of a target's visibility was collected (100 units), that target instantaneously updated its appearance by altering its spatial frequency and/or contrast by one step along a predetermined sweep through CSF space. The evidence counter for that target was then reset to zero and the process repeated. At the same time, a separate global evidence counter determined when no stimuli were visible to the observer and terminated the trial when it fell to a sufficiently negative value (−300 units). This global evidence counter was similarly increased when any target was tracked and decayed at a constant rate when neither tracking condition was met for any stimulus but was further penalized by saccades away from the display or into empty screen space by an amount proportional to the amplitude of the saccade. Saccade events were initially detected using a velocity threshold of 25°; false positives caused by eye-tracker noise were then filtered out by ignoring saccades that (a) were shorter than 50 ms, (b) included an instantaneous angle change of 45° or more, or (c) exceeded 900°/s in instantaneous speed. A threshold was recorded for each target halfway along the interval (in log-log space) between the point representing the last spatial frequency/contrast combination that was successfully tracked and the current (untracked) point. RMS contrast sensitivity (hereafter referred to as CS) was computed as the inverse of the RMS contrast ratio.

The trajectories of the Gradiate sweeps were defined as radial vectors with a common origin, equal magnitude, and directions separated by equal angles in a linear transformation of the log-log CSF space. This transformation was chosen through pilot testing such that a normative CSF was approximately mapped onto a circle of radius 0.5. The log-log values of the point (0.25 cpd, $0.5*10^1$ CS) were mapped to (0, 0) and the log-log values of the point (12 cpd, $10^{3.5}$ CS) to (1, 1). This box also encompassed the range of CSF curves we observed in our previous study (Mooney et al., 2018). Fifteen sweep vectors were defined in this space with unit magnitude and an origin corresponding to a spatial frequency of 1 cpd and $0.5*10^1$ CS (i.e., a starting RMS contrast ratio of 0.2). The 15 sweeps were spread evenly between polar angles 109.703° and 0° (by an angular interval of 7.836°). These sweeps are depicted in log-log CSF space in Figure 1.

Each of the 15 sweeps were divided into 16 steps, with each step determining the spatial frequency and contrast of a single narrow-band noise stimulus. Step size was determined such that a single sweep threshold took approximately 10 s to complete in pilot testing. These stimuli were generated by applying a circular-symmetric Hann window to a filtered noise pattern ($1/f$ amplitude spectrum and random phase) subtending 6° of visual arc at the viewing distance of 620 mm. Each noise pattern was filtered with a band-pass filter centered on the target spatial frequency with a constant width of 0.34 octaves. Temporal aliasing at high spatial frequencies was prevented with an additional anisotropic filter, which removed all Fourier components with horizontal spatial frequency greater than 5.7 cpd (95% of the 6-cpd Nyquist limit of a stimulus moving at 5° per second on a 60-Hz display). The orientation of each stimulus was yoked to its movement direction to ensure that the antialiased direction of the texture always matched the direction of motion. Stimuli with a target frequency above 19.416 cpd were not permitted to appear in the task: After antialiasing, these stimuli contain no power at any orientations that deviate up to 72° from the direction of motion, making them too difficult to track in our pilot testing. Similarly, spatial frequencies below 0.4 cpd could not be displayed with at least two full cycles in a texture subtending 6° of visual arc and were not permitted to appear. Excluding these stimuli truncated several sweeps that would otherwise broach very low or very high spatial frequency values, as shown in Figure 1. In order to keep tracking error thresholds consistent, stimulus envelope size did not vary with spatial frequency, and higher spatial frequency stimuli consequently exhibited more cycles. Prior work with sine gratings has found that holding envelope size constant rather than number of cycles can have a small influence on the CSF (Savoy & McCann, 1975), but as this occurs predominantly at low frequencies, we

determined that the trade-off with consistent error thresholds was worthwhile.

Five targets were presented to the observer simultaneously in each Gradiate trial, with their movements coordinated using a grid motion system. Each target moved from node to node in an invisible diamond grid along straight lines, making smoothed 90° turns at pseudo-random intervals. Movement rules prevented the targets from colliding, making synchronous movements, or repeating the same movement multiple times. If no movement in any direction was possible for a given target (e.g., because it was surrounded), it remained motionless until movement was possible, and any tracking of that target was ignored until movement had resumed. However, the most recently tracked stimulus always received priority for finding a valid movement.

LCA was measured with a variant of the Gradiate task that employed a single horizontal sweep with constant RMS contrast 0.06. The sweep comprised 16 steps of spatial frequency between 0.5 and 16.0 cpd, spaced evenly in the same linear transform of the log-log space as the radial CSF sweeps. For this task, the five targets presented were all separate instances of the same sweep, allowing five repeated measurements of the LCA threshold to be collected within one trial.

Conventional staircase task

Six observers (five with refractive error) also had their CSFs assessed using a conventional four-alternative forced-choice staircase task. Eight of the 15 radial sweeps used in the Gradiate task were measured, with every second Gradiate sweep left out. Stimuli were generated and progressed in the same way as Gradiate. In each trial, a filtered noise patch appeared just upward, rightward, downward, or leftward of a central black fixation point on the same mid-gray background as Gradiate. Each noise patch was again 6° in diameter and, to better emulate Gradiate's effects of motion on sensitivity, scrolled within its fixed envelope at 5° per second. Scrolling direction was randomly set to leftward or rightward in each trial. Observers indicated their best guess of the target's position by pressing the corresponding arrow on the keyboard. If their response was correct, a green dot flashed in the center of the screen, and that sweep was progressed by 0.05 normalized units outward along that radial trajectory in log-log CSF space, which determined its spatial frequency and contrast on its next appearance (i.e., became harder to see). If their response was incorrect, a red dot flashed, and that sweep was regressed by 0.05 units inward (i.e., became easier to see). All eight sweeps were randomly interleaved and, as in Gradiate, began at a common origin point with a spatial frequency of 1 cpd and RMS contrast ratio of 0.2. Each sweep

continued until five reversals had occurred, and the final threshold of that sweep was computed by taking the four midpoints (in log-log space) between each pair of consecutive reversals, discarding the lowest and highest estimates, and averaging the remaining two. The eye tracker was used to pause the task before each trial until the observer returned to the central fixation point (with a tolerance radius of 2°). Observers were given a break after every 100 trials.

Procedure

A single experimental session comprised a calibration step using the standard Tobii Eye Tracking software suite, two repeats of a complete 15-point CSF measurement with Gradiate, and one trial (containing five repeated measurements) of the LCA variant of Gradiate. Each repeat of the complete CSF measurement was conducted by randomly dividing the 15 sweeps into three subsets of 5 sweeps, with those 5 sweeps being presented simultaneously. A random selection of background music was played during the task. Whenever a stimulus was advanced by the observer's eye movements, a musical note was played as helpful feedback for the observer. All audio was controlled using the Audiomath sound toolbox for Python (Hill, Mooney, & Prusky, 2020). When all three sweep sets were completed, the screen showed how many notes had been played. The LCA trial had the same feedback features and was conducted between the two repeats of the full CSF. Finally, LogMAR acuity was measured using a Tumbling "E" eye chart at a viewing distance of 620 mm. In total, a full experimental session took observers approximately 5 min. Observers completed two such sessions on different days or on the same day after a break and were instructed to follow each ball with their eyes until they could no longer see it. All observers with refractive correction completed two additional sessions without their eyewear on the same or different day. Six observers returned at a later date to complete the conventional 4AFC staircase task as described above. Five of those observers required refractive correction and completed sessions with and without their glasses on the same day.

Results

Observers were able to push the stimulus parameters up to their allowable limits on only 10 sweeps out of a total of 5,880 (0.17%) in the main CSF measurement task. These occurred on one of the three rightmost sweep vectors and no more than once per observer. For consistency, thresholds for these sweeps were computed as one half-step beyond the final point. Only 22 sweeps
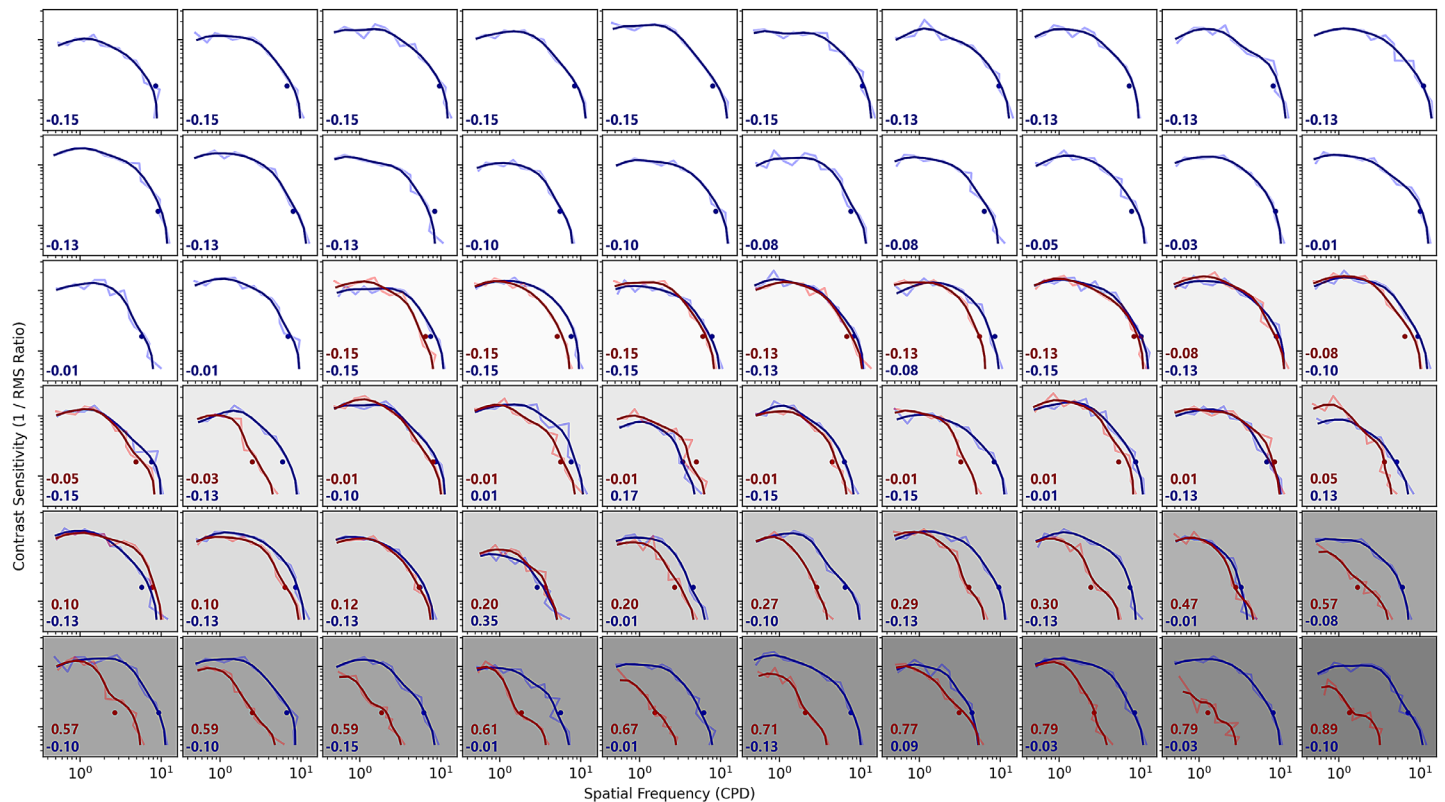
Figure 3. CSFs for all observers. Normal or corrected-to-normal CSFs are depicted in blue, and impaired uncorrected CSFs are in red. The darker lines show the smoothed functions, and the numbers in each plot indicate normal (blue) and impaired (red) LogMAR acuity. The background of each plot becomes darker as impaired acuity gets worse. The singular blue (normal) and red (impaired) dots in each panel of Figure 3 represent the LCA thresholds estimated with horizontal sweeps at an RMS contrast ratio of 0.06, averaged over all 10 repeats of that sweep for each correction type. Observers are sorted from top-left to bottom-right first by their need for corrective lenses and second by either their normal or (where applicable) impaired eye-chart acuity.

had no successes at all, most likely from momentary distraction, and were excluded from the computation of within-session means. None of the 490 LCA sweeps reached the highest possible spatial frequency or had no successes. There was a small but significant correlation between sweep success count and the order in which sweeps were initially tracked (from first to fifth) within each of the 1,176 total trials, $r = 0.107$, $p < 0.001$. Upon closer examination, this effect was driven entirely by a small deficit in mean success count for the first sweep tracked (7.67 stimuli as opposed to 8.47, 8.41, 8.37, and 8.36 stimuli for the second through fifth sweeps). The correlation is not significant when the first sweep is excluded ($r = -0.02$, $p = 0.063$). It is not clear why the first sweep to be tracked exhibited this slight deficit, but as there was no correlation between sweep tracking order and sweep angle ($r = 0.001$, $p = 0.94$), it is unlikely to have biased the thresholds. The average time taken to complete a 15-point CSF was 2 min and 14 s with a standard deviation of 22.6 s.

Figure 3 depicts the CSFs for each observer averaged over all four repeats of Gradiate's 15-point measurement. Blue lines represent CSFs from observers with normal or corrected-to-normal vision, and red lines represent uncorrected CSFs from observers who wear corrective lenses. For simplicity, these two conditions will be referred to hereafter as "normal vision" (blue) and "impaired vision" (red), although not all observers who use refractive correction exhibit impaired CSFs at a 620-mm viewing distance. The lighter lines connect the raw thresholds and the darker lines show the same data smoothed with a Hann window. The red and blue numbers in the bottom left of each plot indicate impaired and normal LogMAR eye chart acuity, respectively.

To allow straightforward comparisons across observers and conditions, all CSF analysis was conducted in the space of radial sweep lengths, which are values between 0 and 1 computed from the progress made along each sweep (as $\frac{n-1}{15}$, where n is the number of successfully tracked stimuli in that sweep). These "polar coordinates" of the CSF were reliable across the two separate experimental sessions conducted for each observer, with a Bland-Altman coefficient of repeatability (CoR) of 0.0562. Figure 4 depicts a Bland-Altman comparison of the measured LCA
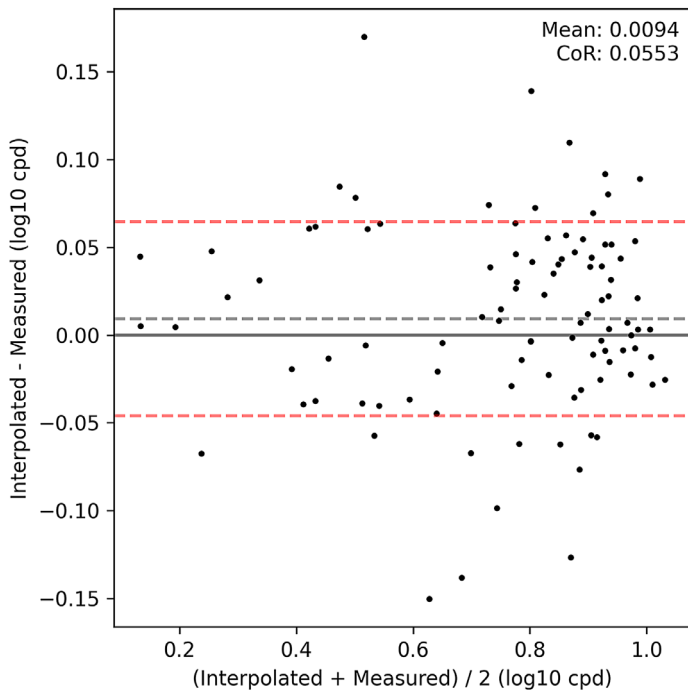
Figure 4. Bland-Altman plot comparing all samples of measured LCA and interpolated LCA. The mean of these measures is shown on the horizontal axis and their difference (interpolated minus measured) on the vertical axis, both in $\log_{10}$ cpd units of spatial frequency. The mean difference and CoR are shown in $\log_{10}$ cpd in the top-right corner. The mean difference and 95% confidence interval are shown by the gray and red dotted lines, respectively, and the solid black line represents a difference of zero.

thresholds (horizontal sweeps) with estimates computed by linearly interpolating the unsmoothed Gradiate CSF measured for that observer and correction (i.e., the point where each horizontal LCA sweep would intersect each smoothed curve in Figure 3). The CoR was 0.055 $\log_{10}$ cpd, which is almost identical to the CoR of the CSF sweeps and less than 5% of the $\log_{10}$ interval between successive spatial frequency values in the LCA task (1.281). This indicates that the two measures were in close agreement, as can be seen directly in Figure 3.

Gradiate factor analyses

We performed two factor analyses to identify variables that account for the variance in all CSFs: one analysis of the raw (not smoothed) sweep lengths and one analysis after normalizing the mean sweep length of each CSF to the grand mean (0.528). The second analysis was conducted to identify factors that account for variation in CSF *shape* rather than size. All mean sweep lengths from the normal and impaired CSFs of all observers were included in both analyses

for a total of 98 CSFs. Factors were identified using the principal component method and an orthogonal equamax rotation. Two factors with an eigenvalue greater than 1 were retained in each analysis. The factor loadings on all 15 sweeps are shown in the top row of Figure 5 and are immediately interpretable. In the raw analysis (left), Factor 1 explained variance in overall CSF radius (79.35%), weighted more heavily toward the high-spatial-frequency end of the curve, while Factor 2 explained some remaining variance (9.49%) in CSF radius at the low-spatial-frequency end. Notably, Factor 1 was highly correlated with both peak contrast sensitivity ($r = .833$) and peak spatial frequency ($r = .978$). Factor 2, however, was moderately correlated with peak contrast sensitivity ($r = .519$) but only weakly and negatively correlated with peak spatial frequency ($r = -.127$, all $p < 0.001$). Both factors explained 96.4% of the variance in peak contrast sensitivity and 97.3% of the variance in peak spatial frequency when combined in linear regression models. We labeled these factors as "CSF radius" and "CSF slope," respectively. Note the results of this factor analysis do not substantially change if the impaired curves are excluded from the analysis, indicating that the same underlying factors account for variation in CSFs among observers with normal or corrected-to-normal refraction. Similarly, performing the same factor analysis on the smoothed CSFs yielded nearly identical results, except that the two factors instead account for a total of 97.56% of the variance in the smoothed sweep lengths. In the normalized analysis (right), Factor 1 explained 55.41% of the total variance and can be interpreted as the "aspect ratio" of the CSF. Higher scores on this factor indicate a greater width-to-height ratio in the CSF; conversely, lower scores indicate a smaller width-to-height ratio, as is the case for CSFs that drop off rapidly from their peak sensitivity. Predictably, this factor was significantly correlated with both CSF radius ($r = 0.828$) and slope ($r = -0.532$) from the raw analysis, both $p < 0.001$. Factor 2 accounted for 10.84% of the variance and can be interpreted as the "curvature" of the CSF. Higher scores on this factor indicate that the CSF has a greater radius along the central (more diagonal) sweeps relative to the low-end and high-end sweeps, while lower scores are typical of CSFs that peak earlier and are more buckled in the center. Curvature was correlated positively and moderately with both CSF radius ($r = 0.321$) and slope ($r = 0.291$), both $p < 0.001$.

In the normalized analysis (right column), Factor 1 explained 55.41% of the total variance and can be interpreted as the "aspect ratio" of the CSF. Higher scores on this factor indicate a greater width-to-height ratio in the CSF; conversely, lower scores indicate a smaller width-to-height ratio, as is the case for CSFs that drop off rapidly from their peak sensitivity. Predictably, this factor was significantly correlated
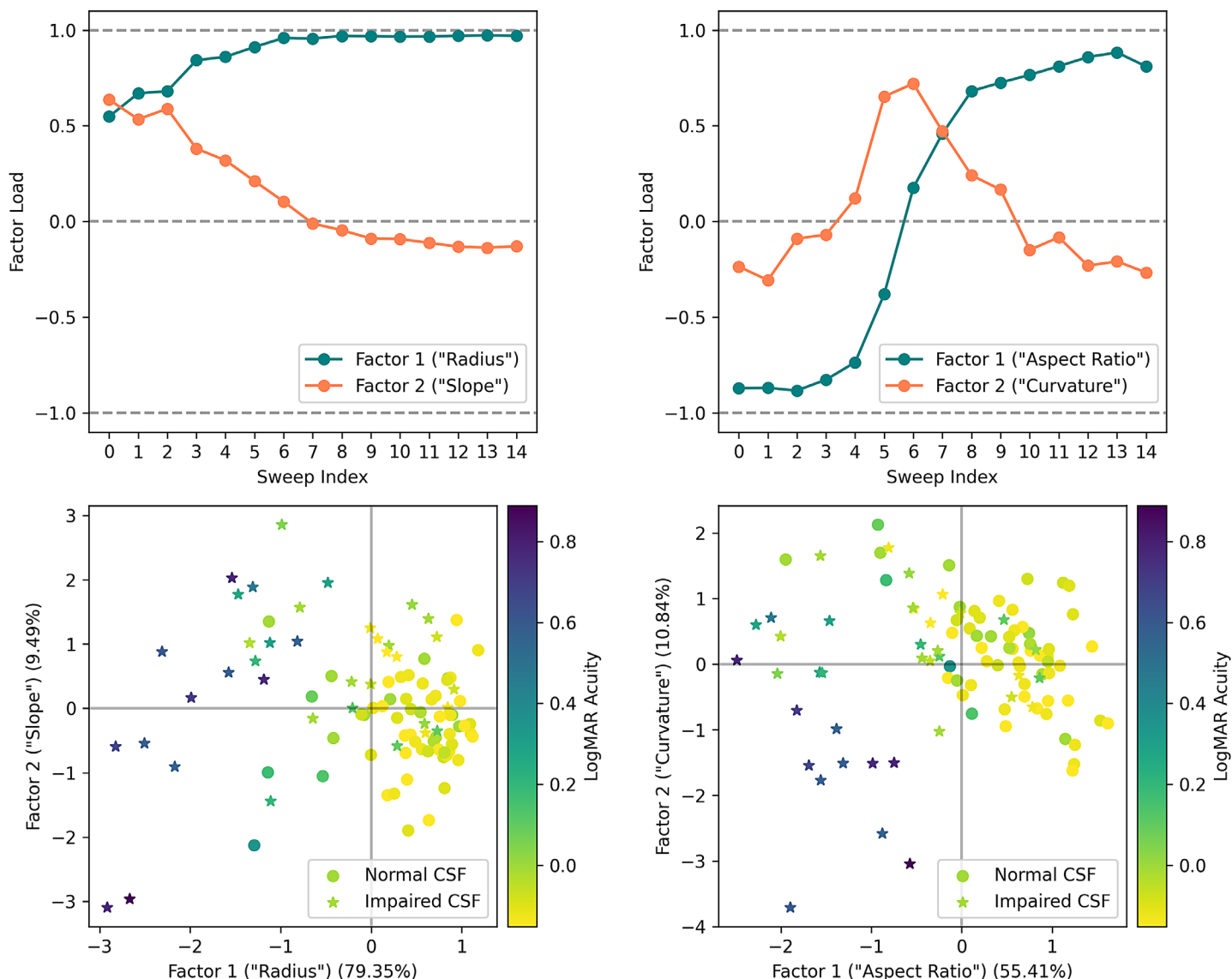
Figure 5. Raw factor analysis (left column) and size-normalized factor analysis (right column) of overall CSF variance. Two factors were retained from each analysis. The plots in the top row depict the loadings of these factors on (i.e., correlations with) the 15 radial sweeps, indexed from 0 to 14. The plots in the bottom row depict the distributions of all CSFs in each two-factor space, with points colored by eye-chart acuity. In these plots, circles represent normal/corrected-to-normal CSFs and stars represent uncorrected CSFs.

with both CSF radius ($r = 0.828$) and slope ($r = -0.532$) from the raw analysis, both $p < 0.001$. Factor 2 accounted for 10.84% of the variance and can be interpreted as the "curvature" of the CSF. Higher scores on this factor indicate that the CSF has a greater radius along the central (more diagonal) sweeps relative to the low-end and high-end sweeps, while lower scores are typical of CSFs that peak earlier and are more buckled in the center. Curvature was correlated positively and moderately with both CSF radius ($r = 0.321$) and slope ($r = 0.291$), both $p < 0.001$.

The scatterplots in the bottom row of Figure 5 reveal how each pair of factors (on the axes) relates to LogMAR eye chart acuity (point color). Each

point represents a normal (circles) or impaired (stars) CSF. In the raw analysis (bottom left panel), almost all normal CSFs had positive radius scores. Impaired CSFs, however, were spread out over negative scores on the radius axis. In a linear regression, CSF radius accounted for 73.6% of the variance in LogMAR acuity ($r = -.858$, $p < 0.001$) among all 98 CSFs. CSF slope accounted for a mere additional 0.1% of variance when added to this regression model, as indicated by the absence of any apparent relationship between acuity and CSF slope in the scatterplot. The scatterplot for the normalized analysis (bottom right panel) resembles the raw analysis but is rotated in this factor space, indicating that better acuity scores are predicted by both greater

CSF aspect ratio and greater curvature. Together, the two orthogonal shape factors account for 67.0% of the variance in LogMAR acuity ($r = .818$, $p < 0.001$), which is only 6.6% less than the variance explained by CSF radius in our previous factor analysis. CSF radius, aspect ratio, and curvature together account for 77.8% of the variance in acuity when combined in one linear model ($r = .882$, $p < 0.001$), which is 4.2% more than radius alone. Figure 6 provides a visualization for how the raw (left column) and normalized (right column) CSFs relate to LogMAR acuity and their two-factor models, which are used to color the curves in each of the three rows.

Finally, we examined the relationship between observer age and the two pairs of CSF components identified above by inspecting which combinations of factors best predicted age. CSF radius accounted for 47.3% of the variance in age across the 60 mean normal CSFs ($r = -0.688$, $p < 0.001$), as shown in the linear regression plot in Figure 7. CSF aspect ratio accounted for 26.6% of variance in age on its own ($r = -0.516$, $p < 0.001$) but only an additional 0.5% of variance when combined with radius. Neither CSF slope nor curvature were significantly correlated with age. When CSF radius was regressed upon both age and corrected-to-normal acuity, rather than age alone, the regression coefficient of age as a predictor of CSF radius decreased from -0.028 radius score change per year of age to -0.019 (still $p < 0.001$), indicating that the relationship between CSF radius and age was only partially mediated by acuity.

Comparison to conventional 4AFC task

The CSFs measured with the conventional task are depicted together with Gradiate CSFs in Figure 8 for the six observers who were assessed with both methods. The uncorrected and corrected CSFs are shown in separate panels for the first five observers. The red curves represent the smoothed Gradiate CSFs, as depicted in Figure 3, and the dotted blue curves represent the smoothed CSFs fitted to the blue 4AFC data points. As we observed in comparisons to the Curveball task in our previous study (Mooney et al., 2018), there is a clear systematic difference in the thresholds reported by the two tasks, even when measured along the same radial sweeps in log-log CSF space. Specifically, Gradiate systematically underestimates spatial vision ability relative to the 4AFC task. As the 4AFC task used the same stimuli as Gradiate, including stimulus motion, this difference can be directly attributed to the eye-tracking methodology: The gains made by observers in the 4AFC task represent degrees of visibility that allow targets to be discriminated from uniform background but not reliably tracked. The solid blue curves in Figure 8 depict

the 4AFC CSFs after applying an affine transformation optimized over all 11 measured CSFs simultaneously. The scaling, shearing, and offsetting effects of this transformation account for a large amount of the disagreement between the two tasks. Figure 9 depicts Bland-Altman plots comparing mean sweep lengths between Gradiate and the untransformed 4AFC task (left) and between Gradiate and the transformed 4AFC task (right). By design, the global transformation naturally eliminates the mean difference between the two tasks, but more important, it reduces the coefficient of repeatability from 0.0575 to 0.0473 in units of normalized sweep length, which is smaller than the CoR between two repeat measures of Gradiate itself (0.0562). This indicates that Gradiate is measuring the same underlying visual ability as the staircase task, although it does so much more quickly: On average, one 8-sweep CSF with the 4AFC task took 10 min and 15 s to measure, which is more than four times greater than the average time taken to measure than one 15-sweep Gradiate CSF.

Discussion

Our findings indicate that the Gradiate task produces reliable, well-defined, detailed CSFs that can provide insight into the underlying dimensions of spatial vision that vary within and between populations. The average time taken to measure a 15-point CSF with Gradiate (2 min, 14 s) was less than half the time needed to measure a 6-point CSF with Curveball (5 min, 15 s) (Mooney et al., 2018) and less than a quarter of the time needed to measure an 8-point CSF with a conventional 4AFC staircase task (10 min, 15 s). No observers were excluded due to poor tracking, which is another improvement over the Curveball task and indicates that Gradiate can accommodate observers who exhibit more frequent saccades while tracking a smoothly moving target. CSFs generated with Gradiate exhibited high agreement with CSFs generated by the conventional staircase task, particularly after applying a global affine transformation to discount the systematic variation between the tracking-based and discrimination-based methodologies. The data also reveal that specific points of interest on the CSF such as LCA can be inferred directly from the full curve with high accuracy, as expected.

We have empirically identified two orthogonal components of the CSF—radius and slope—that account for most of the variance in the 15 sweep length scores that comprise a Gradiate CSF. Unlike CSF radius, slope was not correlated with acuity or age but nevertheless accounted for almost 10% of the overall variance, even among observers with no diagnosed disorders other than refractive error. This demonstrates
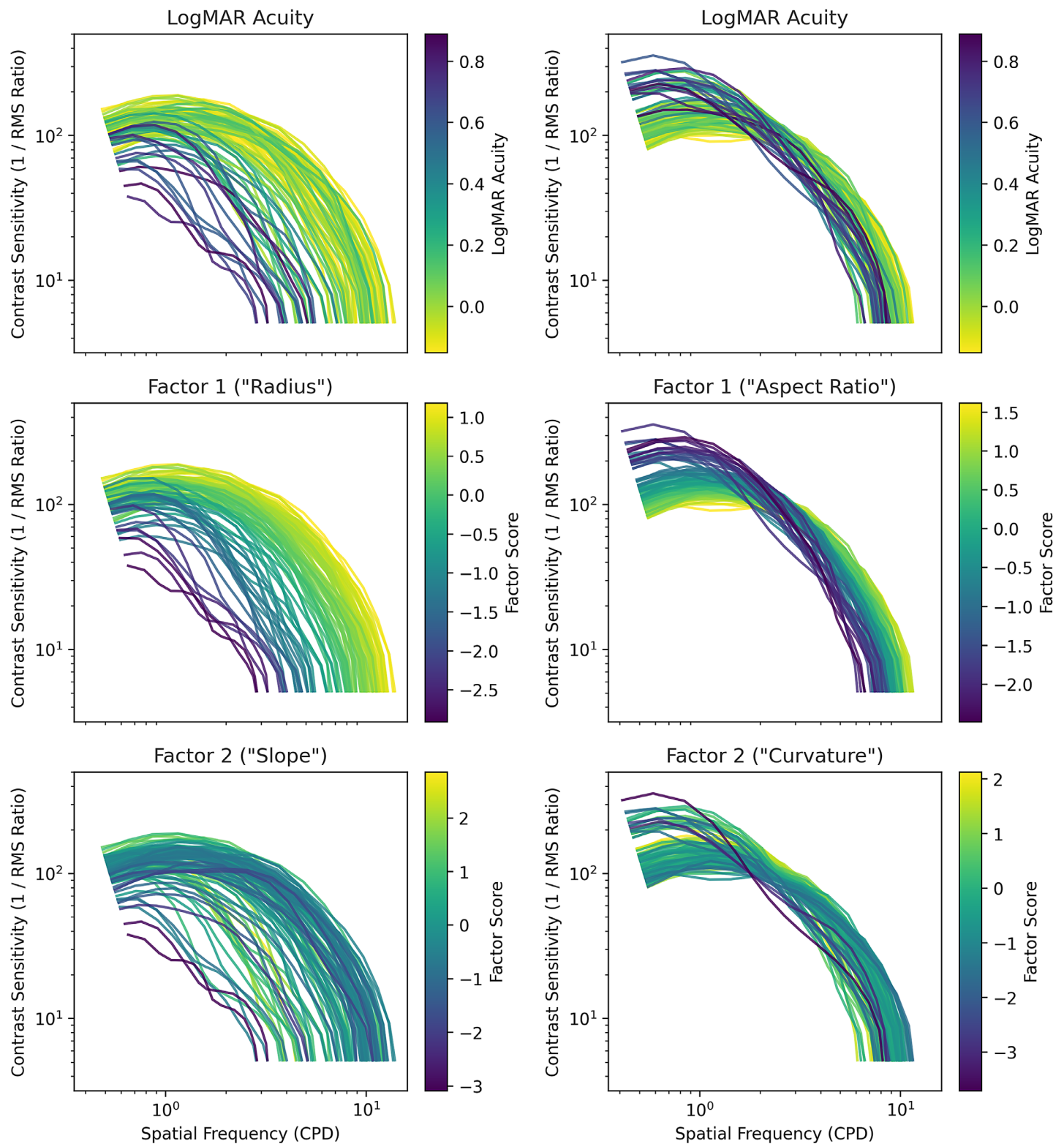
Figure 6. Visualization of the relationship between the identified factors and CSFs. All CSFs from Figure 3 (both normal and impaired) are superimposed in each panel and have additionally had their mean sweep length normalized in the right column. In the top row, the CSFs are colored by acuity, as in the bottom row of Figure 5. In the middle and bottom columns, the CSFs are colored by the first and second factors of the raw (left column) and normalized (right column) factor analyses.
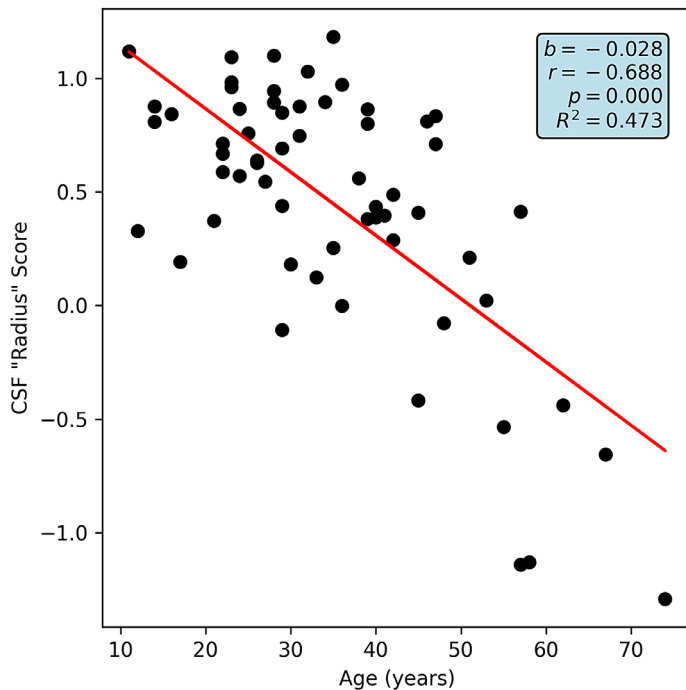
Figure 7. Plot of CSF "radius" score (vertical axis) against observer age in years (horizontal axis) for normal and corrected-to-normal CSFs only. The regression statistics are shown in the top-right and the line of best fit depicted in red.

that CSFs vary significantly between individuals along at least one dimension that is independent of acuity. The slope of the CSF may instead be specifically affected by impaired contrast sensitivity, although its correlation with peak contrast sensitivity ($r = .519$) is still less than that of the CSF radius factor ($r = .833$). We plan to investigate the importance of the CSF's slope in other impaired populations (e.g., glaucoma, age-related macular degeneration, diabetic retinopathy, CVI). Overall CSF size is likely to be sensitive to almost all low-level visual disorders, but for that very reason, it is less likely to be useful for distinguishing one type of disorder from another. The normalized (i.e., shape-based) CSF factors we identified—aspect ratio and curvature—account for almost as much overall CSF variance as the radius factor but are notably much poorer at predicting observer age. We therefore anticipate that these dimensions of CSF shape will be more useful in identifying and subsequently discounting impairment from refractive error than CSF radius (or peak sensitivity, area under the curve, etc.). This distinction is particularly relevant for individuals who cannot complete standard tests of refractive error, and we plan to test this hypothesis in future research with cognitively impaired patients.

We estimated 15-point CSFs from each observer to examine the function's curve in high detail, but our factor analyses provide an empirical basis for even more

efficient variants of the Gradiate procedure. The CSF radius factor, for example, is highly correlated with any of the sweeps that progress toward higher spatial frequencies (top-left panel of Figure 5), peaking at the 14th sweep in clockwise order (i.e., the second to most horizontal sweep in Figure 1), $r = .973$, $p < 0.001$. Measuring this single sweep (which takes only seconds per repeat) would provide most of the information present in the full CSF for healthy participants and would likely be sufficient to determine whether an individual's CSF is impaired by any cause. CSF slope does not load as highly on any one sweep, but further analysis reveals that three sweeps (the 1st, 3rd, and 14th) together predict the slope factor with a correlation above 0.9 ($r = .946$, $p < 0.001$). Similarly, the aspect ratio factor of the normalized-size CSFs is accurately captured by the 3rd and 14th sweeps ($r = .953$, $p < 0.001$), and the curvature factor is mostly captured by the 2nd, 6th, 7th, and 14th sweeps ($r = .932$, $p < 0.001$). This demonstrates that both pairs of CSF factors identified in our data can be accurately estimated from a small subset of the full curve, which is an approach we plan to explore in future work as an alternative to existing "one number" measures that have not been empirically derived from high-resolution CSF samples (such as LCA). Like LCA, these abbreviated tasks may be desirable in situations where time is restricted, where repeated measure quantity is more valuable than individual measure quality, or for basic screening purposes. Of course, it is likely that other dimensions of variation in the CSF will appear when testing populations with impairments other than simple refractive error. We plan to measure detailed CSFs with Gradiate in a variety of clinical populations to identify these factors and describe how they differ between visual disorders.

Finally, we note that other modifications to the task may be necessary for individuals with extensive cognitive impairment. Gradiate replaces the standard verbal or manual feedback with eye movements, but this is not a panacea. Many neurological conditions that affect vision (such as CVI) can cause strabismus, pathological nystagmus, visual neglect, and problems with visual crowding, any of which can make eye-tracking tasks difficult or impossible. Such populations may also simply lack the general attentional or cognitive ability required to track a small, unpredictable target (let alone five such targets simultaneously) or perform even a 1-point gaze calibration with the eye tracker. We are currently testing variants of Gradiate that use single full-screen noise textures and/or limited directions of motion in an effort to capture these observers, who comprise a significant proportion of children with CVI. These variants tend to have a higher false-positive rate due to the absence of a fixation tracking criterion (the cost of their higher true-positive rate for impaired populations), but we anticipate that a balance can
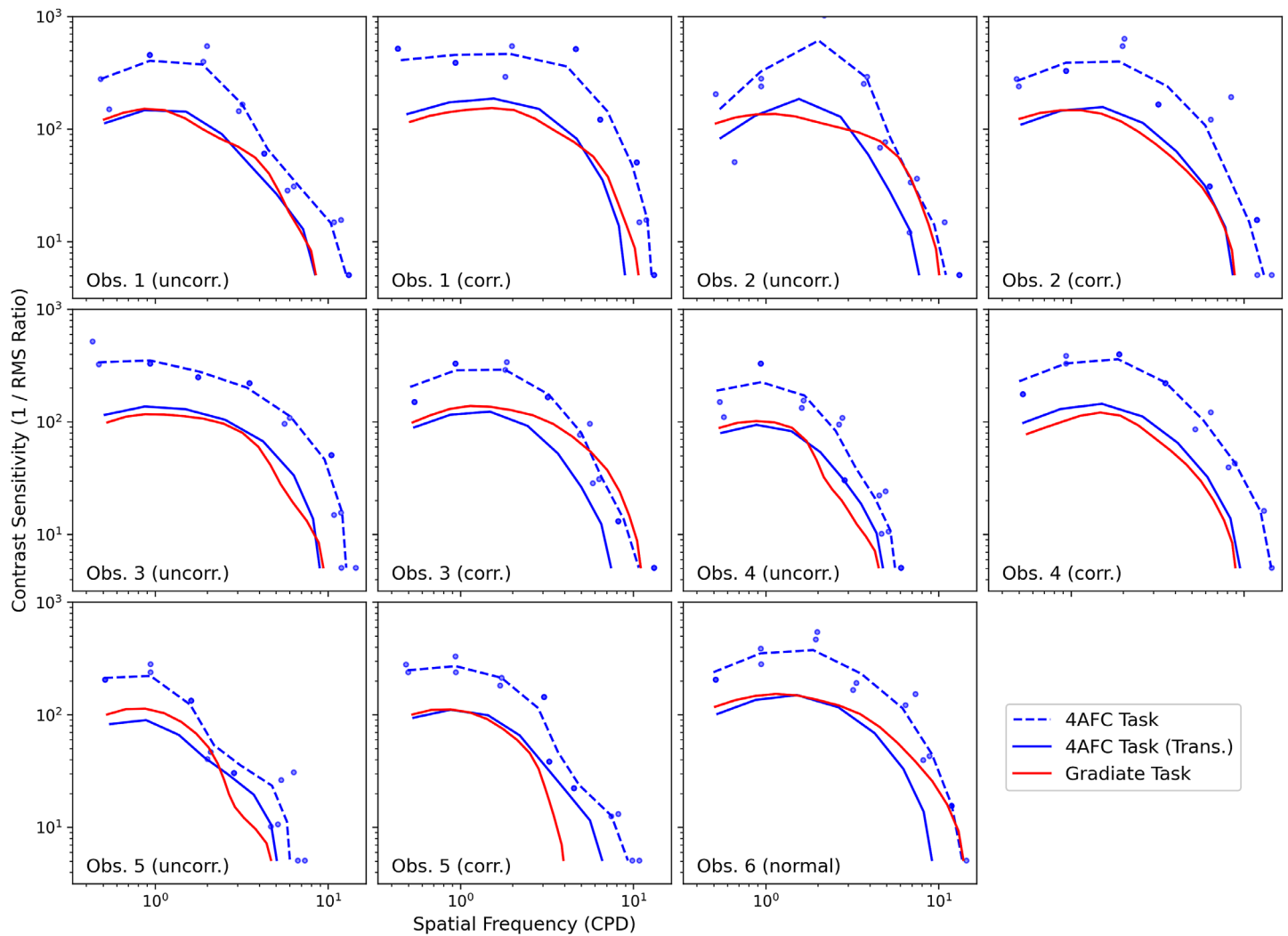
Figure 8. Smoothed Gradiate CSFs (red lines) compared to smoothed 4AFC CSFs (blue data points and dotted blue lines) and the same 4AFC CSFs after applying a global affine transformation (solid blue lines) for six observers. Both uncorrected and corrected CSFs are shown for five observers who require refractive correction. As in previous figures, the horizontal and vertical axes represent log spatial frequency and log RMS contrast sensitivity, respectively.

be reached that maximizes the range of neurological visual disorders that can be compared across impaired and healthy populations of all ages. Some patients, for example, may be able to complete a basic variant of Gradiate that presents only one stimulus at a time; this would sacrifice task efficiency (which could be counteracted by measuring fewer sweeps) but maintains specificity better than a full-screen variant. Such a task would still be more useful than any existing method that is inaccessible to the patient. We also anticipate that the efficiency and intuitiveness of the Gradiate task may make CSF assessment more palatable in general and promote the CSF as a standard, widespread, repeatable measure of visual and neurological health in both hospital settings and the ophthalmologist's office. Our results demonstrate that some individuals with corrected-to-normal eye-chart acuity scores can still

exhibit abnormal CSFs, which helps dispel the myth that acuity is the only dimension of spatial vision that matters for optical correction and speaks to the present underappreciation of comprehensive spatial vision assessment.

As expected, the curves generated by Gradiate exhibit both similarities and differences to CSFs and single thresholds measured in prior studies using different methodologies. Direct quantitative comparison between our polar CSFs and conventional Cartesian CSFs is difficult, but certain shared and disparate features can be readily identified. Like its predecessor Curveball and other pursuit-based tasks (e.g., Dakin & Turnbull, 2016), Gradiate generates normal/corrected-to-normal CSF curves that reliably peak around 1 cpd, which lies between the values typically produced by traditional psychophysical tasks using moving (~0.6 cpd) and
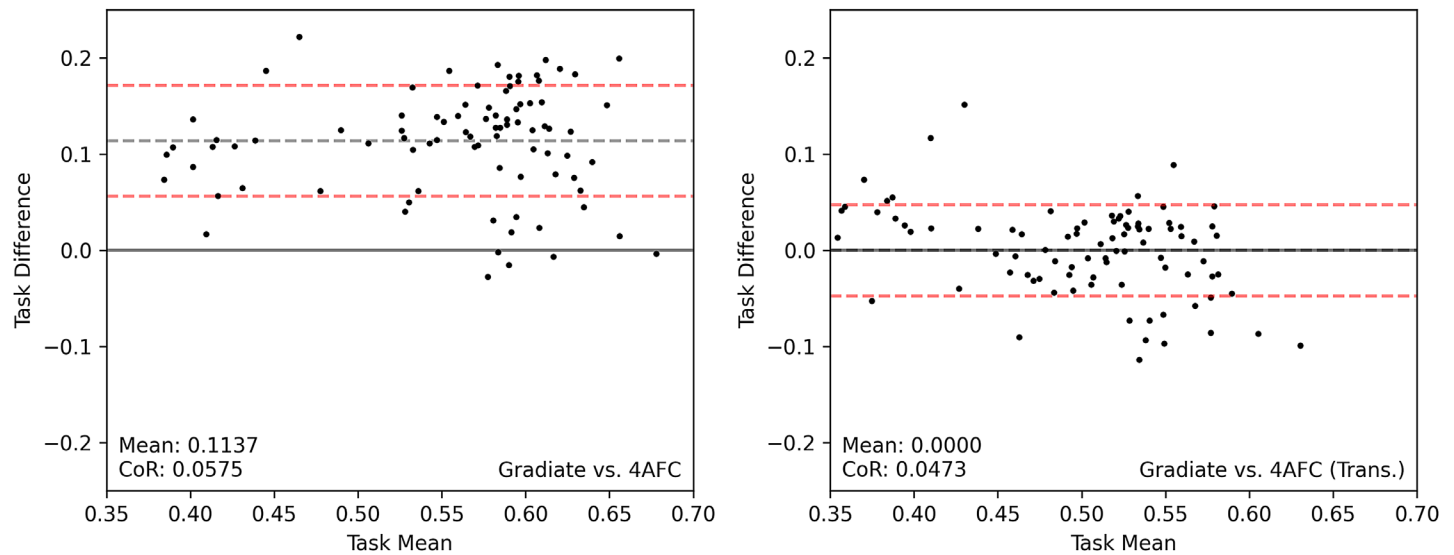
Figure 9. Bland-Altman plots depicting agreement in mean sweep length between Gradiate and the raw 4AFC CSFs (left) or transformed 4AFC CSFs (right). The horizontal and vertical axes represent the mean and difference of corresponding samples from the same observer and sweep, respectively. As in Figure 4, the mean difference and its 95% confidence interval are shown by the gray and red dotted lines, respectively, and the solid black line marks a difference of zero. The mean difference and CoR are also shown in the bottom-left corner.

static (∼2 cpd) stimuli (Burr & Ross, 1982; Mooney et al., 2018). We attribute this to the partial temporal effects induced by the imperfect, ongoing retinal stabilization of moving stimuli in Gradiate. Notably, Gradiate's curves are also downshifted in contrast sensitivity relative to Curveball's, which we attribute to both the use of nonratio RMS contrast values in Curveball (an effective increment of $0.3 \log_{10}$ CS units) and a systematic overestimation of contrast sensitivity by Curveball caused by the continuous fade approach (Mooney et al., 2018). Gradiate also underestimated CSF size relative to a conventional 4AFC staircase task that matched the stimulus dimensions used in Gradiate. This difference likely represents a genuine gulf between stimuli that are visible enough to discriminate above chance when their location is restricted to a few possible positions but not visible enough to track as they move across a wider field. It is not obvious that the 4AFC task's implicit definition of "visible" is more useful or externally valid than Gradiate's gaze-based definition: In natural settings, stimuli are generally less predictable in both position and velocity than they are in forced-choice experiments, and an inability to detect or stabilize a "visible" stimulus on the retina as it moves may significantly impair an individual's capacity to respond to it properly. Nevertheless, the systematic difference between the tasks is easily quantified and places no limitations on Gradiate's usefulness in comparing CSFs between and within individuals.

Finally, we note that multiple observers with refractive error exhibited localized "notches" at medium spatial frequencies in their uncorrected CSF. Several clear examples are visible toward the bottom-left of Figure 3 (red lines). The reliability of their location and shape across repeats within observers, as well as differences between observers, lead us to believe that they are not artifacts of the task. Past studies have shown that notched CSFs can be exhibited by participants with neurological disorders (Regan et al., 1977), but as our participants are cognitively healthy and the notches vanish when their vision is corrected, we instead suggest that these are idiosyncratic deficits caused by distinct refractive/diffractive errors, as have been observed previously (Campbell & Green, 1965). If so, this is affirming for Gradiate's potential usefulness in assessing spatial vision at a detailed level and suggests that it is well positioned to detect any similar notches caused by individual cases of CVI. Other efficient CSF assessment tasks would likely miss these notches due to low threshold resolution, too few points on the CSF, or an inability to model local deviations from a parameterized CSF model (e.g., qCSF; Lesmes et al., 2010). However, further work is required to better characterize these subtle CSF features, confirm their reliability (e.g., through extensive repeat testing of specific observers), and analyze their potential causes in more detail.

Overall, our findings indicate that Gradiate permits a rapid, intuitive, detailed, and time-flexible assessment of the contrast sensitivity function and may offer an opportunity to increase the general penetrance of contrast sensitivity testing into the mainstream. Existing tasks that measure the CSF have largely been held back by lengthy testing procedures and a reliance

on verbal feedback that excludes the most impaired populations from assessment. Gradiate addresses these problems while remaining relatively agnostic about the exact form of the CSF and can be administered without instruction. While some basic eye-tracker calibration is still needed, the task leans on tracking detection algorithms that tolerate a moderate degree of calibration error, which we plan to improve further in the future. Gradiate provides a rational way to move forward in clinical studies aiming to link specific diseases to specific changes in the shape of the CSF, and we plan to employ the task in such future work, including the first quantitative assessment of CSFs in children with cortical visual impairment.

Keywords: contrast sensitivity, smooth pursuit, psychophysics, Gradiate, visual acuity

Acknowledgments

Commercial relationships: none.
Corresponding author: Glen T. Prusky.
Email: glp2004@med.cornell.edu.
Address: Burke Neurological Institute, White Plains, NY, USA.

References

Bonnen, K., Burge, J., Yates, J., Pillow, J., & Cormack, L. K. (2015). Continuous psychophysics: Target-tracking to measure visual sensitivity. *Journal of Vision, 15*(3), 14, https://doi.org/10.1167/15.3.14.

Bradley, A., Hook, J., & Haeseker, J. (1991). A comparison of clinical acuity and contrast sensitivity charts: Effect of uncorrected myopia. *Ophthalmic and Physiological Optics, 11*(3), 218–226, https://doi.org/10.1111/j.1475-1313.1991.tb00536.x.

Brown, B., & Lovie-Kitchin, J. E. (1989). High and low contrast acuity and clinical contrast sensitivity tested in a normal population. *Optometry and Vision Science, 66*(7), 467–473.

Burr, D. C., & Ross, J. (1982). Contrast sensitivity at high velocities. *Vision Research, 22*(4), 479–484, https://doi.org/10.1016/0042-6989(82)90196-1.

Campbell, F. W., & Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *Journal of Physiology, 181*(3), 576–593.

Chandrakumar, M., Colpa, L., Reginald, Y. A., Goltz, H. C., & Wong, A. M. F. (2013). Measuring contrast sensitivity using the M&S Smart System II versus the Pelli-Robson chart. *Ophthalmology, 120*(10), 2160–2161.e1, https://doi.org/10.1016/j.ophtha.2013.07.022.

Cimmer, C., Szendi, I., Csifcsák, G., Szekeres, G., Ambrus Kovács, Z., Somogyi, I., . . .,, . . . Kéri, S. (2006). Abnormal neurological signs, visual contrast sensitivity, and the deficit syndrome of schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 30*(7), 1225–1230, https://doi.org/10.1016/j.pnpbp.2006.03.021.

Collins, J. W., & Carney, L. G. (1990). Visual performance in high myopia. *Current Eye Research, 9*(3), 217–224, https://doi.org/10.3109/02713689009044516.

Dakin, S. C., & Turnbull, P. R. K. (2016). Similar contrast sensitivity functions measured using psychophysics and optokinetic nystagmus. *Scientific Reports, 6*(1), 34514, https://doi.org/10.1038/srep34514.

de Faria, J. M. L., Katsumi, O., Arai, M., & Hirose, T. (1998). Objective measurement of contrast sensitivity function using contrast sweep visual evoked responses. *British Journal of Ophthalmology, 82*(2), 168–173, https://doi.org/10.1136/bjo.82.2.168.

Dorr, M., Lesmes, L. A., Elze, T., Wang, H., Lu, Z.-L., & Bex, P. J. (2017). Evaluation of the precision of contrast sensitivity function assessment on a tablet device. *Scientific Reports, 7*(1), 1–11, https://doi.org/10.1038/srep46706.

Dosso, A. A., Bonvin, E. R., Morel, Y., Golay, A., Assal, J. P., & Leuenberger, P. M. (1996). Risk factors associated with contrast sensitivity loss in diabetic patients. *Graefe's Archive for Clinical and Experimental Ophthalmology, 234*(5), 300–305, https://doi.org/10.1007/BF00220704.

Elliott, D. B., Sanderson, K., & Conkey, A. (1990). The reliability of the Pelli-Robson contrast sensitivity chart. *Ophthalmic and Physiological Optics, 10*(1), 21–24, https://doi.org/10.1111/j.1475-1313.1990.tb01100.x.

Freeman, E. E., Muñoz, B., Turano, K. A., & West, S. K. (2005). Measures of visual function and time to driving cessation in older adults. *Optometry and Vision Science, 82*(8), 765–773, https://doi.org/10.1097/01.opx.0000175008.88427.05.

Gegenfurtner, K. R. (2016). The interaction between vision and eye movements. *Perception, 45*(12), 1333–1357, https://doi.org/10.1177/0301006616657097.

Geruschat, D. R., Turano, K. A., & Stahl, J. W. (1998). Traditional measures of mobility performance and

retinitis pigmentosa. *Optometry and Vision Science, 75*(7), 525–537.

Good, W. V., Jan, J. E., DeSa, L., Barkovich, A. J., & Groenveld, M. (1994). Cortical visual impairment in children. *Survey of Ophthalmology, 38*(4), 351–364, https://doi.org/10.1016/0039-6257(94)90073-6.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377–381, https://doi.org/10.1016/j.jbi.2008.08.010.

Hess, R., & Woo, G. (1978). Vision through cataracts. *Investigative Ophthalmology & Visual Science, 17*(5), 428–435.

Hill, N. J., Mooney, S. W. J., & Prusky, G. T. (2020). *audiomath: A Neuroscientist's Sound Toolkit*. Advance online publication. https://doi.org/10.31234/osf.io/fhma9.

Hill, N. J., Mooney, S. W. J., Ryklin, E. B., & Prusky, G. T. (2019). Shady: A software engine for real-time visual stimulus manipulation. *Journal of Neuroscience Methods, 320*, 79–86, https://doi.org/10.1016/j.jneumeth.2019.03.020.

Huo, R., Burden, S. K., Hoyt, C. S., & Good, W. V. (1999). Chronic cortical visual impairment in children: Aetiology, prognosis, and associated neurological deficits. *British Journal of Ophthalmology, 83*(6), 670–675, https://doi.org/10.1136/bjo.83.6.670.

Hyvärinen, L., Laurinen, P., & Rovamo, J. (1983). Contrast sensitivity in evaluation of visual impairment due to diabetes. *Acta Ophthalmologica, 61*(1), 94–101, https://doi.org/10.1111/j.1755-3768.1983.tb01399.x.

Jones, P. R., Kalwarowsky, S., Atkinson, J., Braddick, O. J., & Nardini, M. (2014). Automated measurement of resolution acuity in infants using remote eye-tracking. *Investigative Ophthalmology & Visual Science, 55*(12), 8102–8110, https://doi.org/10.1167/iovs.14-15108.

Leat, S. J., Yadav, N. K., & Irving, E. L. (2009). Development of visual acuity and contrast sensitivity in children. *Journal of Optometry, 2*(1), 19–26, https://doi.org/10.3921/joptom.2009.19.

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

Mooney, S. W. J., Hill, N. J., Tuzun, M. S., Alam, N. M., Carmel, J. B., & Prusky, G. T. (2018). Curveball: A tool for rapid measurement of contrast sensitivity

based on smooth eye movements. *Journal of Vision, 18*(12), 7, https://doi.org/10.1167/18.12.7.

Owsley, C. (2003). Contrast sensitivity. *Ophthalmology Clinics, 16*(2), 171–177, https://doi.org/10.1016/S0896-1549(03)00003-8.

Pelli, D. G., Robson, J. G., & Wilkins, A. J. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences, 2*(3), 187–199.

Pelli, Denis G., & Bex, P. (2013). Measuring contrast sensitivity. *Vision Research, 90*, 10–14, https://doi.org/10.1016/j.visres.2013.04.015.

Pesudovs, K., Marsack, J. D., Donnelly, W. J., Thibos, L. N., & Applegate, R. A. (2004). Measuring visual acuity-mesopic or photopic conditions, and high or low contrast letters? *Journal of Refractive Surgery, 20*(5), S508–S514, https://doi.org/10.3928/1081-597X-20040901-20.

Regan, D. (1988). Low-contrast visual acuity test for pediatric use. *Canadian Journal of Ophthalmology, 23*(5), 224–227.

Regan, D., Raymond, J., Ginsburg, A. P., & Murray, T. J. (1981). Contrast sensitivity, visual acuity and the discrimination of Snellen letters in multiple sclerosis. *Brain, 104*(2), 333–350, https://doi.org/10.1093/brain/104.2.333.

Regan, D., Silver, R., & Murray, T. J. (1977). Visual acuity and contrast sensitivity in multiple sclerosis—Hidden visual loss: An auxiliary diagnostic test. *Brain, 100*(3), 563–579, https://doi.org/10.1093/brain/100.3.563.

Ross, J. E., Bron, A. J., & Clarke, D. D. (1984). Contrast sensitivity and visual disability in chronic simple glaucoma. *British Journal of Ophthalmology, 68*(11), 821–827, https://doi.org/10.1136/bjo.68.11.821.

Savoy, R. L., & McCann, J. J. (1975). Visibility of low-spatial-frequency sine-wave targets: Dependence on number of cycles. *Journal of the Optical Society of America, 65*(3), 343–350, https://doi.org/10.1364/JOSA.65.000343.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision, 11*(5), 9, https://doi.org/10.1167/11.5.9.

Spering, M., & Carrasco, M. (2015). Acting without seeing: Eye movements reveal visual processing without awareness. *Trends in Neurosciences, 38*(4), 247–258, https://doi.org/10.1016/j.tins.2015.02.002.

Spering, M., & Montagnini, A. (2011). Do we track what we see? Common versus independent processing for motion perception and smooth pursuit eye movements: A review. *Vision Research, 51*(8), 836–852, https://doi.org/10.1016/j.visres.2010.10.017.

Stamper, R. L. (1984). The effect of glaucoma on central visual function. *Transactions of the American Ophthalmological Society, 82*, 792–826.

Teller, D. Y., McDonald, M. A., Preston, K., Sebris, S. L., & Dobson, V. (1986). Assessment of visual acuity in infants and children; the acuity card procedure. *Developmental Medicine & Child Neurology, 28*(6), 779–789, https://doi.org/10.1111/j.1469-8749.1986.tb03932.x.

Thomas, B. B., Seiler, M. J., Sadda, S. R., Coffey, P. J., & Aramant, R. B. (2004). Optokinetic test to evaluate visual acuity of each eye independently. *Journal of Neuroscience Methods, 138*(1), 7–13, https://doi.org/10.1016/j.jneumeth.2004.03.007.

Wender, M. (2007). Value of Pelli-Robson contrast sensitivity chart for evaluation of visual system in multiple sclerosis patients. *Neurologia i Neurochirurgia Polska, 41*(2), 141–143.

West, S. K., Rubin, G. S., Broman, A. T., Muñoz, B., Bandeen-Roche, K., & Turano, K. (2002). How does visual impairment affect performance on tasks of everyday life? The SEE project. *Archives of Ophthalmology, 120*(6), 774–780, https://doi.org/10.1001/archopht.120.6.774.

Witton, C., Talcott, J. B., & Henning, G. B. (2017). Psychophysical measurements in children: Challenges, pitfalls, and considerations. *PeerJ, 5*, e3231, https://doi.org/10.7717/peerj.3231.

Zimmern, R. L., Campbell, F. W., & Wilkinson, I. M. (1979). Subtle disturbances of vision after optic neuritis elicited by studying contrast sensitivity. *Journal of Neurology, Neurosurgery & Psychiatry, 42*(5), 407–412, https://doi.org/10.1136/jnnp.42.5.407.

Supplementary Material

Supplementary Movie S1. Video of one 5-stimulus trial of Gradiate, which produces a 5-point CSF in less than 30 seconds. The five noise stimuli progress their spatial frequency and contrast along predefined radial sweeps in log-log CSF space when they are tracked by the observer's gaze (green dot). Each stimulus successfully tracked produces a musical note as feedback. The inset plot in the bottom-left shows points on these sweeps (blue) being tracked in real time. The stimuli move in a way that avoids collisions, and the stimulus that is currently being tracked makes smooth rather than abrupt turns as often as possible. The trial ends when no stimuli are tracked for an extended period. In the study, fifteen sweeps were measured per CSF, divided over three trials of five simultaneous stimuli.