The Bouma law accounts for crowding in fifty observers

Jan W. Kurzawski[1], Augustin Burchell[3], Darshan Thapa[2], Jonathan Winawer[1,2], Najib J. Majaj[2], Denis G. Pelli[1,2]

[1]Department of Psychology, New York University, New York, NY, USA
[2]Center for Neural Science, New York University, New York, NY, USA
[3]Cognitive Science & Computer Science, Swarthmore College

Corresponding Author: Denis G. Pelli, Department of Psychology, 6 Washington Place, New York University, NY, 10003, USA (denis.pelli@nyu.edu)

Conflict of interest: The authors declare no competing financial interests.

Keywords: crowding, critical spacing, crowding distance, Bouma law, object recognition, statistics of crowding, asymmetries around the visual field.

Funding: This research was funded by the US National Eye Institute R01-EY027964 to DGP and R01-EY031446 to NJM. It was also supported by an NIH core vision grant P30-EY013079.

Abstract

Crowding is the failure to recognize an object due to surrounding clutter. We surveyed visual crowding by measuring 13 crowding distances in each of 50 observers. The 13 measurements reflected combinations of three eccentricities (0, 5, and 10 deg), four cardinal meridians, two orientations (radials and tangential), and two fonts (Sloan and Pelli). We also measured foveal acuity. Remarkably, fitting a two-parameter model, the well-known Bouma law — crowding distance grows linearly with eccentricity — explains 82% of the variance for all 13 x 50 measured log crowding distances, cross-validated. An enhanced Bouma model, with factors for meridian, flanking direction, target kind, and observer explains 94% of the variance, again cross-validated. These additional factors reveal several asymmetries which can be expressed as ratios of crowding distance: 0.62 horizontal:vertical, 0.77 lower:upper, 0.79 right:left, 0.63 tangential:radial, and 0.79 Sloan font:Pelli font. The three meridional asymmetries and the orientation asymmetry are consistent with prior reports. In our results, peripheral crowding is independent of foveal crowding and acuity. At first glance, our best estimate of the radial Bouma factor on the right meridian (0.19), as in other modern studies, seemed lower than Bouma's original reports (0.5 and 0.4). In fact, the many studies from which the Bouma factor can be estimated differ in several parametric ways that one can correct for. After correcting for threshold criterion, number of choice alternatives, and flanker placement, Bouma's estimate becomes 0.35, and five representative studies, including ours, have an average Bouma factor of 0.34, showing that Bouma's study remains prototypical. We found that guaranteeing fixation by gaze-contingent display halved the standard deviation across observers of the log Bouma factor. The reduction in standard deviation is explained by a peeking model that assumes that in 40% of unmonitored fixation trials the observer looked near the anticipated target location. Individual differences are robust as evidenced by the nearly 3:1 ratio of 0.16 SD across observers to 0.06 SD of test-retest. This enhances crowding's promise as a biomarker for dyslexia and visual health.

Introduction

Crowding is the failure to recognize an object due to surrounding clutter (Bouma, 1973; Pelli et al., 2004; Pelli & Tillman, 2008; Strasburger et al., 1991; Stuart & Burian, 1962). Crowding has been studied with several different tasks including letter identification (Bouma, 1970; Flom et al., 1963), Vernier acuity (Levi et al., 1985; Westheimer & Hauske, 1975), face recognition (Louie et al., 2007; Martelli et al., 2005), and orientation discrimination (Andriessen & Bouma, 1976; Toet & Levi, 1992; Westheimer et al., 1976). It is invariant with the size of target and flankers (Levi & Carney, 2009; Pelli et al., 2004; Pelli et al., 2007; Strasburger et al., 1991; Tripathy & Cavanagh, 2002). Crowding is usually measured by sandwiching the target between two similar flanking objects, or *flankers*, and is characterized by the *crowding distance* (or "critical spacing"), which is the center-to-center distance from target to flanker at which recognition attains a criterion level of performance. Crowding distance increases linearly with eccentricity (Bouma, 1970; Kooi et al., 1994; Levi & Carney, 2009; Pelli et al., 2004; Toet & Levi, 1992), and increases with target-flanker similarity (Andriessen & Bouma, 1976; Chastain, 1982; Kooi et al., 1994; Leat et al., 1999; Nazir, 1992; Pelli et al., 2004) as well as the number of distractors (Felisbert et al., 2005; Grainger et al., 2010; Strasburger et al., 1991). Crowding also occurs for moving stimuli (Bex & Dakin, 2005; Bex et al., 2003). For review on crowding literature see (Herzog et al., 2015; Levi, 2008; Pelli & Tillman, 2008; Strasburger, 2020; Strasburger et al., 2011; Whitney & Levi, 2011). Among the normally sighted, crowding was initially reported in the periphery, and, after long debate, has now been convincingly demonstrated in the fovea (Coates et al., 2018; Flom et al., 1963 1963; Liu & Arditi, 2000; Malania et al., 2007; Pelli et al., 2016; Siderov et al., 2013; Toet & Levi, 1992). Crowding is one of several tasks whose threshold increases linearly with radial eccentricity, and such a task can be summarized by an E_2 value that is the eccentricity at which threshold reaches twice its foveal value (Levi et al., 1985).

We are interested in relating psychophysical measures of crowding to brain physiology, especially cortical magnification measured by fMRI (functional Magnetic Resonance Imaging) in areas V1, V2, V3, and hV4. For this purpose, we tested crowding in 50 observers to characterize the statistics of crowding, looking at parameters across individuals. The comparison with fMRI will be reported separately (Kurzawski et al., 2021). Here we report only the psychophysics. We tested with letters, which, after little or no training, provide many possible targets, which is needed for quick testing (Pelli & Robson, 1991). Long term, we are interested in testing crowding in children (Waugh et al., 2018) as an early biomarker for susceptibility to visual problems like dyslexia. Crowding distance is highly conserved across object kind (Kooi et al., 1994; Pelli & Tillman, 2008), which suggests that letters, vernier, and gabors may give similar mean crowding distances. On the other hand, some previous work speculated that crowding might affect letters and symbols differently (Grainger et al., 2010).

Crowding exhibits several striking asymmetries. Crowding distance measured radially (along a line passing through fovea) is roughly twice that measured tangentially; the orthogonal direction (Greenwood et al., 2017; Kwon et al., 2014; Pelli, 2008; Petrov & Meleshkevich, 2011; Toet & Levi, 1992). Crowding distance has often been reported to be smaller in the lower visual field (Fortenbaugh et al., 2015; Greenwood et al., 2017; He et al., 1996; Petrov

& Meleshkevich, 2011) and on the horizontal midline (Chung, 2013; Coates et al., 2021; Liu et al., 2009; Petrov & Meleshkevich, 2011; Toet & Levi, 1992; Wallis & Bex, 2012).

Crowding distance is a potentially valuable biomarker for several reasons. Crowding severely limits what we see, how fast we read, and is correlated with dyslexia. There are large individual differences in crowding distance and correspondingly large physiological differences in the sizes of relevant areas of visual cortex, which invite analysis by correlation (Kurzawski et al., 2021). Here we measured crowding in 50 observers. Previous in-person crowding surveys (Grainger et al., 2010; Greenwood et al., 2017; Petrov & Meleshkevich, 2011; Toet & Levi, 1992) included at most 27 observers. The only remote crowding survey tested 793 observers but did not report any spatial variations (Li et al., 2020). The above cited works used various kinds of stimuli, including letters of various fonts. The original reports of the crowding phenomenon were mostly letter-based (Anstis, 1974; Bouma, 1970, 1973; Ehlers, 1936; Ehlers, 1953; Korte, 1923; Stuart & Burian, 1962), historical review of crowding is described elsewhere (Strasburger, 2020; Strasburger et al., 2011). Here, we too use letters, because they do not require training, and provide a large number of stimulus alternatives, which speeds threshold estimation in laboratory and clinical testing (Pelli et al., 1988).

Whether crowding can be explained by neural computations in any particular cortical location remains unknown, but several candidate areas have been suggested: V1 (Millin et al., 2014), V2 (Freeman & Simoncelli, 2011; He et al., 1996), V3 (Bi et al., 2009; Tyler & Likova, 2007), V4 (Burchell et al., 2019; Liu et al., 2009; Motter, 2006; Zhou et al., 2017) and higher order areas (Aghdaee, 2005; Louie et al., 2007). Recently, crowding distances were found to differ when stimulus properties were tuned to stimulate either the parvo or magno cellular pathway (Atilgan et al., 2020). Finding that crowded (vs uncrowded) stimuli produce suppression of the BOLD signal in V1 suggests that crowding distance might be partly explained by low-level mechanisms (i.e., the size of the receptive fields) and thus related to acuity. Both crowding and acuity increase with eccentricity, which might suggest a common physiological origin. However, although acuity and crowding distance both grow linearly with eccentricity, the two lines have very different intercepts with the eccentricity axis, which seems inconsistent with a common origin (Latham & Whitaker, 1996; Petrov & Meleshkevich, 2011; Rosenholtz, 2016; Song et al., 2014; Strasburger, 2020). Here, we use our data from 50 observers to study the relationship between acuity and crowding in the fovea.

In our 50 participants, we measured 13 crowding distances (12 in the periphery and 1 in the fovea) and acuity in the fovea. We used the Sloan and the Pelli font. As far as we know, Pelli is still the only font skinny enough to measure crowding distance in the fovea. We also assessed crowding's variation along the four cardinal meridians, with flanking direction and across individuals. We find that the crowding varies across meridians, producing several asymmetries and report two-fold individual differences.

METHODS

Measuring crowding

Our goal was to acquire data on many participants. Here, we describe two ways of approaching data collection. Details about Quest and stimulus presentation were similar for both and are described in the sections below. Here we focus on the differences. In the first approach (*unmonitored fixation*) we acquired data on multiple subjects at the same time on laptop computers without the use of eye tracker. The experimenter was not present in the room during data collection. We measured the viewing distance only once in the beginning of each session and participants used the keyboard to provide their responses. We did not use the chin rest or forehead support. Next trial was displayed immediately after the subjects' response. Participants were completely naive to the task and did not do any training. Data was acquired in two blocks where we always interleaved the target position between left and right meridian or upper and lower meridian. In that way we acquired data on 100 participants at 4 cardinal meridians with the Sloan font and radial flankers. This way of collecting the data was motivated by the fact that psychophysics is shifting to online experiments where many aspects cannot be controlled.

In the second approach to data collection (*awaited or guaranteed fixation*), to start the next trial participants had to continuously fixate for at least 250 ms, we followed a more conventional method of measurement and used a more controlled environment. Data was always acquired from one participant at a time and the experimenter was always present during the data acquisition. Viewing distance was measured once before each session and we used the forehead support and the chinrest, so participants did not move their heads during data collection. This time participants responded with a mouse click (letter choices appeared on the screen after each trial) to minimize looking at the keyboard. We used a gaze tracker to discard fixations that were 1.5 deg away from the fixation cross. For each trial in which bad fixation was detected, participants saw a red fixation cross, and another trial was added to the quest sequence. Each participant had to do at least 10 correct trials before they started the experiment. All observers were introduced to the stimuli, locations and fonts before the experiment. We now randomly interleaved all 4-meridian position within one block and again acquired data with Sloan font and radial flankers. Each trial was followed by a 1.2 s delay, so participants had time to fixate back at the fixation cross. For each measured location we acquired two thresholds to estimate test-retest reliability.

By comparing these two methods (thresholds acquired at 5 deg of eccentricity and 2 meridians – left and right) we see obvious differences in the mean and distribution of the Bouma factor (**Figure 1**). Using the first method the mean of log Bouma factor was -0.77 (0.17) with a log SD of 0.33. In the second method the mean was equal to -0.60 (0.25) and the SD dropped by almost half to 0.16. Our interpretation of the broader histogram and lower mean Bouma factor in *unmonitored fixation* is that observers occasionally "peek", that is fixate on the anticipated location of the target instead of the fixation cross as instructed. This introduces a reduced Bouma factor and results in a broader distribution. In the next paragraph we show a peeking model where we prove that looking directly at the target location decreases the mean of Bouma factor and increases the SD, consistent we observed results.

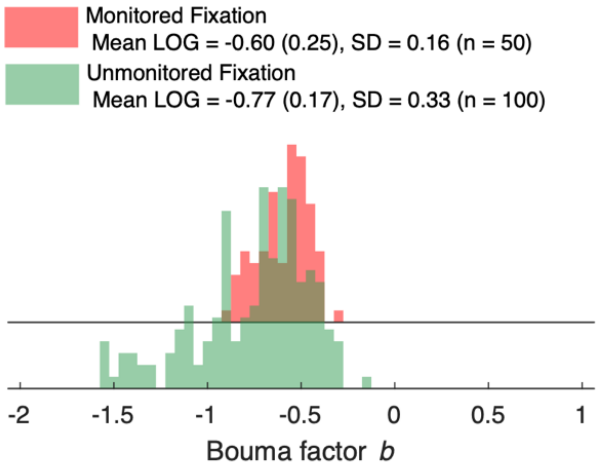

Figure 1. Histograms of Bouma factor estimated by two methods of measurement. Each histogram shows Bouma factor from left and right meridian at 5 degrees of eccentricity.

In this paper we focus on the data acquired with the second method as it is a more standard approach. Data can be downloaded from OSF (future osf link).

The peeking-observer model

We wondered how the estimated Bouma factor depends on fixation accuracy. People rarely do fine discrimination in the periphery. Despite instructing observers to fixate on the central cross, we know that gaze could be elsewhere during the target presentation. The observer is torn between a desire to follow instructions to fixate the cross and a natural impulse to fixate an anticipated target location. We model this by two distributions, first, *the awaited fixation distribution* (participants had to fixate in the central crosshair for at least 250 ms to star the trials), which is compact and centered on the fixation cross, and, second, an assumed *"peeking" distribution* with modes at each possible target location. We use these distributions to build a peeking-observer model. This model assumes that observers peek on a certain fraction of the trials and on these trials eye position is a random sample from the peeking distribution. In the spirit of the Bouma law, our peeking-observer model assumes that the probability of identifying the target depends solely on the ratio of target-flanker spacing to actual eccentricity.

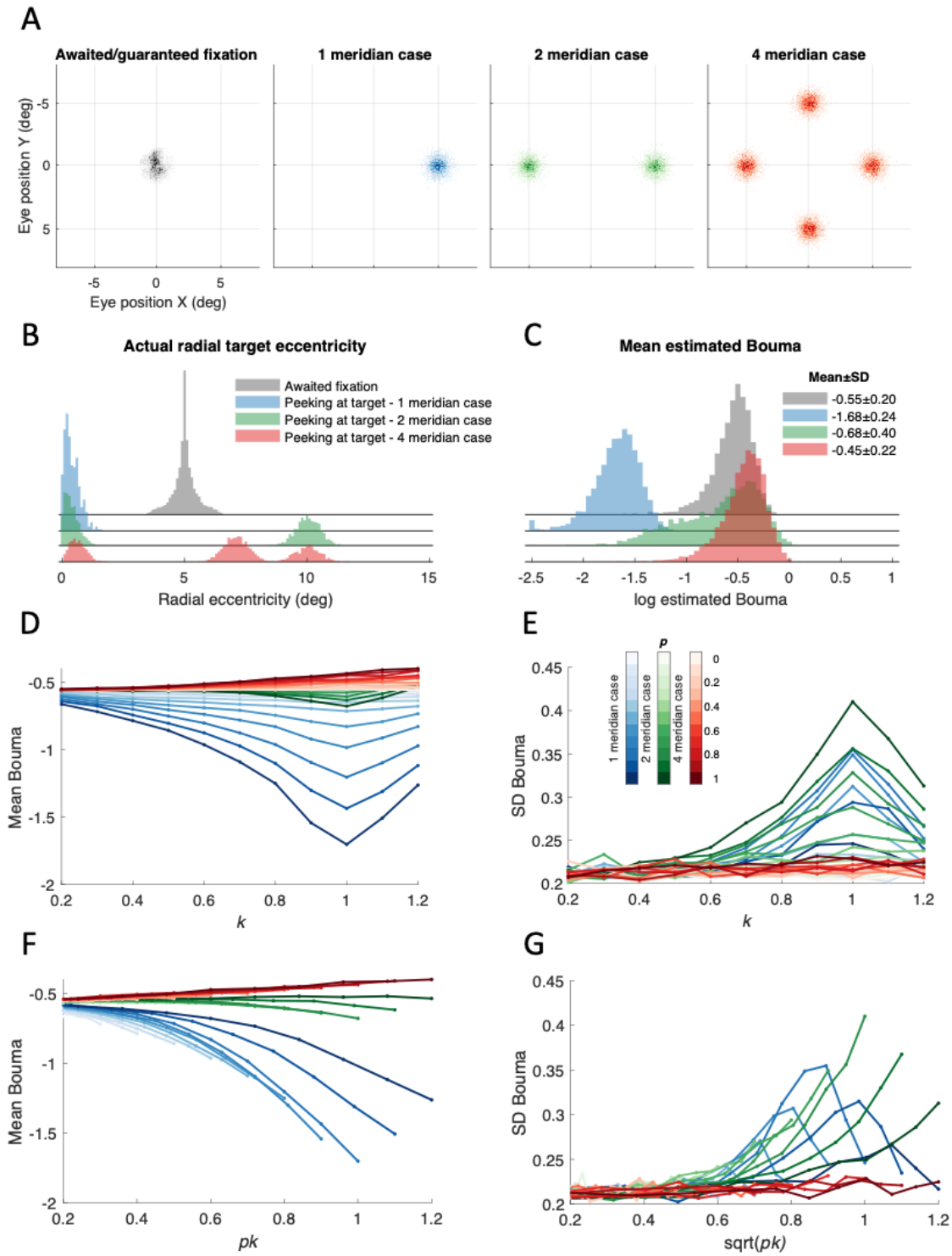

Figure 2. The peeking model. A) Plots show distributions of eye position depending on how many target locations were present in one testing block. Gray histogram was extracted from the Eye-link Eye tracker while colored distributions were simulated. Simulated eye position was a 2D gaussian cantered at target location with SD = 0.5 deg. **B)** Four histograms showing actual radial eccentricity of the target based on the histograms from panel A. **C)** 5000 Bouma factors estimated using Quest where at each trial we substituted target eccentricity with its actual eccentricity corrected for eye position. **D)** Mean Bouma factor for each simulated case plotted vs. peeking eccentricity. We use saturated colors to indicate what was the percentage of trials at

which observers peak. **E)** SD of Bouma factor vs. peeking eccentricity. **F-G)** Same as Panel D and E however Mean Bouma and SD of Bouma is plotted vs combination of peeking probability with estimated eccentricity.

We estimated *awaited fixation distribution* from our measured eye positions which were restricted to be within 1.5 deg from center of the crosshair. We used 3500 actual eye-positions (35 trials x 50 participants) measured with an EyeLink eye tracker.

We considered several cases that varied in the number of possible target locations (**Figure 2A**). First, target was always presented at one location (right meridian at 5 deg). Second – target was randomly presented at ±5 deg on horizontal meridian, third – target was randomly at ±5 radial eccentricity on one of the four meridians (right, left, upper, lower). When participants "peek", depending on the number of possible locations, they have a 100%, 50%, or 25% chance of being near the target location. Of course, gaze position affects target eccentricity.

Nominal eccentricity of the target is relative to the crosshair. *Actual* eccentricity of the target is relative to gaze position when the target is presented. Peeking near the target position will practically zero the actual eccentricity, while fixating at a wrong target location could produce an actual eccentricity greater than nominal. **Figure 2B** shows the actual target eccentricities calculated based on the four distributions of gaze position in **Figure 2A**. The nominal radial eccentricity of the target is always 5 deg, so we simulate an observer with a threshold spacing of 1.5 deg by using a psychometric function provided by Quest (beta = 2.30, delta = 0.01, gamma = 0.11). Thus, our model assumes a Bouma factor of 1.5/5 = 0.3 which corresponds to log Bouma of -0.52. This is the parameter we try to estimate. The point of this exercise is to evaluate how various methods estimate the true Bouma factor. We simulate a block of 35 trials using Quest to estimate the 70%-correct spacing threshold from which we calculate the Bouma factor. We repeat this many times to get a histogram of estimated Bouma factor (**Figure 2C**) for *awaited fixation* (gray) and *peeking* (colored) distributions.

Our modelling shows that the mean Bouma factor is lowest for *peeking* at one location and highest for *peeking* at 4 locations (log Bouma -1.68 vs. -0.45), whereas the true log Bouma factor is -0.52. Using eye position from *awaited fixation* distribution the model estimate of log Bouma factor is -0.55, matching the true value of -0.52. The standard deviation is highest for *peeking* at two locations and lowest for 4 locations (log Bouma 0.40 vs. 0.22).

In order to model the data acquired with *awaited fixation* and *unmonitored* fixation, we extended our peeking model to have two distributions, so eye position is sampled from the *peeking* distribution on a proportion *p* of trials and otherwise from the *awaited fixation* distribution. The model has a shortfall proportion parameter *k* that allows for the possibility that observers systematically undershoot the actual target eccentricity, so that the modal gaze position is a fraction k of nominal target eccentricity.

The error in estimated Bouma grows with the proportion *p* and fraction *k* of peeking (**Figure 2D-E**). Observing that the error of estimated Bouma factor grows proportionally with *k* and that the SD grows proportionally with k^2, we produced new figures (**Figure2 F-G**) showing that mean is approximately determined by kp and SD by sqrt(p)k.

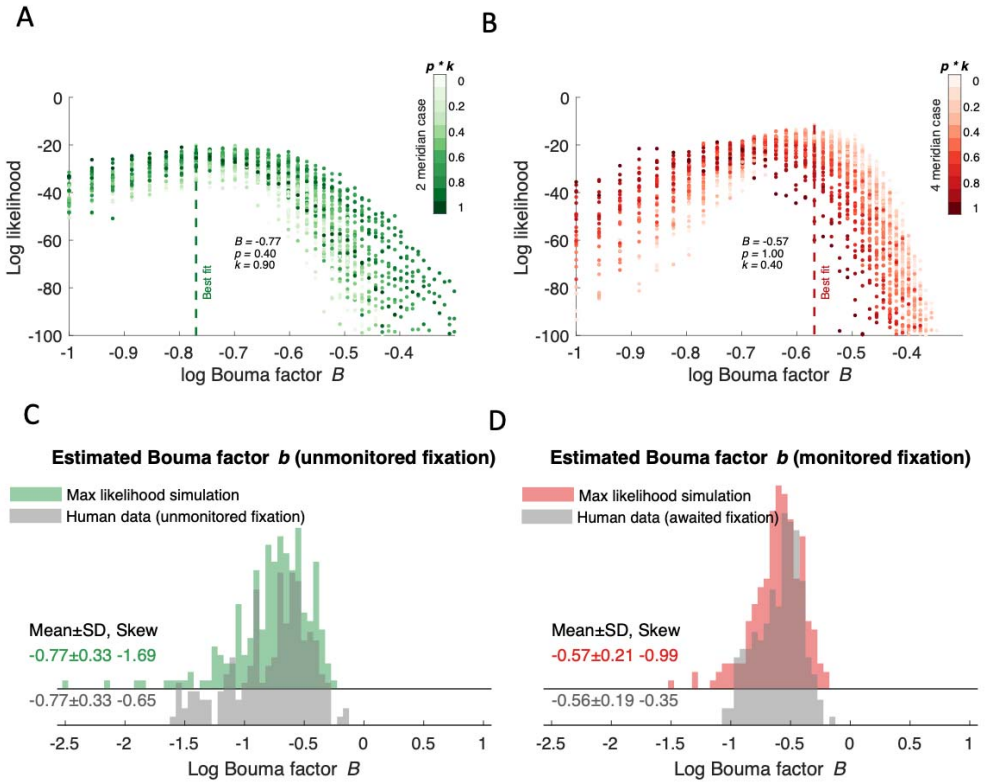

Figure 3. Comparing the peeking model with data. We used log-likelihood estimation to find which model fits our two methods of measurement best. The higher the likelihood the better the fit. **A-B)** Log-likelihood values are plotted vs. estimated log Bouma factor *B*. We use saturated colors to indicate the product of *p* and *k*. **C-D)** Compares histogram of best fitting simulated and acquired data. For each histogram we show the mean, SD, and skew. Note that the red distribution in panel D is different from red distribution in Figure 1. Here we plot data from all four meridians. Data plotted in Figure 1 is extracted from right and left meridian only.

Finally, we used maximum likelihood to discover which model parameters best fit the human data. We wanted to know how far the observers peek *k*, how often *p* and what is the log Bouma factor *B* predicted by the model. Data acquired with *unmonitored* fixation are best fit by simulated data with $k = 0.9$, $p = 0.40$, and $B = -0.77$ (log-like = -20.1599, **Figure 3A**). Observers were peeking at full target eccentricity on 40% of trials. Data acquired with *monitored* fixation is best fit by simulated data with $k = 0.40$, $p = 1$, and $B = -0.57$ (log-like. = -11.2005, **Figure 3B**). They peeked close to the range allowed by the eye tracker. **Figure 3C-D** shows that the simulated histograms are a good match to the human data for both methods. The mean and standard deviation match and the skews have the same sign. Our models of eye position and crowding together provide a simulation of estimated Bouma factors that fits both cases: good fixation and high peeking. This exercise shows that one simple model of crowding fits all our data.

It is our impression that as observers gain experience with peripheral viewing, they peek less. For example, Bouma's original paper reports a Bouma factor of 0.5 which is higher than later estimates of the Bouma factor (Coates et al., 2021). Based on **Figure 3**, Bouma's peeking rate must have been practically zero.

Crowding dataset

Data were acquired with the *CriticalSpacing* software (Pelli et al., 2016) using Quest (Watson & Pelli, 1983), allowing for reliable and relatively fast measurement of crowding distances. Our crowding database consists of measurements of crowding distance with the Sloan font (with radial and tangential flankers) and with the Pelli font (radial flankers) in 50 observers. With Sloan font we measured crowding at 8 different peripheral locations in the visual field: 2 eccentricities (5, 10 deg) along the four cardinal meridians (upper, lower, left and right). Sloan tangential crowding was only measured at 5 deg of eccentricity on the left and right meridian. With Pelli font we measured crowding at the fovea and on the left and right meridian (5 deg). The Sloan font acuity size is too big to allow measuring foveal crowding distance in adults. The Pelli font was specially designed for measuring foveal crowding distance (Pelli et al., 2016). Additionally, we measured acuity in the fovea. Spatial map of testing is shown in **Figure 6**. To estimate test-retest reliability of our measurements we measured each threshold twice in two sessions. Sessions were scheduled at least a day apart with a maximum of 5 days apart. We report our results as Bouma factor *b* (slope of crowding distance vs. eccentricity) estimated from **Equation 7** where crowding distance is defined as the required center-to-center spacing (in deg) for 70% correct report of the middle letter in a triplet.

Participants

Our dataset is described in **Table 1**, and its spatial coverage of the visual field is plotted in **Figure 6**. The study tested 50 observers, mostly New York University undergraduate students (mean age = 23). Each observer had normal or corrected-to-normal vision. All experiments were conducted in accordance with the Declaration of Helsinki and were approved by New York University's ethics committee on activities involving human observers. In all analyses except the test-retest section, we average thresholds between first and second session.

N	Measure	Font	Flanking direction	Radial eccentricity (deg)	Cardinal meridians	Thresholds per observer (measured twice)	Gaze tracking
50	Crowding	Sloan	radial	5, 10	all	8	Yes
	Crowding	Sloan	tangential	5	right, left	2	Yes
	Crowding	Pelli	radial	5	right, left	2	Yes
	Crowding	Pelli	horizontal	0	–	2	No
	Acuity	Sloan	–	0	–	2	No

Table 1. Data summary. In the periphery, we measured crowding distance radially and tangentially with Sloan font. With Pelli font we measured crowding in the periphery and fovea. We also measured foveal acuity with the Sloan font. Each threshold was measured once in two sessions separated by at least 24 hours. For peripheral thresholds we used gaze tracking. For foveal we instructed participants to fixate in the center since long viewing distance did not allow for correct gaze tracking.

Apparatus

Each testing session was completed on an Apple iMac 27" with an external monitor for stimulus presentation. Screen resolutions was 5120 x 2880. Apple iMac has an AMD graphics for optimal compatibility with the Psychtoolbox imaging software. The Systems Preference: Displays: Brightness slider was set (by calling MacDisplaySettings.m in the Psychtoolbox) to 0.86 (range 0 to 1) to achieve a white background luminance of about 250 cd/m^2. The observer viewed the screen binocularly at one of two different viewing distances. The software required a special keypress by the experimenter at the beginning of every block with a new observer or a new viewing distance, to affirm that the new viewing distance (eye to screen) was correct as measured with a tape measure, and that the screen center was orthogonal to the observer's line of sight. To measure crowding and acuity in the fovea, the viewing distance was 200 cm. For ±5 deg and ±10 deg eccentricity the distance was 40 cm, and for ±20 deg and ±30 deg was 20 cm. The long viewing distance gives good rendering of small stimuli; the short viewing distance results in a wide angular subtense of the display, to allow presentation of peripheral targets on either side of a central fixation. Stimuli were rendered using *CriticalSpacing.m* software (Pelli et al., 2016) implemented in MATLAB 2021 using the Psychtoolbox (Brainard, 1997). The smallest size of letter was always limited to be at least 8 pixels wide.

Stimuli and procedure

To measure acuity, we show one letter. To measure crowding we show a trigram of three letters or digits. For each trial, the three letters or digits are drawn randomly, without replacement, from the 9 letters or 9 digits available. Letters and digits are rendered as black in the Sloan or Pelli font, presented on a uniform white background of about 250 cd/m^2 (Pelli et al., 2016; Sloan et al., 1952). For the Sloan font we used the letters "DHKNORSVZ" as possible targets and flankers, and for the Pelli font we used the digits "123456789". We omit the C in the Sloan font because it's too easily confused with the O (Elliott et al., 1990). For crowding, each trigram was arranged either radially or tangentially. Each testing session included several blocks and was about an hour long. Most blocks measured 4 thresholds, interleaved, usually four crowding thresholds at 4 cardinal meridians. For Pelli font and tangential crowding we measured 2 thresholds along the horizontal meridian. To minimize the temptation to look away from fixation toward an expected target location, we randomly interleaved conditions measuring thresholds at all 4 cardinal meridian locations or 2 cardinal meridian locations with respect to fixation. A sample stimulus sequence appears in **Figure 4A**. Fixation is displayed until the observer presses a key to initiate the trial. Then the letter trigram appears on the screen for 150 ms and the computer waits for the observer to identify the middle letter by choosing the letter using a mouse click. Observers are instructed to return their eyes to fixation before clicking their response. The computer waits indefinitely for each response. The next trial began 300 ms after the observer's response, only if the observer was fixating in the center for at least 250 ms.

Measuring threshold for crowding distance

To estimate the crowding distance we used the procedure of Pelli et al., (Pelli et al., 2016). Letter spacing is controlled by QUEST. Letter size scales with spacing, maintaining a fixed ratio of 1.4:1. Based on fitting of psychometric functions for radial crowding by two

observers, we estimate the steepness parameter at 2.3, and we provided that value to QUEST for all the data reported here. We set the guessing rate parameter gamma to the reciprocal of the characters in the test alphabet for that font, usually 9. We set the "finger error" probability delta to 1%, so that QUEST can better cope with an occasional reporting mistake. At the end of each block, QUEST estimates threshold (crowding distance in deg, **Figure 4B**). To measure acuity, we followed a similar procedure, except that the target was presented without flankers. Threshold was defined as the letter spacing (crowding distance) or letter size (acuity) that achieved 70% correct identification, using Quest to control the stimulus parameters trial by trial, and make the final estimate. We used the Sloan and Pelli fonts to measure crowding distance. Acuity was always measured with the Sloan font.

We devote 35 trials to measure the threshold for a condition. We call that sequence a block, in which we randomly interleave meridian locations, to get a threshold for each condition, while keeping the observer uncertain as to which condition is present on each trial. We do this to minimize the urge to "peek" away from fixation. Until the target (and flankers on a crowding trial) are presented, there is nothing about the display that distinguishes which of the interleaved conditions this trial belongs to. (In principle, an observer could count the trials of each condition over the block and thus predict the condition of the last trial, since there are 35 trials of each condition in the block, but we do not explain the fixed number of trials per condition, we do not think they count, and this would affect only the last few trials.

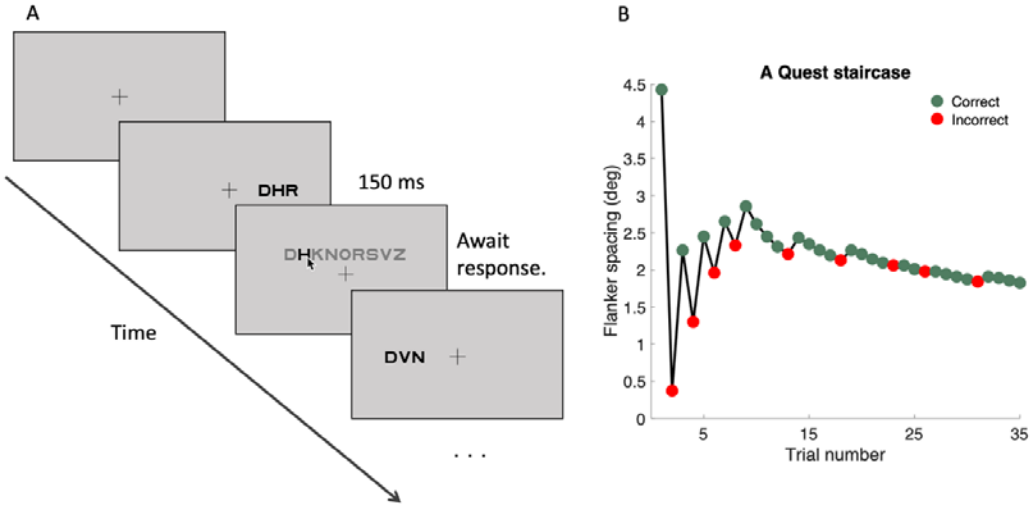

Figure 4. Stimulus and procedure. A) The image sequence for a peripheral trial. While gazing at the fixation mark, which is always present, the observer presses a key, which begins the first trial. The target presentation is accompanied by a low-pitched purr. Then the observer identifies the target by mouse click by highlighting and clicking on one letter out of all possible letters that appear above the fixation, hears a brief beep of reward if correct and silence otherwise, followed by the next target presentation. The actual target sequence is random (one out of 4 meridians), so the observer does not know which condition comes next. **B)** The staircase sequence of spacings tested on 35 successive trials of one condition (5 deg eccentricity), under control of QUEST. QUEST picks the most informative spacing to test on each trial to minimize variance of its final estimate. Finally, after 35 trials, QUEST estimates the crowding distance, i.e., spacing to achieve 70% correct. Notice that testing quickly homes in on threshold.

Eye tracking

For all peripheral thresholds we controlled for eye-movements. To measure fixation, we used an Eyelink 1000 eye tracker (SR Research, Ottawa, Ontario, Canada) with a sampling rate of 1000 Hz. To allow for short viewing distance (40 cm) we used their Tower mount setup with a 25 mm lens mounted on the Eyelink camera. We allowed participants to move their eyes withing 1.5 deg radius from fixation. We checked eye position during stimulus presentation. If during the stimulus presentation, observers moved their eyes more that 1.5 deg away from fixation, fixation cross turned red, and observers were asked to repeat the trial. Each threshold was estimated based on 35 trials with correct fixation. To start the trial observers had to fixate for at least 250ms. We did not use eye tracking with foveal measurements due to long viewing distance (200 cm).

Model fitting

Each fit minimizes the RMS error in fitting the log crowding distance S = log10(s), where s is the crowding distance in deg. Because the models are not linear functions of S, the fitting is nonlinear, for which we use MATLAB's fmincon function. It is generally found that the error of measurement of crowding distance is approximately proportional to the crowding distance. Because of the proportionality, the SD of log crowding distance is conserved across eccentricities, therefore it is recommended to fit the log of crowding distance. We estimate the participant, meridional, and font factors by solving several models (for equations see Table 3)**.**

We calculate the variance explained by each model using **Equation 1**.

$$\mathrm{R} = 1 - \sum_i \left(\vec{S}_i - \vec{S}_i'\right)^2 / \sum_i \left(\vec{S}_i - <\vec{S}>\right)^2 \tag{1}$$

where $\vec{S}$ is the vector of log crowding distances, $\vec{S}$ and $<\vec{S}>$ is mean log crowding distance, a scalar. Our fitting minimizes the RMS error in predicting the log crowding distances.

Cross validation

First, we divide data into randomly partitioned 6 subsets equal in size. With every cross-validation step one subset of data is retained as the validation set for testing the model, while the remaining subsets are used as training data. Each subset is chosen only once for testing. We repeat leave one out testing 6 times to obtain the full dataset. Variance explained R^2 is calculated by **Equation 1.**

Model comparison

We compare models using an F-test. First, we calculate sum of square errors (**Equation 2**) for two models that we want to compare. We refer to them as SSE_s (simple) for model with less parameters and SSE_f (full) for model with more parameters.

$$SSE = \sum_i \left(\vec{S}_i - \vec{S}_i'\right)^2 \tag{2}$$

We then calculate the F-statistic using **Equation 3**.

$$F = \frac{(SSE_s - SSE_f)/(p_f - p_s)}{SSE_f/(N - p_f)} \quad (3)$$

Where p_f is the number of parameters in the full model, p_s is number of parameters in the simple model and N is the number of observations. We estimate the p-value using F-distribution. P-value lower than 0.05 indicates that model with more parameters is significantly better in explaining the data.

Estimating the correction factor to compare crowding distance across studies

We define the *universal* Bouma factor *b*' as a measured ratio of crowding distance to radial eccentricity with a correction for differences to Bouma's original experiment: threshold criterion, number of choices, and linear vs. log spacing.

Table 2 compares Bouma factor *b* across studies that used various numbers of response choices (e.g., 9 Sloan letters vs. 2 orientations of a tumbling T), various threshold criteria (e.g., 70% vs. 75% correct) and linear or log flanker spacing. Including the present one, we know of five studies that compare crowding distance across meridians. We include Bouma 1970 as reference.

Study	**Target**	**Choices**	**Prob. Of guessing** γ	**Proportion correct at threshold** ***P***	**After correction for guessing** ***P****	**ΔB re Bouma 1970**	***b* ratio w.r.t. to Bouma 1970 (antilog of ΔB)**	**Correction factor (1/*b* ratio)**	**Correction factor for Log vs. linear flanker spacing**
Our data	Sloan or Pelli letter	9	1/9	0.70	0.66	-0.01	0.85	1.17	1.18
Greenwood et al., 2017	Tumbling clock	4	1/4	0.80	0.73	0.02	0.92	1.10	1
Toet & Levi, 1992	Tumbling T	2	1/2	0.75	0.50	-0.01	0.70	1.43	1
Grainger et al., 2010	Courier New letter or symbol	9	1/9	0.75	0.72	0.01	0.91	1.10	1
Coates et al., 2021	Tumbling T	4	1/4	0.625	0.50	-0.09	0.70	1.43	1
Bouma 1970	Courier 10 letter	25	1/25	0.75	0.79	0	1	1	1

Table 2. Factors for correction of the Bouma factor. The number of choices (e.g., the 9 Sloan letters or the 2 orientations of the Tumbling T) and the threshold criterion come from each study. We illustrate the

corresponding threshold criterion P^* to be applied after correction for guessing as in Equation 3. The relative log threshold ΔB is illustrated in Fig. 3A. The b ratio is $10^{\Delta B}$. We align the results of each of the studies to match those of Bouma (Bouma, 1970). We calculated Bouma's estimate at 75% threshold criterion based on Bouma's data reanalyzed by (Coates et al., 2021). For precise estimation of Bouma's estimate see *Bouma factor* paragraph in the discussion. Correction factor to account for linear vs. log flanker spacing is estimated in the Supplementary material.

To account for the different number of choices and threshold criterion, we assume that, with accurate fixation, the proportion correct P is a Weibull function of the log spacing X,

$$P(X) = \gamma + (1-\gamma)[1 - exp\,(-10^{\beta(X-S)})] \tag{4}$$

with a threshold log crowding distance S, where γ is the guessing rate (the reciprocal of the number of target choices, which is 1/9=0.11 when we use 9 letters of the Sloan font), and β is a steepness parameter, e.g., we set to 2.3, based on fitting psychometric functions to hundreds of trials at several spacings by two experienced observers. **Figure 5A** shows psychometric functions for the four studies, taking the guessing rate as the reciprocal of number of choices, and using our own estimate of steepness β. The horizontal shift is arbitrary. To accommodate various guessing rates gamma, we correct for guessing,

$$P^* = (P - \gamma)/(1 - \gamma) \tag{5}$$

This is a popular transformation of psychometric data, usually justified by assuming that the guessing rate can be modelled as an independent process. That makes sense for a yes-no task, but not for an identification task. Here we proceed regardless, since this rough correction helps align the Toet and Levi results with those of later studies. Applying this transformation to each threshold criterion P gives us the corresponding criterion $P^*(X)$ to apply after correction for guessing (Table 2). Applying the transformation to the psychometric function we get

$$P^*(X) = 1 - \exp(-10^{\beta(X-S)}) \tag{6}$$

Figure 5B plots this function, now the same function for all studies, and, for each study, reads off the log threshold spacing S corresponding to the threshold criterion P^*. As reference we use results from the paper by Toet & Levi (1992). The horizontal scale has been shifted to put zero at the Toet & Levi threshold. For each study, the value of log threshold spacing on this graph indicates how much bigger the reported threshold is relative to what the reference study would report.

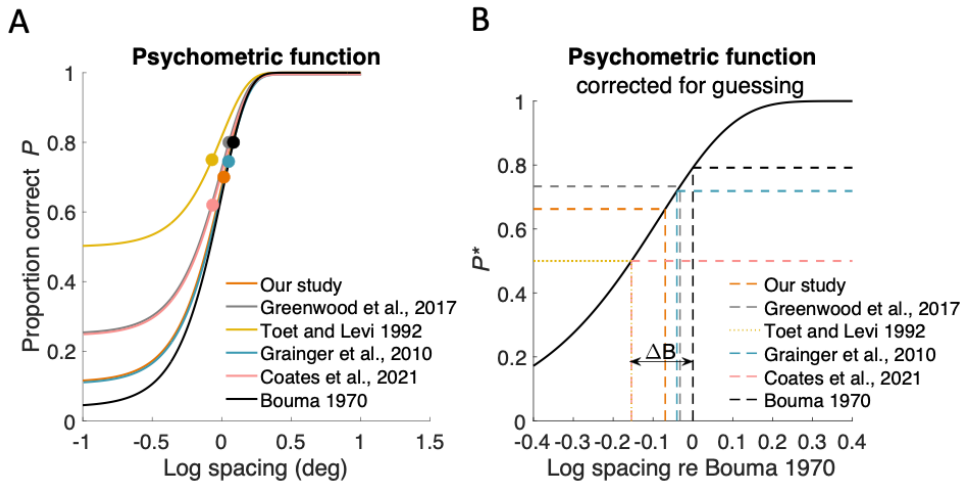

Figure 5. Effect of criterion on threshold log crowding distance. A) The six studies used various numbers of response choices and threshold criteria. **B)** Each horizontal dashed line represents the threshold criterion from one of the studies. Each studies threshold is indicated by vertical dashed line at the log spacing where the psychometric function intercepts the threshold criterion. Relative to the Bouma's parameters, threshold indicates how much each study's parameters increase the reported threshold. These numbers appear in the *ΔB* column of Table 2.

Since this analysis produced various correction factors, we shift the estimates of each study accordingly when comparing Bouma factor across studies.

Log-symmetric spacing of flankers

Since Bouma (Bouma, 1970), most crowding studies measure crowding distance as the minimum center-to-center spacing between the target and each of two flankers on opposite sides of the target to achieve a criterion level of performance. However, when crowding is measured in the radial direction, the Bouma law tells us that crowding distance increases linearly with eccentricity. Several studies have documented that, when flankers are arranged symmetrically about the target on a radial line from fixation, the outer flanker has much more effect. This is to be expected, since crowding distance grows with eccentricity and the outer flanker is more eccentric. In fact, crowding distance on the cortical surface—the product of crowding distance in deg and cortical magnification in mm/deg—is conserved across eccentricity (for eccentricities above 5 deg) because psychophysical crowding distance scales with eccentricities. Due to logarithmic cortical mapping of the visual field (Fischer, 1973), when measuring radial crowding we space the trigram so that the log eccentricity of the target is midway between the log eccentricities of the flankers and report the inner spacing. This raises the question of how to compare crowding distances between experiments that spaced the flankers linearly vs logarithmically. Given the Bouma law (**Equation 8** below) and the assumption that crowding distance depends primarily on the flanker-to-flanker distance, and negligibly on the target position between them, we show in the Supplement ("Effect of symmetric placement of flankers with regard to either linear or log eccentricity") that the crowding distance is expected to be 1.18 times larger when measured with linearly-spaced flankers than with log-spaced flankers. We account for this difference when comparing our results to other reports of the Bouma factor.

Does attention help explain differences in the reported Bouma factor.

Attention is an important factor in estimating psychometric thresholds (Carrasco, 2011). The attentional effects of crowding have been widely studied; however previous work yielded inconsistent results. On one hand, several found an attentional benefit in crowding tasks (Bacigalupo & Luck, 2015; Kewan-Khalayly et al., 2022) including reduction of crowding distance (Yeshurun & Rashal, 2010), while others failed to report these effects (Scolari et al., 2007; Strasburger, 2005; Strasburger & Malania, 2013). Here we did not use any attentional cues. Attention spreads between randomized locations at which we test, thus we do not expect different effects of attention across tested locations.

Obtaining data from other studies

We used WebPlotDigitizer (Rohatgi, 2020) to extract data of Toet and Levi (1992). Data from Grainger et al. (2010) was extracted from their Figure 7. Data from Greenwood et al. (2017) and Coates et al. (2021) was obtained through personal communication with the authors. Data from Bouma (1970) was estimated based on a recent reanalysis of Bouma' original data by Coates et al. (2021).

Statistical analysis

Statistics on the Bouma factors were assessed by an ANOVA with Bouma factor as dependent variable. Two sample comparisons are made with Wilcoxon rank sum test. For test-retest reliability and crowding correlations we report Pearson's r correlation coefficient.

RESULTS

Crowding and acuity data

As shown in the map of testing (**Figure 6**), we measured radial crowding thresholds in 50 adults at 9 locations of the visual field. With Sloan font we tested at four cardinal meridians at 5 and 10 deg eccentricity. On the left and right meridian at 5 deg eccentricity we also tested crowding with tangential flankers. With the Pelli font we tested crowding in the fovea and on the right and left meridians at 5 deg eccentricity. We also measured foveal acuity with the Sloan font. Each threshold was measured twice.

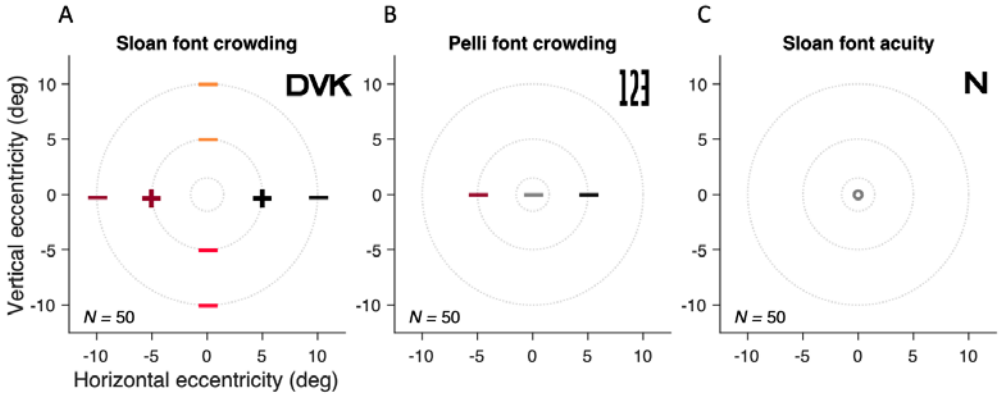

Figure 6. Map of testing. Each panel title indicates the measured threshold. Number in the lower left part of the plot represent the number of observers tested. The +, – and o symbols indicate testing of crowding with radial (—) or tangential (|) flankers, and testing of acuity (o). Typical stimuli appear in the upper right of each panel.

Test-retest reliability of visual threshold

To assess the reliability of our measurements, we measured each threshold twice on different days, at least 1 day apart with a maximum of 5 days apart. For crowding thresholds results are presented as the Bouma factor (see **Equation 7** below). Foveal crowding and acuity are presented as crowding distance (deg) and acuity as letter size (deg). **Figure 7** plots estimates from session one vs. estimates from session two for all tasks. All reported correlations between session 1 and session 2 are statistically significant ($p<0.001$). For the Pelli font (**Figure 7B**) we found a clear improvement of measured crowding distance in the second session (ratio of retest/test = 1.23). This effect was much smaller for the Sloan font (1.07). In general, we find very good reproducibility of measured thresholds, even though each threshold is derived only from one Quest staircase with 35 trials. The analyses performed in the following sections are based on the average threshold across sessions 1 and 2.

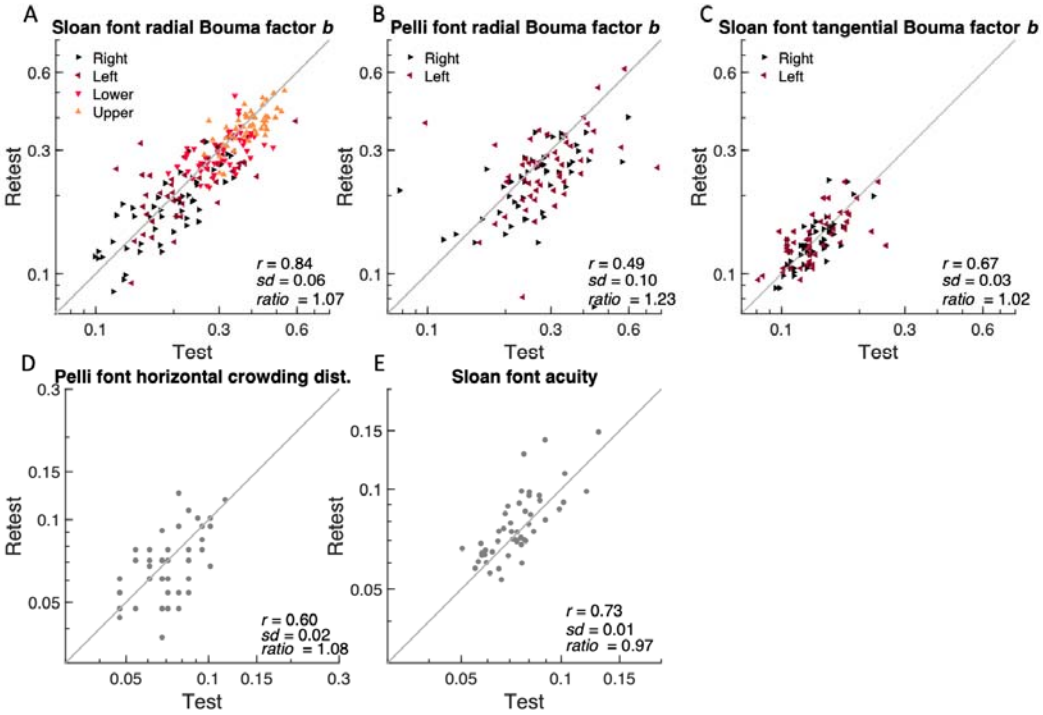

Figure 7. Test-retest reliability of acquired thresholds. We estimated Pearson's *r* correlation coefficient, standard deviation and ratio of test-retest based on the logarithm of the variable named at the top of each panel. For peripheral crowding (**A-C**) each measurement is represented by a triangle pointing towards the tested meridian. Black line is the equality line.

Bouma model explains 82% of variance

Bouma discovered the linear relationship between crowding distance and eccentricity (Bouma, 1970). He reported that the fitted slope (Bouma factor) *b* (**Equation 7**) between these two variables is equal to 0.5, suggesting that an observer can recognize a target at say

5 deg of eccentricity if flankers are spaced at least 2.5 deg away from the target. Bouma later reduced his estimate from 0.5 to 0.4 (Andriessen & Bouma, 1976).

The Bouma model (**Equation 7**), can be written as:

$$\hat{s} = (\varphi_0 + \varphi)\, b \tag{7}$$

where s is crowding distance in deg, φ is radial eccentricity in deg, and φ_0 and b are fitted constants. The slope b is the Bouma factor (Bouma, 1970; Rosen et al., 2014) and φ_0 is the fitted foveal offset (Liu & Arditi, 2000; Strasburger et al., 2011; Toet & Levi, 1992). Fitting the Bouma model to our large database of visual crowding (**Table 1**) results in an excellent fit (R^2 = 82.45%), showing that most of crowding's variation can be explained by eccentricity. Note that we used just two parameters, b and φ_0, to fit all 650 data points (13 thresholds measured in each of 50 observers). The estimated slope b was equal to 0.23 which is just over half of Bouma's 0.4. Although estimates lower than 0.3 have been reported before (Chung et al., 2001; Kooi et al., 1994; Strasburger et al., 1991; Toet & Levi, 1992), they were usually estimated for crowding measured in one location with one stimulus. Our database consists of measurements along four meridians with radial and tangential flankers. To capture the effect of these parameters on the Bouma factor we propose an extended version of the Bouma model.

Extended Bouma model explains 94% of variance

Crowding depends on more than just eccentricity. Crowding varies substantially across meridians (e.g., horizontal or vertical meridian), flanker direction (e.g., radial or tangential), target kind (e.g., letters vs. symbols) and across individuals. Here, we enhance the Bouma model by including these other variables. In our extension we consecutively add model parameters for meridian, flanking direction, target kind, and observer. All models and the variance that they account for are summarized in **Table 3**.

Equation 8 adds b_θ to the Bouma model which depends on meridian θ.

$$\hat{s} = \varphi_0 + \varphi b_\theta \tag{8}$$

where "b" now represents the geometric mean of b_θ, i.e., b = 10^(mean($\log_{10}(b_\theta)$)). Note that the meridian is undefined at the fovea.

Equation 9 adds an extra factor f_{dir} for flanking direction f (i.e., radial and tangential).

$$\hat{s} = (\varphi_0 b + \varphi b_\theta) f_{dir} \tag{9}$$

Equation 10 extends our model to accommodate dependence on the target kind (i.e., Sloan vs. Pelli font), by adding a factor t_{kind} for each font.

$$\hat{s} = (\varphi_0 b + \varphi b_\theta) f_{dir}\, t_{kind} \tag{10}$$

Finally, **Equation 11** adds the parameter o_i for each observer

$$\hat{s} = (\varphi_0 b + \varphi b_\theta) f_{dir} t_{kind} o_i \quad (11)$$

where $\prod_i \; o_i = 1$

Consecutively adding factors to the original model of Bouma, each extended model accounts for more variance. Comparing the simplest to the most enhanced model (**Equations 7 and 11**), we find that explained variance increases from 82% to 94% (**Table 3**). Model performance is improved by adding the meridian factor (R^2 = 89%) and flanking direction (R^2 = 93%) A negligible improvement of explained variance occurs after adding a factor for the target kind from 92.54% to 92.63%. Finally, our most enhanced model, with an observer factor, explains 94% of variance. The models are all cross validated, so the additional variance explained is not a necessary consequence of the increase in parameters. If the additional parameters were overfitting the training data, then we would find less variance accounted for in the left-out test data.

Model	**Equation**	***R^2(%)***	***R***	**RMSE**	**No. of parameters**
Bouma law	$s = (\varphi_0 + \varphi)\, b$ (7)	82.45	0.90	4.80	2
× **meridional factor**	$s = \varphi_0\, b + \varphi\, b_\theta$ (8)	88.96	0.94	3.80	5
× **flanking direction**	$s = (\varphi_0\, b + \varphi\, b_\theta)\, f_{dir.}$ (9)	92.54	0.96	3.13	7
× **target-kind factor**	$s = (\varphi_0\, b + \varphi\, b_\theta)\, f_{dir}\, t_{kind}$ (10)	92.63	0.96	3.11	9
× **observer factor**	$s = (\varphi_0\, b + \varphi\, b_\theta)\, f_{dir}\, t_{kind}\, o_i$ (11)	93.86	0.97	2.84	59

Table 3. How well our models predict crowding distance. We begin with the Bouma law (**Equation 7**). Successive models try to account for more variance by adding factors that depend on the meridian, flanking direction, font, and observer. Each row gives a significantly better fit than the row above (assessed with F-test using **Equation 3**). The R^2 (**Equation 1**) column shows cross-validated variance accounted for in predicting log crowding distance over the whole visual field (13 thresholds per observer). Pearson's *R* (the square root of R^2) shows the correlation between acquired and predicted data and RMSE is a root mean square mean error.

Since model parameters can be added in any order to form the final model (**Equation 11**), we asked how each of the additional parameters contributes to the increase of explained variance. To do so, we remove one parameter at the time, calculate the explained variance for the reduced model and compare it with the variance estimated for the full model. In this way, a drop in model performance reveals the contribution of an individual factor. We find that, after eccentricity, meridian contributes the most (4.5%) and target kind contributes the least (0.2%). Results are shown in **Table 4**.

Removed factor	Equation	Decrease in R^2(%)
meridian	$s\square = (\varphi_0 + \varphi)\, b\, f_{dir}\, t_{kind}\, o_i$	4.5%
flanking direction	$s\square = (\varphi_0\, b + \varphi\, b_\theta)\, t_{kind}\, o_i$	3.1%
observer	$s\square = (\varphi_0\, b + \varphi\, b_\theta)\, f_{dir}\, t_{kind}$	1.4%
target kind	$s\square = (\varphi_0\, b + \varphi\, b_\theta)\, f_{dir}\, o_i$	0.2%

Table 4 – Parameter contribution to the most extended Bouma law. We measured the contribution of each parameter to the full model.

Since the enhanced model accounts for more variance than the original Bouma law, we looked in the data for systematic effects of these parameters. Except for the target kind each of the proposed factors accounts for approximately a two-fold variation in the Bouma factor (**Figure 8**). We explore these effects further in the following sections.

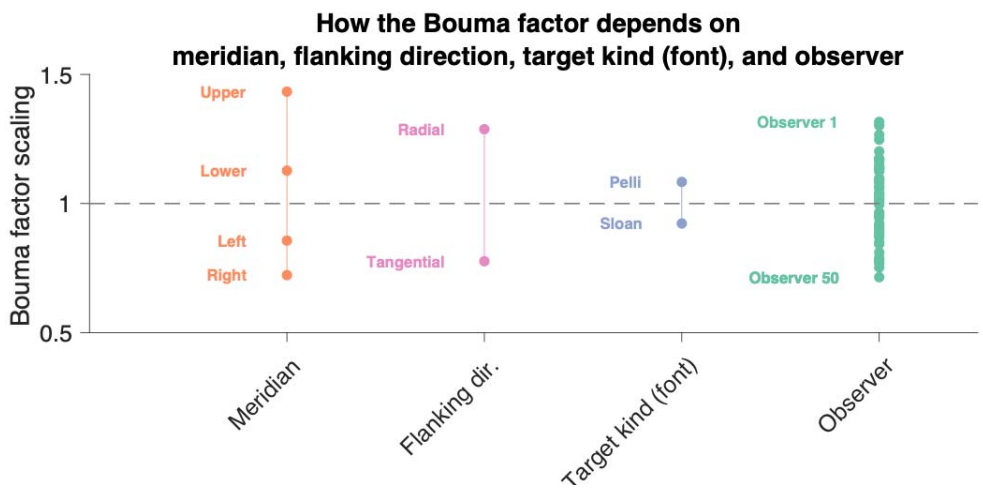

Figure 8. How several parameters scale the Bouma factor. To reveal the effect of each parameter, (horizontal axis), we refit the full model (**Equation 12**), but forcing that parameter to have a geometric mean of 1. The vertical axis plots model's estimates of that parameter. Because the model is multiplicative, the final Bouma factor is proportional to the product of all the parameters. We find a similar variation of Bouma factor with meridian, flanking direction, and observer. Target kind has the least effect.

Bouma factor varies twofold across meridians

Most of the thresholds in our dataset are for the Sloan font and with radial flankers. Using these data, we explore the variation of Bouma factor across meridians. We approximate the Bouma factor by fitting **Equation 7** for each subject and meridian independently (**Figure 9A**). The Bouma factor is smallest along the right meridian and highest along the upper meridian

(**Figure 9A**). Bouma factor is 0.19 right, 0.24 left, 0.30 lower and 0.39 upper meridian (**Figure 9A**; μ = 0.28).

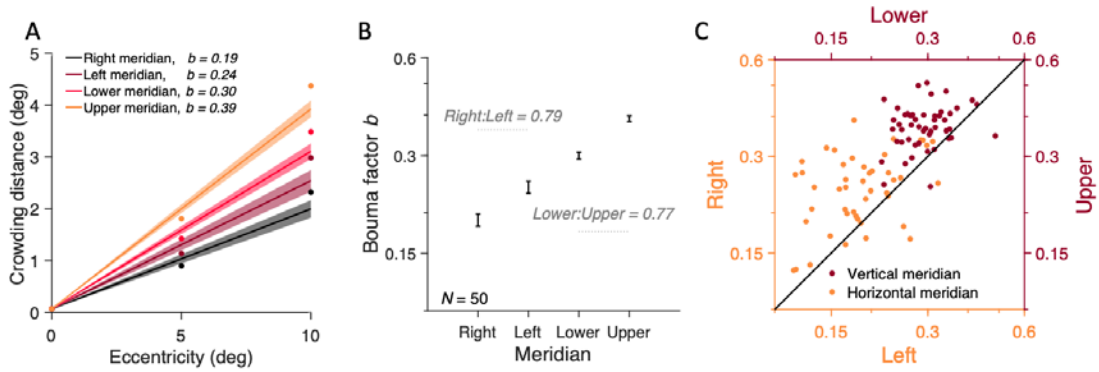

Figure 9. Bouma factor vs. meridian. A) Bouma factor estimates for radial crowding with the Sloan font. Each point represents the mean and error bars represent standard error across subject. **B)** Bouma factor vs. meridian. **C)** Individual-subject data plotted for the vertical (dark red) and horizontal meridian (orange).

Therefore, we report three asymmetries. 1) Along the vertical meridian Bouma factor was 0.77 smaller in the lower vs. the upper visual field. 2) Along the horizontal meridian Bouma factor was 0.79 smaller in the right vs. the left visual field. 3) Finally, we calculated that Bouma factor was 0.62 smaller along the horizontal than the vertical meridian. All reported asymmetries were highly consistent across subjects (**Figure 9C**). ANOVA test revealed that there was a significant effect of meridian $F(3,196) = 92.76$, $p < 0.001$ and post-hoc analysis showed that Bouma estimates at each meridian were significantly different from each other (all $p < 0.001$; corrected for multiple comparisons). For each meridian we also estimated, the eccentricity at which the crowding distance reaches twice its foveal value. E_2 was 0.37 deg ± 0.02 for right, 0.29 deg ± 0.02 for left, 0.22 deg ± 0.01 for lower and 0.17 deg ± 0.01 for upper meridian (N=50).

Bouma factor halves with tangential flankers

Tangential Bouma factor was much smaller than radial: 0.13 for the right and 0.14 for the left meridian. The tangential:radial ratio was 0.63 (ratio taken at each meridian and averaged). In contrast to radial crowding, we did not find any difference between left and right meridian as assessed with the Wilcoxon rank sum test ($z = -0.73$, $p = 0.49$).

Bouma factor varies with target kind

Although Pelli and Tillman (2008) suggested that crowding is invariant to stimulus kind, other work showed stimulus dependent differences (e.g. letters vs. symbols; Grainger et al., 2010). Along the horizontal meridian, Bouma factor for Sloan font was 0.19 on the right and 0.24 on the left, and slightly higher for Pelli font: 0.25 on the right and 0.29 for the left meridian. In general, Bouma factor for Pelli font was 1.26 larger than for Sloan (ratio taken at each meridian and averaged) and the difference between fonts was statistically significant ($z = 3.58$, $p < 0.001$). this, the model performance is slightly improved by adding the target-kind factor (**Equation 10**). This factor contributed little to the overall variance explained of the model because most of the data came from trials with the same target kind (Sloan letters). So even though excluding target kind as a factor from the model caused

inaccurate predictions for the Pelli font, the reduction in variance explained is negligible. We anticipate that the target-kind factor will account for more variance in datasets with more diverse stimuli. For the Pelli font, the right:left asymmetry was not statistically significant (z = -1.92, p = 0.06).

Bouma factor varies twofold across observers

Bouma factor varies with meridian, flanking direction, and target kind. Here, we quantify differences between observers. First, we estimated how well Bouma law fits individual-subject data. Fitting **Equation 7** to the right meridian data for each individual results, on average, in 97% explained variance, confirming that individual crowding data are well described by the linear model. Next, for each observer we fit the whole model to estimate the observer's overall Bouma factor (**Figure 10A-C**). Radial Bouma factor for the Sloan font varies approximately two-fold across observers (**Figure 10A**). This variation increases for tangential flankers (**Figure 10B**) and almost doubles for the Pelli font (**Figure 10C**). Individual differences are expressed as standard deviation of log Bouma factor (std b). Foveal acuity a and crowding s also vary two-fold (**Figure 10D-E**). The E_2 values also vary two-fold and range between 0.17 and 0.43.

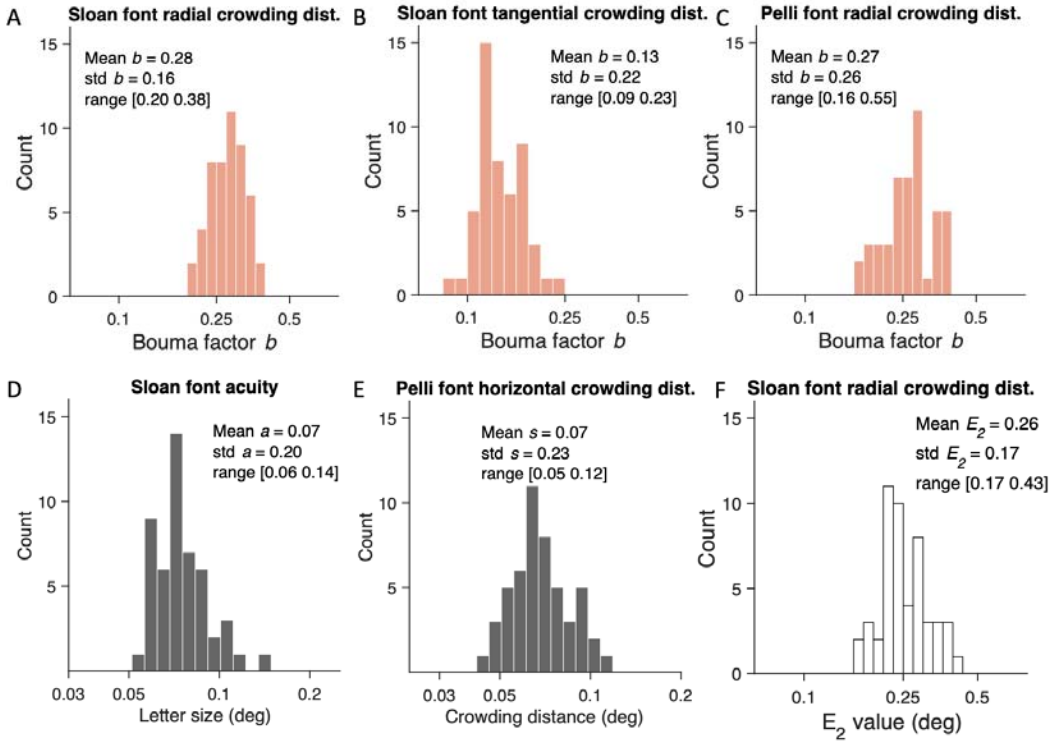

Figure 10. Histograms of crowding and acuity. Histograms of log Bouma factor for **A)** radial Sloan, **B)** radial Pelli **C)** tangential Sloan. **D)** Log acuity a (letter size), **E)** horizontal crowding distance s for Pelli. **F)** Histogram of E_2 values for Sloan font.

Bouma factor increases with eccentricity

Bouma reported the linear increase of crowding distance with eccentricity. We have seen that this linear equation fits our data well. However, seeing that we have 50 participants a

reviewer suggested that we examine how well Bouma factor is conserved across eccentricity. To estimate the Bouma factor we fit **Equation 7** for Sloan font with radial flankers to our data at 0 deg plus either 5 or 10 deg (**Figure 11A**). On average Bouma factor is 1.4 higher at 10 than 5 deg eccentricity. This effect was statistically significant ($F(1,398) = 42.3$ $p < 0.001$). This shows that the growth of crowding distance with eccentricity is not quite linear. Motivated by this finding we also measured crowding distance in 10 observers (already included in the main dataset) at 20 and 30 deg eccentricity. We modified **Equation 7** and added an exponent that allows the Bouma factor to grow with eccentricity (**Equation 12**). Comparing how well **Equations 7 and 12** fit data of 10 participants we find that replacing **Equation 7** by **Equation 12** increases the explained variance from 90% to 95%. The exponent parameter *n* was equal to 1.55.

$$\hat{s} = (\varphi_0 + \varphi)^n b \tag{12}$$

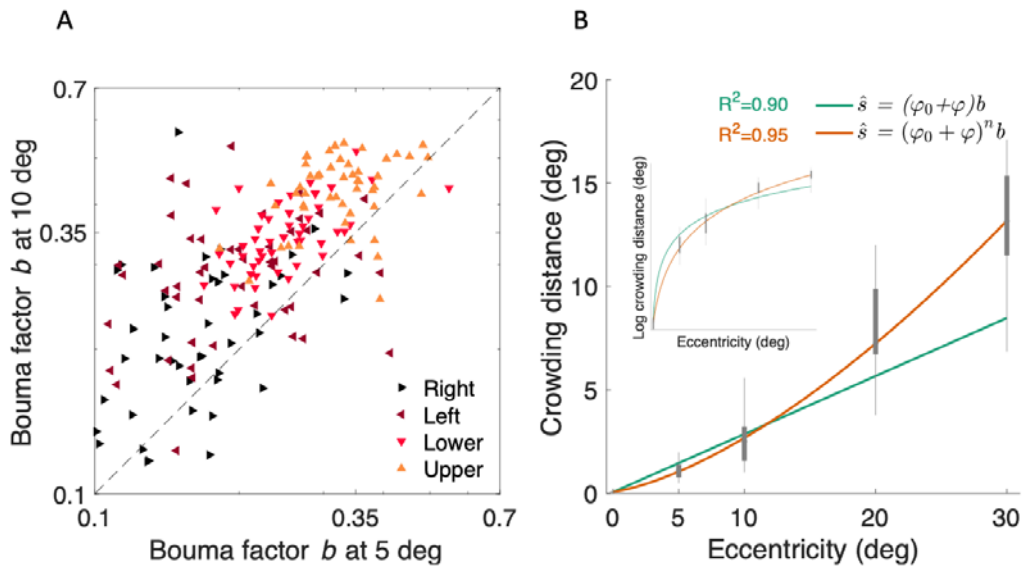

Figure 11. Bouma factors grows with eccentricity. A) Bouma factor at 5 vs. 10 deg eccentricity for radial flankers and Sloan font. Each color shows data at different meridian. Data is plotted for all 50 subjects included in the main study. **B)** Crowding distance for 10 participants between 0 and 30 deg eccentricity. Linear Bouma law (green), non-linear Bouma law (orange) fits to the data. Adding an exponent to Bouma law increases the explained variance from 90 to 95%. The inset shows the same fits on a log scale because the fit minimizes the error in log coordinates. Don't worry, this is not an artifact of perspective transformation: The computation of target angular size and eccentricity was done correctly using the arc tangent function.

Correlation of log crowding distance across visual field, flanking direction, and target kind

We explored the pattern of correlations of log crowding distance for visual field locations, flanking directions, and target kinds. These correlations are shown in **Figure 12A** where each cell shows Pearson's *r* between two measurements. Rows are sorted so that the average correlation decreases from top to bottom. We find that crowding distance measured on the right meridian at 10 deg eccentricity with Sloan font and radial flankers yields the highest average correlation with other measurements. Foveal crowding measured with Pelli font yields the smallest average correlation. To summarize how these correlations depend on stimulus properties we estimated an average correlation across measurements when 1, 2 or 3 stimulus properties (eccentricity, meridian, target kind, flanking direction) are modified

(**Figure 12B**). The test-retest correlation 0.54 is plotted at 0 changes. The average correlation with 1 change drops to 0.30 and keeps dropping with more changes to 0.25 at 2 changes and 0.18 at 3 changes.

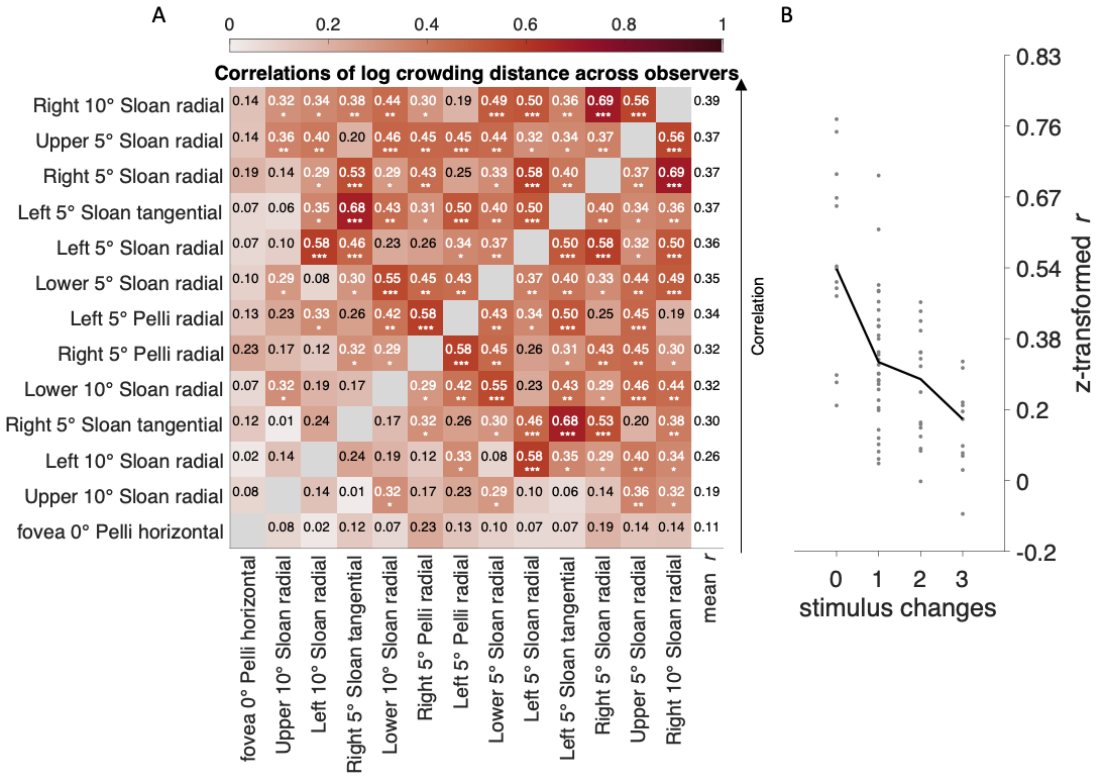

Figure 12. Correlations of log crowding distance. A) Pair-wise correlations between crowding distances (Table 1). Rows are sorted so that the average correlation decreases from top to bottom. **B)** Average correlation as a function of the number of stimulus property differences (radial eccentricity, meridian, flanking direction, and target kind) e.g., comparing two eccentricities is a 1-parameter difference.

Corrected Bouma factor and its asymmetries across different studies

We compare our estimates of the Bouma factor to previous studies. We included all studies that measured any of the crowding asymmetries. The eccentricities were all larger than 2 deg, so we estimate their Bouma factors by simply dividing the crowding distance by the radial eccentricity of the target. This calculation approximates the foveal intercept φ_0 as zero because it is easier to think about and has negligible consequence on the answer at large eccentricities (> 2 deg). Bouma factors were multiplied by a correction factor to account for criterion differences and log vs. linear flanker spacing (**Table 2**). In general, after the correction the average Bouma factor across studies increases (e.g., from 0.27 to 0.34 in the right meridian; **Table S2**). After correcting we obtain a new *universal* Bouma factor *b'* for each study. We compare results to those of others by plotting the universal Bouma factor vs. meridian. The studies are summarized below in **Table 5** and **Figure 13**.

Row	Study	N	Gaze tracking	Target	Flanking direction	mean universal Bouma factor (*b'*) [SEM]			
						Right	**Left**	**Lower**	**Upper**
1	Our data	50	Yes	Sloan font	radial	0.26 [0.24 0.27]	0.33 [0.31 0.34]	0.41 [0.40 0.42]	0.54 [0.52 0.54]
2	Our data	50	Yes	Sloan font	tangential	0.18 [0.17 0.20]	0.19 [0.17 0.29]	– –	– –
3	Our data	50	Yes	Pelli font	radial	0.34 [0.28 0.35]	0.40 [0.35 0.42]	– –	– –
4	Greenwood et al., 2017	12	Yes	tumbling clock	radial	0.34 [0.30 0.36]	0.37 [0.33 0.41]	0.51 [0.46 0.55]	0.69 [0.63 0.70]
5	Greenwood et al., 2017	12	Yes	tumbling clock	tangential	0.19 [0.17 0.20]	0.17 [0.14 0.19]	0.25 [0.22 0.27]	0.31 [0.28 0.36]
6	Toet & Levi, 1992	6	No	tumbling T	radial	– –	0.46 [0.40 0.50]	0.69 [0.66 0.70]	– –
7	Toet & Levi, 1992	6	No	tumbling T	tangential	– –	0.22 [0.21 0.24]	0.31 [0.29 0.34]	– –
8	Grainger et al., 2010	27	No	Courier New letter	radial	0.24 [0.19 0.29]	0.36 [0.29 0.40]	– –	– –
9	Grainger et al., 2010	27	No	Courier New symbol	radial	0.36 [0.30 0.40]	0.46 [0.42 0.51]	– –	– –
10	Coates et al., 2021	4	No	tumbling T	radial	0.41 [0.40 0.43]	– –	0.60 [0.58 0.61]	– –
B	Bouma 1970	25	No	Courier 10	radial	0.35			

Table 5 – Universal Bouma factor vs. meridian across studies. We compare our results with those of 4 studies that estimated crowding distance on more than one of the four cardinal meridians and report the Bouma factor by dividing crowding distance by the target's eccentricity. We corrected each all Bouma factors to account for differences in number of target choices and threshold criterion (see last column of **Table 2**). Data from Greenwood et al. (2017) and Coates et al. (2021) were obtained through personal communication. Data from Toet and Levi (1992) and Grainger et al. (2010) were digitized from figures in their papers. Data from Bouma 1970 were replotted from Coates et al. We estimated the standard error of Bouma factors in the Grainger et al. (2010) study based on the significance they reported for meridian differences.

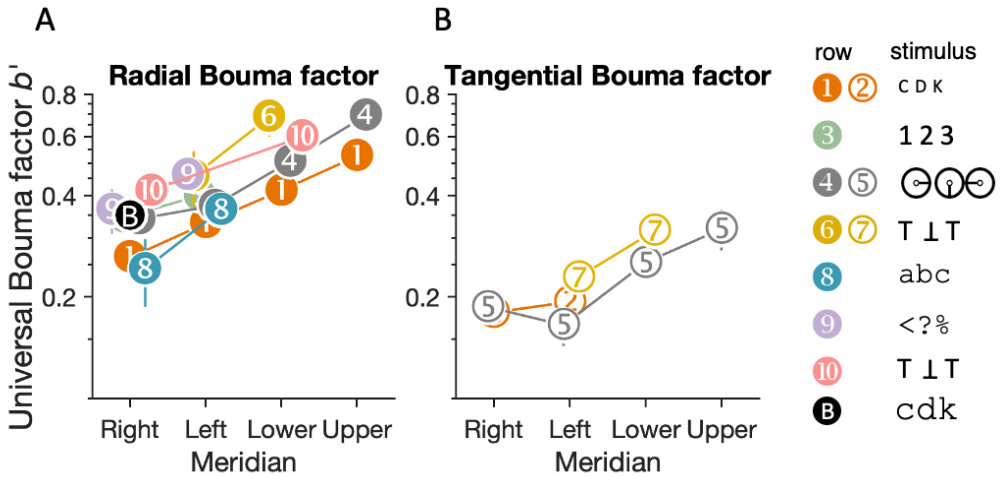

Figure 13. Universal Bouma factor vs. meridian for various studies and target kinds. Each numbered point corresponds to a numbered row of data in Table 5. **A)** Comparison across studies for radial flanking direction and **B)** tangential flanking direction. Both panels plot Bouma factor vs. meridian. Legend shows stimuli used in crowding experiments. B symbol shows Bouma factor reported by Bouma (1970) estimated at 4 deg eccentricity using threshold criterion of 75%. [See Bouma's data reanalysed in Coates et al. (2021).]

Only two rows in **Table 5 —** 1 (Sloan letter) and 4 (Tumbling clock) **—** report the radial universal Bouma factor for all four cardinal meridians. **Figure 13A** shows that although the universal Bouma factor is higher for the clock than for Sloan, by a factor of 1.3 (ratio of means for the right meridian 0.34/0.26 = 1.3), the two curves are otherwise similar, showing the same dependence on meridian. The 1.3:1 difference is not an artifact of number of choices or threshold criterion (**Table 2**). Both studies used gaze tracking to exclude fixation errors, so the difference is not a consequence of bad fixation. Thus, this seems to be a real 1.3:1 difference in universal Bouma factor between target kinds, precisely what the target-kind factor t_{kind} is meant to account for in **Equation 10**. The tumbling clocks may be more like each other than the 9 Sloan letters are and therefore produce larger crowding distance. Other studies (rows 3, 6, 9, 10) cluster above the Sloan font and show the same dependence on meridian. In general, we find that Courier new letters (row 8) produce the lowest universal Bouma factor (0.24 on the right meridian) and Tumbling T (row 10) produce the largest universal Bouma factor (0.41).

We also compared universal Bouma factors estimated with tangential flankers across studies (**Figure 13B**). Tangential Bouma factor did not vary as much as radial, and this is especially visible for the left meridian (**Figure 13B**, Rows 2,5,7). Radial crowding estimates, even with the same stimuli, showed more variation in the universal Bouma factor (**Figure 13A**, Rows 1, 4, 6). Although our data did not show any difference between right and left meridians (row 2), data extracted from Grainger et al. (2010) show a right:left asymmetry (row 5).

Table 6 and **Figure 14** report three Bouma-factor asymmetries (horizontal:vertical, right:left, and, lower:upper Bouma-factor ratios) for our and four selected studies. On average, the Bouma factor asymmetry is larger radially than tangentially. Radially, there is an advantage of horizontal over vertical meridian, right over left meridian, and lower over upper meridian in every study. The horizontal advantage seems to be insensitive to object kind as the estimates are clustered around ratio of 0.6-0.7. Similarly, lower visual field advantage is close to 0.8 for both studies that tested at this location. The right:left asymmetry is most variable. The right meridian advantage is largest for Courier new letters (row 8) and smallest for clocks (row 4) and the Pelli font (Row 3).

Row	Study	*N*	Gaze tracking	Target	Flanking direction	Ratio of Bouma factors [SEM]		
						Horizontal :vertical	Right:left	Lower:upper
1	Our data	50	Yes	Sloan font	radial	0.62 [0.60 0.64]	0.79 [0.78 0.91]	0.77 [0.76 0.82]
2	Our data	50	Yes	Sloan font	tangential	– –	0.95 [0.93 0.96]	– –

3	Our data	50	Yes	Pelli font	radial	– –	0.86 [0.82 0.90]	– –
4	Greenwood et al., 2017	12	Yes	tumbling clock	radial	0.59 [0.57 0.64]	0.92 [0.88 1.05]	0.74 [0.70 0.84]
5	Greenwood et al., 2017	12	Yes	tumbling clock	tangential	0.68 [0.6 0.75]	1.12 [0.97 1.40]	0.80 [0.73 0.91]
6	Toet & Levi, 1992	6	–	tumbling T	radial	0.67 [0.61 0.78]	– –	– –
7	Toet & Levi, 1992	6	–	tumbling T	tangential	0.70 [0.67 0.84]	– –	– –
8	Grainger et al., 2010	27	–	Courier New letter	radial	– –	0.67 [0.56 0.76]	– –
9	Grainger et al., 2010	27	–	Courier New symbol	radial	– –	0.78 [0.70 0.90]	– –
10	Coates et al., 2021	4	–	tumbling T	radial	0.68 [0.66 0.70]	– –	– –

Table 6 – Bouma factor and its asymmetries, compared with the literature. As in **Table 5**, we compare our results with those of studies that measured crowding distance on more than one of the four cardinal meridians. Each row lists the target type, flanking direction, and mean ratio with standard error across subjects.

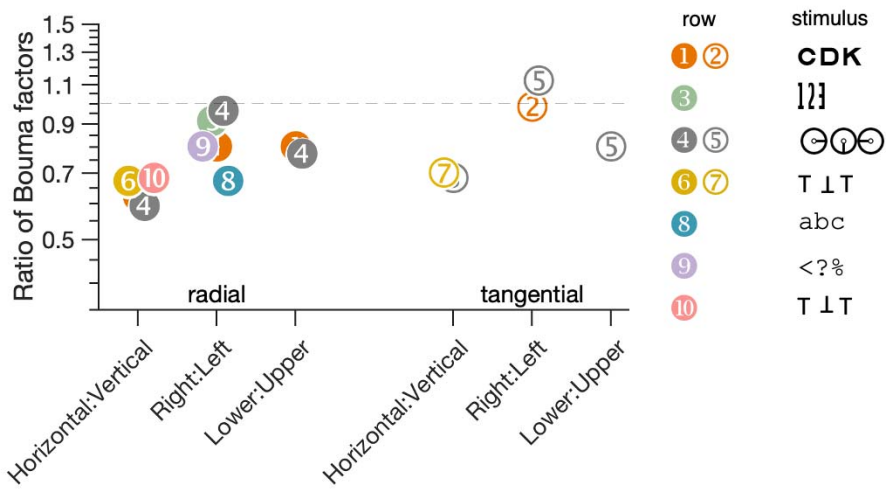

Figure 14. Plot of the three asymmetries, expressed as the ratio of Bouma factors. Each numbered point corresponds to a numbered row of data in Table 6. The horizontal dashed line at 1 represents no asymmetry. Each point is the ratio between Bouma factors; the error bars indicate the standard error across subjects. As in Figure 11, the legend shows examples of the objects that were used to measure crowding distance. All the error bars are small and hidden by the symbols.

What does peripheral crowding distance add to foveal acuity?

Any evaluation of the usefulness of crowding as a biomarker must assess what crowding tells us about the observer over and above what can be gleaned from foveal acuity. This is because all optometric and ophthalmic exams already include foveal acuity. For our 50 observers, foveal acuity failed to predict peripheral crowding, with an insignificant average correlation of 0.04 (**Figure 15;** red line). More generally both acuity and crowding

measured in the fovea fail to predict peripheral crowding (average foveal-peripheral crowding correlation is an insignificant correlation of 0.15). Thus, any exploration of the idea that peripheral crowding limits reading speed will need to measure peripheral crowding because it cannot be predicted from foveal measurements.

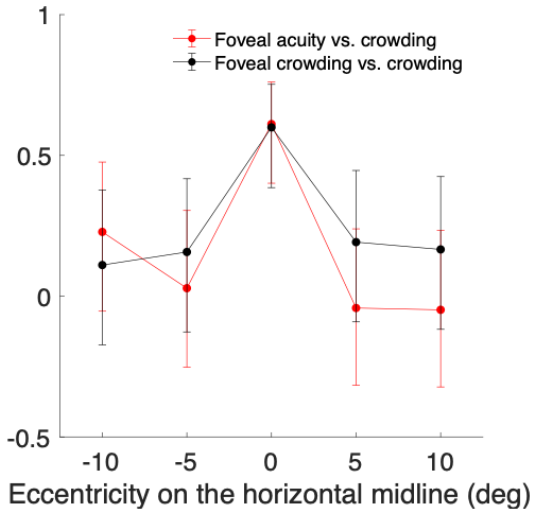

Figure 15. Foveal acuity and crowding fail to predict peripheral crowding. Correlation of foveal acuity and foveal crowding vs. peripheral crowding. For foveal crowding at 0 deg we plot a correlation for test-retest (**Figure 7D**).

DISCUSSION

The Bouma law and factor

Bouma law. The Bouma law describes the linear increase of crowding distance with eccentricity (Bouma, 1970, 1973; Levi, 2008; Pelli & Tillman, 2008; Rosen et al., 2014). The Bouma factor is the slope of that line (Rosen et al., 2014). The Bouma law is robust when fit to individual observer's data (Pelli et al., 2004; Rosen et al., 2014; Strasburger, 2020; Strasburger et al., 1991). In this study, for the first time, we fit Bouma law to data that includes measurements from 50 observers tested with two flanking directions at 9 locations of the visual field. Despite variation in crowding distance introduced by these factors Bouma law is an excellent fit to our data and explains 82.5% of the variance. We tried adding terms to the Bouma law to account for known factors: flanking direction (Greenwood et al., 2017; Kwon et al., 2014; Petrov & Meleshkevich, 2011; Toet & Levi, 1992), meridional location of the stimulus (Fortenbaugh et al., 2015; Greenwood et al., 2017; He et al., 1996), target-kind (Coates et al., 2021; Grainger et al., 2010) and individual differences (Petrov & Meleshkevich, 2011; Veríssimo et al., 2021). We find that the enhanced model explains a bit more variance (increased from 82.5% to 94%). Eccentricity remains the dominant factor, accounting for 82.5% of the variance.

Bouma factor. Bouma reported a "roughly" 0.5 slope for radial letter crowding vs. eccentricity (Bouma, 1970). Andriessen and Bouma (1976) later reported a slope of 0.4 for crowding of lines. Modern studies all use letters or symbols and typically report a Bouma factor of 0.3 (Chung et al., 2001; Coates et al., 2021; Kooi et al., 1994; Strasburger et al.,

1991; Toet & Levi, 1992), like our estimate of 0.28, averaged across observers and visual field locations (**Figure 10A**). In his original paper, Bouma said "For complete visual isolation of a letter presented at an eccentricity of φ, it follows that no other letters should be present within (roughly) 0.5φ distance."(Bouma, 1970). Note that Bouma says "roughly" and does not provide a quantitative threshold criterion, but Andriessen and Bouma later used a 75% threshold criterion for their estimation of Bouma factor. Coates et al. reanalyzed Bouma's original data with various threshold criteria and we can interpolate between the 70% and 80% thresholds to estimate the 75% threshold. Estimating the Bouma factor from Bouma's original data using this criterion yields a Bouma factor of 0.35, in line with modern estimates.

Reported Bouma factors range from 0.1 to 0.7 across studies (Pelli et al., 2004; Strasburger & Malania, 2013). Some of this variation can be explained by differences in target kind, number of possible targets, and threshold criterion. We define the *universal* Bouma factor b' as the reported Bouma factor b (ratio of crowding distance to radial eccentricity) multiplied by correction factors that account for differences in task and stimulus of the published study relative to Bouma's original paper. We correct for: number of alternatives, threshold criterion, and log vs. linear spacing of flankers.

Figure 13A shows that the corrected Bouma factor b' ranges from 0.24 for Courier new letters to 0.41 for Tumbling T. That residual difference may be due to target kind (Coates et al., 2021; Grainger et al., 2010). This is further supported by our finding that the Bouma factor was 1.23 higher for the Pelli font than for the Sloan font.

Linearity of the Bouma law. The linearity of the Bouma law implies that the Bouma factor is constant across eccentricities. The Coates et al. (2021) reanalysis of Bouma's data found that the Bouma factor increases slightly with eccentricity. They suggested that this inconsistency is likely because Bouma used constant size stimuli at all eccentricities. In 10 subjects we measured crowding distance between 0 and 30 deg eccentricity, as usual, scaling stimulus size with spacing (see Methods). Our results with proportional letter size replicate the trend in Bouma's data which were collected with fixed letter size. Enhancing the original model of Bouma for 10 observers to allow a nonlinear dependence on eccentricity improves the fit slightly, increasing the variance accounted from 90% to 95%. The Bouma factor doubles from 5 to 30 deg eccentricity. This effect is small and detectable in data from 0, 5 and 10 deg (**Figure 7A**), and becomes important at eccentricities of 20 and 30 deg (**Figure 7B**).

Crowding asymmetries

At any given eccentricity, the Bouma factor varies with polar angle. The Bouma factor is lower along the horizontal than vertical meridian (Greenwood et al., 2017; Petrov & Meleshkevich, 2011; Toet & Levi, 1992), is higher in the upper than lower meridian (Fortenbaugh et al., 2015; Greenwood et al., 2017; He et al., 1996; Toet & Levi, 1992), tends to be lower in the right than left meridian (Grainger et al., 2010; White et al., 2020) and approximately halves with tangential flankers (Greenwood et al., 2017; Kwon et al., 2014). In this work, we replicated all these asymmetries (**Figure 14**). The horizontal vs. vertical advantage and better performance in the lower vs. upper visual field is found for many

visual tasks (Abrams et al., 2012; Barbot et al., 2021; Carrasco et al., 2001; Chakravarthi et al., 2022; Himmelberg et al., 2020) and these asymmetries parallel those found in population receptive field size, cortical magnification, retinal ganglion cell density, and the BOLD signal magnitude (Benson et al., 2020; Himmelberg et al., 2021; Kupers et al., 2022; Kupers et al., 2019; Kurzawski et al., 2022; Kwon & Liu, 2019; Liu et al., 2006; Silva et al., 2018). The right:left asymmetry seems to be least described and does not generalize across all tasks. Beyond crowding, right visual field advantages have been reported: For native readers of left-to-right written languages, like English, the right meridian outperforms left in word recognition (Mishkin & Gorgays, 1952). Worrall and Coles (1976) examined letter recognition across the visual field and found a significant right hemifield advantage only along the right horizontal midline. The similarities in asymmetry suggest a common mechanism, and the differences may be useful hints toward the cortical substrate of crowding.

Standard deviation of measured acuity and crowding

To estimate the reliability of our measurements we acquired each threshold twice. Previous work showed improved performance in crowding tasks for repeated measurements (Chung, 2007; Malania et al., 2020). From their figures we estimate the second-block benefit to be 13% for Malania et al. and 20% for Chung. [Chung shows thresholds before and after 60 100-trial blocks. She shows percent correct for each block. By eye, we estimate that the benefit from first to second block is about a third that provided by the 60 blocks of training. Thus her 62% advantage (see average data from Table 1 – Chung 2007) after 60 blocks corresponds to the 20% advantage after the first block.] We find a modest second-threshold improvement for crowding thresholds measured with Sloan font at all tested locations and with the Pelli font in the fovea (less than 10%). Thresholds measured with Pelli font in the periphery yielded the highest improvement (23%). In our data the improvement is likely not due to acquiring familiarity with the task, since all observers participated in a training session, which consisted of repeated trials until 10 answers are correct. We find no improvement in acuity.

Overall, we find very good reproducibility of crowding and acuity thresholds (**Figure 7**). The standard deviation of log crowding distance measured with Sloan font for test-retest is much lower than the standard deviation across observers (0.06 vs 0.16). For acuity the standard deviation across test-retest was 0.01 vs 0.2 across observers.

Individual differences

Estimating individual differences requires data from many observers. In this paper we measured crowding in 50 observers, which is the biggest dataset of crowding measurements to date. Previous in-person crowding surveys included at most 27 observers (Grainger et al., 2010; Greenwood et al., 2017; Petrov & Meleshkevich, 2011; Toet & Levi, 1992). [An online crowding study tested 793 observers, but did not report individual differences (Liu et al., 2009).] To capture individual differences in the Bouma factor we included an observer factor o_i in the enhanced model. Adding the observer factor improved the explained variance from 92.6% to 94% (**Table 3**). Although this effect may seem negligible at first, we find that the Bouma factor varies twofold across observers, ranging from 0.20 to 0.38 (**Figure 10A**).

Bouma factor as a biomarker. Large individual differences enhance crowding's potential as a biomarker for studying cortical health and development. Specifically, crowding varies across children too (Kalpadakis-Smith et al., 2022) and predicts RSVP reading speed (Pelli et al., 2007). If crowding correlates with reading speed of beginning readers, then preliterate measures of crowding might help identify the children who need extra help before they learn to read. Measuring crowding distance across individuals in several diverse populations might expose any limit that crowding imposes on reading, yielding a norm for the development of crowding. Huge public interventions seek to help dyslexic children read faster and to identify them sooner. A virtue of crowding distance as a potential biomarker for dyslexia and cortical health, is that it can be measured in four minutes.

Crowding correlations

This paper reports 9 crowding thresholds for each of 50 observers. Such a comprehensive dataset allows for a correlation analysis to assess how well each crowding threshold predicts the others. We find a strong correlation of crowding between peripheral locations ($r = 0.39$) and hardly any between fovea and periphery ($r = 0.11$). We also find that crowding measured with radial flankers correlates highly with crowding measured with tangential flankers at the same location ($r = 0.52$ for the right meridian, $r = 0.48$ for the left meridian). In practice, data acquired on the right meridian with Sloan font and radial flankers at 10 deg eccentricity predicts all other peripheral thresholds with a correlation of (r = 0.41; correlation recalculated excluding the fovea), which is higher than for any other location. Thus, the most valuable single measurement to predict peripheral crowding is radial crowding at 10 deg along the right meridian.

The peeking-observer model

We always asked the observer to fixate on the crosshair during each trial. We acquired data with two methods: *unmonitored fixation*, without gaze tracking, and *awaited fixation,* in which the stimulus was only presented when the eye was near the fixation cross. Both methods are described in detail in the Methods section. The two methods yield different Bouma factor distributions. Upgrading from *unmonitored* to *awaited* fixation increased the Bouma factor mean by a factor of 1.5 (from 0.17 to 0.25) and halved the standard deviation (from 0.33 to 0.16). Histograms are shown in **Figure 1**. This model assumes firstly that performance on each trial depends solely on target eccentricity (relative to gaze position) and secondly that the observer peeks on a fraction p of the trials and fixates near the crosshair on the rest of the trials. When peeking, the observer fixates near one of the possible target locations. We assume that the target is too brief to allow an eye movement to the target during stimulus presentation. Thus, if there are n possible target locations, the peek will arrive at the correct target location with a probability of $1/n$.

In unmonitored fixation, the observer peeks with probability p. In awaited fixation, gaze-contingent display prevents peeking. Consider the condition with two possible target locations. The Bouma factor distribution is unimodal for low values of p and becomes bimodal for high values for p. *Unmonitored fixation* yields a bimodal Bouma distribution and is best fit with a peeking probability of 40% at target location and *awaited fixation* is unimodal and best fit by a peeking close to maximum allowed by the eye tracker. Replacing

the bimodal with the unimodal distribution raises the mean log Bouma factor, from -0.77 to -0.56 and halves the standard deviation from 0.33 to 0.19. Thus, eliminating peeking produces Bouma factor estimates consistent with Bouma's original data.

Why measure crowding?

Peripheral crowding provides additional information about visual health. Acuity is the threshold size of a target for recognition, while crowding is the threshold spacing. Clinical assessment routinely includes acuity and not crowding. Both limit recognition of everyday objects. Our results show that peripheral crowding is independent of foveal acuity and can be potentially used to provide additional insights about visual health. Specifically, peripheral crowding might predict dyslexia (Bouma & Legein, 1977; Martelli et al., 2009; O'Brien et al., 2005). There are hints that crowding tends to be worse in dyslexia (Pelli et al., 2007). If crowding correlates with reading speed of beginning readers, then preliterate measures of peripheral crowding might help identify the children who need extra help before they learn to read.

What about foveal crowding? In healthy individuals, foveal crowding correlates with foveal acuity but there are some conditions in which the two are dissociated. Strabismic amblyopia makes crowding worse in the fovea, but not in the periphery (Song et al., 2014). This suggests that fovea is the best place to detect increase in crowding associated with amblyopia. Traditional tests for crowding are mostly peripheral and use a fixation mark and a brief peripheral target, which are poorly suited for testing children and dementia patients whose attention may wander. Such participants will fixate much more reliably on a foveal target.

We hope there will be clinical studies to assess the diagnostic benefit of measuring crowding, which takes 3 minutes.

CONCLUSIONS

1. The well-known Bouma law — crowding distance depends linearly on radial eccentricity — explains 82% of the variance of log crowding distance, cross-validated. Our enhanced Bouma model, with factors for observer, meridian, and target kind, explains 94% of the variance, cross-validated.

2. Bouma factor varies twofold across observers, meridians, and flanking directions.

3. The meridional factors confirm previously reported asymmetries. Calculating asymmetry as the ratio of Bouma factors, we find a 0.62 horizontal:vertical advantage, a 0.77 lower:upper advantage, and a 0.79 right:left advantage.

4. For crowding individual differences are robust as evidenced by the nearly 3:1 ratio of 0.16 SD across observers to 0.06 SD of test-retest. For foveal acuity we have a 20:1 ratio of 0.20 SD across observers to 0.01 SD of test-retest.

5. Crowding distance measured at 10 deg of eccentricity along the right meridian is the best predictor of average crowding distance elsewhere (r = 0.39).

6. Preventing peeking, by use of gaze-contingent display, increased the mean of the log Bouma factor distribution from -0.77 (b = 0.17) to -0.56 (b = 0.28), and approximately halved its standard deviation (from 0.33 to 0.19).

7. Peripheral crowding is independent of foveal crowding and acuity.

ACKNOWLEDGEMENTS

We want to thank Nina Hanning for helping with Eyelink hardware and software and for providing code to extract eye position at each trial. We thank Barbara Finlay for helpful discussion of how psychophysical crowding correlations might relate to retinal developmental. We thank Ashley Feng and Laura Suciu for comments on the manuscript. The authors contributed as follows: Jan Kurzawski collected data with *awaited fixation*, analyzed data with *awaited* and *unmonitored fixation* and wrote the paper; Augustin Burchell performed initial analysis of *unmonitored fixation* data and participated in writing. Darshan Thapa supervised testing for data with *unmonitored fixation*; Najib Majaj participated in discussions throughout, including design, analysis, and writing; Jonathan Winawer participated in discussions throughout, including design, analysis, and writing; Denis Pelli directed the project, including design, software, data collection, analysis, and writing.

SUPPLEMENTARY MATERIALS

Effect of equating either linear or log spacing of the flankers

When measuring radial crowding, the target lies between two flankers on a radial line from fixation. Bouma spaced the flankers equally, and most investigators have followed suit. However, we spaced the flankers symmetrically on a logarithmic rather than linear scale. This raises the question of how to compare crowding distances between experiments that spaced the flankers linearly vs logarithmically. Given the Bouma law (**Equation 8**) and assuming that crowding distance depends primarily on the flanker-to-flanker distance, and negligibly on the target position between them, we show here that the crowding distance is expected to be 1.18 times larger when measured with linearly-spaced flankers than with log-spaced flankers.

Specifically, for a target at φ we choose the two flanker eccentricities φ_{in} and φ_{out} so that

$$\log(\varphi_0+\varphi) = [\log(\varphi_0+\varphi_{in}) + \log(\varphi_0+\varphi_{out})] / 2, \quad \text{(S1)}$$

where φ_0=0.15 deg, and report the spacing $s = \varphi-\varphi_{in}$. We can rearrange

$$\log(\varphi_0+\varphi_{out}) = 2 \log(\varphi_0+\varphi) - \log(\varphi_0+\varphi_{in}) \quad \text{(S2)}$$

Thus, the two flankers are at different (linear) distances from the target, but we suppose that they are equally effective in crowding the target. We report the center-to-center spacing from the inner flanker to the target as the "spacing" *s*.

We now estimate the relation of crowding distances measured in these two ways. We suppose that degree of crowding is determined by the separation between flankers on the log scale, as expected from the Bourma law:

$$\Delta\Phi = \log(\varphi_0+\varphi_{out}) - \log(\varphi_0+\varphi_{in}) \quad \text{(S3)}$$

If we are in the periphery, i.e. $\varphi_{in} > 1$, then φ_0=0.15 is negligible, and we can simplify,

$$\Delta\Phi \approx \log(\varphi_{out} / \varphi_{in}) \quad \text{(S4)}$$

Of course, increasing spacing alleviates crowding, so the degree of crowding will drop as log spacing $\Delta\Phi$ grows. Note that this model is at best an approximation, as it neglects position of the target. We are using it solely to compare crowding for two different ways of centering the target between flankers so the two target positions won't differ by much.

Linear flanker spacing: With flankers spaced symmetrically about the target on a linear scale, both at distance *s* from the target:

$$\Delta\Phi \approx \log (\varphi+s)/(\varphi-s) \quad \text{(S5)}$$

Log flanker spacing: With flankers spaced symmetrically on a log scale, the log flanker-to-flanker spacing is twice the log target-to-flanker spacing:

$$\Delta\Phi \approx 2 \log \varphi/(\varphi-s') \quad \text{(S6)}$$

Now we equate the two log flanker spacings, one with linearly symmetric spacing *s*, the other with log-symmetric spacing *s'*.

$$\log (\varphi+s)/(\varphi-s) = 2 \log \varphi/(\varphi-s') \quad \text{(S7)}$$

Solve for *s'*,

$$s' = \varphi - \varphi\,[(\varphi-s)/(\varphi+s)]^{0.5} \quad \text{(S8)}$$

Now substitute $b=s/\varphi$ and $b'=s'/\varphi$,

$$b' = 1 - [(1-b)/(1+b)]^{0.5} \quad \text{(S9)}$$

Figure S1 shows that, to a good approximation, this is a proportionality, with error of at most 0.014 over the relevant range $0 \le b \le 0.9$,

$$b' \approx 0.845\, b \quad \text{(S10)}$$

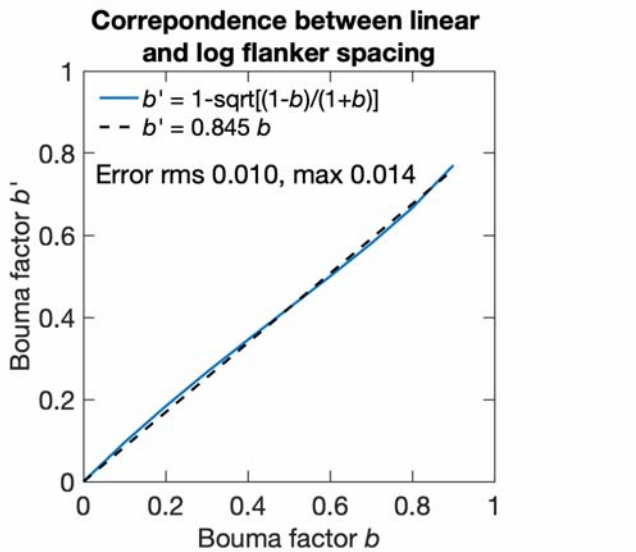

Figure S1. Log-symmetric spacing $s'=\varphi b'$ for same crowding effect as each linearly symmetric spacing $s = \varphi b$.

Thus, our log-symmetric spacing $s'=\varphi b'$ is approximately 0.845 times the linearly-symmetric spacing $s=\varphi b$ that is traditionally reported.

Log-log version of Bouma's model

Model	Equation	R^2(%)	Pearson's R	No. of parameters
Bouma law	$S \approx \Phi + B$	53.53	0.78	1
× meridional factor	$S \approx \Phi + B_\theta$	71.38	0.85	4
× flanking direction	$S \approx \Phi + B_\theta + F_d$	80.47	0.89	6
× target-kind factor	$S \approx \Phi + B_\theta + F_d + T_{kind}$	81.13	0.90	8
× observer factor	$S \approx \Phi + B_\theta + F_d + T_{kind} + O_i$	84.32	0.92	58

Table S1 – Fitting the log-log version of the Bouma law (excluding the foveal data). Uppercase variables are the log10 of corresponding lowercase variables. R_2 represents variance explained after model cross-validation. All our fitting minimizes error in log crowding distance so fitting log-log version can be fit by linear regression.

Correcting the Bouma factor

Row	Correction factor	Bouma factor / **Universal Bouma factor**			
		Right	**Left**	**Lower**	**Upper**
1	**x 1.38 (1.17x1.18)**	0.19 **0.26**	0.24 **0.33**	0.30 **0.41**	0.39 **0.54**
3	**x 1.38 (1.17x1.18)**	0.25 **0.35**	0.29 **0.40**		
4	**x 1.1**	0.31 **0.34**	0.34 **0.37**	0.46 **0.51**	0.63 **0.69**
8	**x 1.1**	0.22 **0.24**	0.33 **0.36**		
9	**x 1.1**	0.33 **0.36**	0.42 **0.46**		
10	**x 1.43**	0.32 **0.46**	– –	0.48 **0.69**	
Mean		0.27 **0.34**	0.32 **0.38**	0.41 **0.54**	0.51 **0.62**

Table S2 – Radial Bouma factor before and after the correction. Row numbers correspond to numbers in Table 5. Each cell contains two numbers. First is the Bouma factor before accounting for measurement differences and second is the Universal Bouma factor (recalculated after multiplying by the correction factor). Correction factors are calculated in Table 2. After the correction the average Bouma factor across studies increases (e.g., 0.27 to 0.34 on the right meridian).

Abrams, J., Nizam, A., & Carrasco, M. (2012). Isoeccentric locations are not equivalent: the extent of the vertical meridian asymmetry. *Vision Res*, *52*(1), 70-78. https://doi.org/10.1016/j.visres.2011.10.016

Aghdaee, S. M. (2005). Adaptation to spiral motion in crowding condition. *Perception*, *34*(2), 155-162. https://doi.org/10.1068/p5298

Andriessen, J. J., & Bouma, H. (1976). Eccentric vision: adverse interactions between line segments. *Vision Res*, *16*(1), 71-78. https://doi.org/10.1016/0042-6989(76)90078-x

Anstis, S. (1974). Letter: A chart demonstrating variations in acuity with retinal position. *Vision Research*, *14 7*, 589-592.

Atilgan, N., Yu, S. M., & He, S. (2020). Visual crowding effect in the parvocellular and magnocellular visual pathways. *J Vis*, *20*(8), 6. https://doi.org/10.1167/jov.20.8.6

Bacigalupo, F., & Luck, S. J. (2015). The allocation of attention and working memory in visual crowding. *J Cogn Neurosci*, *27*(6), 1180-1193. https://doi.org/10.1162/jocn_a_00771

Barbot, A., Xue, S., & Carrasco, M. (2021). Asymmetries in visual acuity around the visual field. *Journal of Vision*, *21*(1), 2-2. https://doi.org/10.1167/jov.21.1.2

Benson, N. C., Kupers, E. R., Barbot, A., Carrasco, M., & Winawer, J. (2020). Cortical Magnification in Human Visual Cortex Parallels Task Performance around the Visual Field. *bioRxiv*, 2020.2008.2026.268383. https://doi.org/10.1101/2020.08.26.268383

Bex, P. J., & Dakin, S. C. (2005). Spatial interference among moving targets. *Vision Res*, *45*(11), 1385-1398. https://doi.org/10.1016/j.visres.2004.12.001

Bex, P. J., Dakin, S. C., & Simmers, A. J. (2003). The shape and size of crowding for moving targets. *Vision Res*, *43*(27), 2895-2904. https://doi.org/10.1016/s0042-6989(03)00460-7

Bi, T., Cai, P., Zhou, T., & Fang, F. (2009). The effect of crowding on orientation-selective adaptation in human early visual cortex. *J Vis*, *9*(11), 13 11-10. https://doi.org/10.1167/9.11.13

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177-178. https://doi.org/10.1038/226177a0

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Res*, *13*(4), 767-782. https://doi.org/10.1016/0042-6989(73)90041-2

Bouma, H., & Legein, C. P. (1977). Foveal and parafoveal recognition of letters and words by dyslexics and by average readers. *Neuropsychologia*, *15*(1), 69-80. https://doi.org/10.1016/0028-3932(77)90116-6

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis*, *10*(4), 433-436. https://www.ncbi.nlm.nih.gov/pubmed/9176952

Burchell, A., Benson, N. C., Zhou, J. Y., Winawer, J. A., & Pelli, D. G. (2019). Using fMRI to link crowding to hV4. *Journal of Vision*, *19*(10), 14a-14a. https://doi.org/10.1167/19.10.14a

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Res*, *51*(13), 1484-1525. https://doi.org/10.1016/j.visres.2011.04.012

Carrasco, M., Talgar, C. P., & Cameron, E. L. (2001). Characterizing visual performance fields: effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spat Vis*, *15*(1), 61-75. https://doi.org/10.1163/15685680152692015

Chakravarthi, R., Papadaki, D., & Krajnik, J. (2022). Visual field asymmetries in numerosity processing. *Atten Percept Psychophys*, *84*(8), 2607-2622. https://doi.org/10.3758/s13414-022-02585-1

Chastain, G. (1982). Confusability and interference between members of parafoveal letter pairs. *Perception & Psychophysics*, *32*(6), 576-580. https://doi.org/10.3758/BF03204213

Chung, S. T. (2007). Learning to identify crowded letters: does it improve reading speed? *Vision Res*, *47*(25), 3150-3159. https://doi.org/10.1016/j.visres.2007.08.017

Chung, S. T. (2013). Cortical reorganization after long-term adaptation to retinal lesions in humans. *J Neurosci*, *33*(46), 18080-18086. https://doi.org/10.1523/JNEUROSCI.2764-13.2013

Chung, S. T., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Res*, *41*(14), 1833-1850. https://doi.org/10.1016/s0042-6989(01)00071-2

Coates, D. R., Levi, D. M., Touch, P., & Sabesan, R. (2018). Foveal Crowding Resolved. *Sci Rep*, *8*(1), 9177. https://doi.org/10.1038/s41598-018-27480-4

Coates, D. R., Ludowici, C. J. H., & Chung, S. T. L. (2021). The generality of the critical spacing for crowded optotypes: From Bouma to the 21st century. *J Vis*, *21*(11), 18. https://doi.org/10.1167/jov.21.11.18

Ehlers, H. (1936). V: THE MOVEMENTS OF THE EYES DURING READING [https://doi.org/10.1111/j.1755-3768.1936.tb07306.x]. *Acta Ophthalmologica*, *14*(1-2), 56-63. https://doi.org/https://doi.org/10.1111/j.1755-3768.1936.tb07306.x

Ehlers, H. (1953). Clinical testing of visual acuity. *AMA Arch Ophthalmol*, *49*(4), 431-434. https://doi.org/10.1001/archopht.1953.00920020441007

Elliott, D. B., Whitaker, D., & Bonette, L. (1990). Differences in the legibility of letters at contrast threshold using the Pelli-Robson chart [https://doi.org/10.1111/j.1475-1313.1990.tb00877.x]. *Ophthalmic and Physiological Optics*, *10*(4), 323-326. https://doi.org/https://doi.org/10.1111/j.1475-1313.1990.tb00877.x

Felisbert, F. M., Solomon, J. A., & Morgan, M. J. (2005). The role of target salience in crowding. *Perception*, *34*(7), 823-833. https://doi.org/10.1068/p5206

Fischer, B. (1973). Overlap of receptive field centers and representation of the visual field in the cat's optic tract. *Vision Res*, *13*(11), 2113-2120. https://doi.org/10.1016/0042-6989(73)90188-0

Flom, M. C., Heath, G. G., & Takahashi, E. (1963). Contour Interaction and Visual Resolution: Contralateral Effects. *Science*, *142*(3594), 979-980. https://doi.org/10.1126/science.142.3594.979

Fortenbaugh, F. C., Silver, M. A., & Robertson, L. C. (2015). Individual differences in visual field shape modulate the effects of attention on the lower visual field advantage in crowding. *J Vis*, *15*(2). https://doi.org/10.1167/15.2.19

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nat Neurosci*, *14*(9), 1195-1201. https://doi.org/10.1038/nn.2889

Grainger, J., Tydgat, I., & Issele, J. (2010). Crowding affects letters and symbols differently. *J Exp Psychol Hum Percept Perform*, *36*(3), 673-688. https://doi.org/10.1037/a0016888

Greenwood, J. A., Szinte, M., Sayim, B., & Cavanagh, P. (2017). Variations in crowding, saccadic precision, and spatial localization reveal the shared topology of spatial vision. *Proc Natl Acad Sci U S A*, *114*(17), E3573-E3582. https://doi.org/10.1073/pnas.1615504114

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*(6598), 334-337. https://doi.org/10.1038/383334a0

Herzog, M. H., Sayim, B., Chicherov, V., & Manassi, M. (2015). Crowding, grouping, and object recognition: A matter of appearance. *J Vis*, *15*(6), 5. https://doi.org/10.1167/15.6.5

Himmelberg, M. M., Kurzawski, J. W., Benson, N. C., Pelli, D. G., Carrasco, M., & Winawer, J. (2021). Cross-dataset reproducibility of population receptive field (pRF) estimates and cortical magnification asymmetries. *bioRxiv*, 2021.2004.2012.439348. https://doi.org/10.1101/2021.04.12.439348

Himmelberg, M. M., Winawer, J., & Carrasco, M. (2020). Stimulus-dependent contrast sensitivity asymmetries around the visual field. *J Vis*, *20*(9), 18. https://doi.org/10.1167/jov.20.9.18

Kalpadakis-Smith, A. V., Tailor, V. K., Dahlmann-Noor, A. H., & Greenwood, J. A. (2022). Crowding changes appearance systematically in peripheral, amblyopic, and developing vision. *J Vis*, *22*(6), 3. https://doi.org/10.1167/jov.22.6.3

Kewan-Khalayly, B., Migo, M., & Yashar, A. (2022). Transient attention equally reduces visual crowding in radial and tangential axes. *J Vis*, *22*(9), 3. https://doi.org/10.1167/jov.22.9.3

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spat Vis*, *8*(2), 255-279. https://doi.org/10.1163/156856894x00350

Korte, W. (1923). Über die Gestaltauffassung im indirekten Sehen. *Zeitschrift für Psychologie*, *93*, 17-82.

Kupers, E. R., Benson, N. C., Carrasco, M., & Winawer, J. (2022). Asymmetries around the visual field: From retina to cortex to behavior. *PLoS Comput Biol*, *18*(1), e1009771. https://doi.org/10.1371/journal.pcbi.1009771

Kupers, E. R., Carrasco, M., & Winawer, J. (2019). Modeling visual performance differences 'around' the visual field: A computational observer approach. *PLoS Comput Biol*, *15*(5), e1007063. https://doi.org/10.1371/journal.pcbi.1007063

Kurzawski, J. W., Gulban, O. F., Jamison, K., Winawer, J., & Kay, K. (2022). Non-neural factors influencing BOLD response magnitudes within individual subjects. *J Neurosci*, *42*(38), 7256-7266. https://doi.org/10.1523/JNEUROSCI.2532-21.2022

Kurzawski, J. W., Pelli, D. G., & Winawer, J. A. (2021). Conservation across individuals of cortical crowding distance in human V4. *Journal of Vision*, *21*(9), 2675-2675. https://doi.org/10.1167/jov.21.9.2675

Kwon, M., Bao, P., Millin, R., & Tjan, B. S. (2014). Radial-tangential anisotropy of crowding in the early visual areas. *J Neurophysiol*, *112*(10), 2413-2422. https://doi.org/10.1152/jn.00476.2014

Kwon, M., & Liu, R. (2019). Linkage between retinal ganglion cell density and the nonuniform spatial integration across the visual field. *Proc Natl Acad Sci U S A*, *116*(9), 3827-3836. https://doi.org/10.1073/pnas.1817076116

Latham, K., & Whitaker, D. (1996). Relative roles of resolution and spatial interference in foveal and peripheral vision. *Ophthalmic and Physiological Optics*, *16*(1), 49-57. https://doi.org/https://doi.org/10.1016/0275-5408(95)00124-7

Leat, S. J., Li, W., & Epp, K. (1999). Crowding in central and eccentric vision: the effects of contour interaction and attention. *Invest Ophthalmol Vis Sci*, *40*(2), 504-512. https://www.ncbi.nlm.nih.gov/pubmed/9950611

Levi, D. M. (2008). Crowding--an essential bottleneck for object recognition: a mini-review. *Vision Res*, *48*(5), 635-654. https://doi.org/10.1016/j.visres.2007.12.009

Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: why bigger is better. *Curr Biol*, *19*(23), 1988-1993. https://doi.org/10.1016/j.cub.2009.09.056

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Res*, *25*(7), 963-977. https://doi.org/10.1016/0042-6989(85)90207-x

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for Participants' Viewing Distance in Large-Scale, Psychophysical Online Experiments Using a Virtual Chinrest. *Sci Rep*, *10*(1), 904. https://doi.org/10.1038/s41598-019-57204-1

Liu, L., & Arditi, A. (2000). Apparent string shortening concomitant with letter crowding. *Vision Res*, *40*(9), 1059-1067. https://doi.org/10.1016/s0042-6989(99)00247-3

Liu, T., Heeger, D. J., & Carrasco, M. (2006). Neural correlates of the visual vertical meridian asymmetry. *J Vis*, *6*(11), 1294-1306. https://doi.org/10.1167/6.11.12

Liu, T., Jiang, Y., Sun, X., & He, S. (2009). Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. *Curr Biol*, *19*(2), 127-132. https://doi.org/10.1016/j.cub.2008.11.065

Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: selective interference between configural representations of faces in crowded scenes. *J Vis*, *7*(2), 24 21-11. https://doi.org/10.1167/7.2.24

Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *J Vis*, *7*(2), 1 1-7. https://doi.org/10.1167/7.2.1

Malania, M., Pawellek, M., Plank, T., & Greenlee, M. W. (2020). Training-Induced Changes in Radial-Tangential Anisotropy of Visual Crowding. *Transl Vis Sci Technol*, *9*(9), 25. https://doi.org/10.1167/tvst.9.9.25

Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *J Vis*, *9*(4), 14 11-18. https://doi.org/10.1167/9.4.14

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *J Vis*, *5*(1), 58-70. https://doi.org/10.1167/5.1.6

Millin, R., Arman, A. C., Chung, S. T., & Tjan, B. S. (2014). Visual crowding in V1. *Cereb Cortex*, *24*(12), 3107-3115. https://doi.org/10.1093/cercor/bht159

Mishkin, M., & Gorgays, D. G. (1952). Word recognition as a function of retinal locus. *J Exp Psychol*, *43*(1), 43-48. https://doi.org/10.1037/h0061361

Motter, B. C. (2006). Modulation of transient and sustained response components of V4 neurons by temporal crowding in flashed stimulus sequences. *J Neurosci*, *26*(38), 9683-9694. https://doi.org/10.1523/JNEUROSCI.5495-05.2006

Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Res*, *32*(4), 771-777. https://doi.org/10.1016/0042-6989(92)90192-l

O'Brien, B. A., Mansfield, J. S., & Legge, G. E. (2005). The effect of print size on reading speed in dyslexia. *J Res Read*, *28*(3), 332-349. https://doi.org/10.1111/j.1467-9817.2005.00273.x

Pelli, D., & Robson, J. (1991). Are letters better than gratings? *Clin Vis Sci*, *6*, 409-411.

Pelli, D., Waugh, S., Martelli, M., Crutch, S., Primativo, S., Yong, K., Rhodes, M., Yee, K., Wu, X., Famira, H., & Yiltiz, H. (2016). A clinical test for visual crowding. *F1000Research*, *5*. https://doi.org/10.12688/f1000research.7835.1

Pelli, D. G. (2008). Crowding: a cortical constraint on object recognition. *Curr Opin Neurobiol*, *18*(4), 445-451. https://doi.org/10.1016/j.conb.2008.09.008

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis*, *4*(12), 1136-1169. https://doi.org/10.1167/4.12.12

Pelli, D. G., Robson, J. G., & Wilkins, A. J. (1988). The design of a new letter chart for measuring contrast sensitivity [Article]. *Vision Research*, *2*(3), 187-199. http://www.scopus.com/inward/record.url?scp=0023948250&partnerID=8YFLogxK http://www.scopus.com/inward/citedby.url?scp=0023948250&partnerID=8YFLogxK

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nat Neurosci*, *11*(10), 1129-1135. https://doi.org/10.1038/nn.2187

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *J Vis*, *7*(2), 20 21-36. https://doi.org/10.1167/7.2.20

Petrov, Y., & Meleshkevich, O. (2011). Asymmetries and idiosyncratic hot spots in crowding. *Vision Res*, *51*(10), 1117-1123. https://doi.org/10.1016/j.visres.2011.03.001

Rohatgi, A. (2020). Webplotdigitizer: Version 4.4. In.

Rosen, S., Chakravarthi, R., & Pelli, D. (2014). The Bouma law of crowding, revised: Critical spacing is equal across parts, not objects. *Journal of Vision*, *14*. https://doi.org/10.1167/14.6.10

Rosenholtz, R. (2016). Capabilities and Limitations of Peripheral Vision. *Annu Rev Vis Sci*, *2*, 437-457. https://doi.org/10.1146/annurev-vision-082114-035733

Scolari, M., Kohnen, A., Barton, B., & Awh, E. (2007). Spatial attention, preview, and popout: which factors influence critical spacing in crowded displays? *J Vis*, *7*(2), 7 1-23. https://doi.org/10.1167/7.2.7

Siderov, J., Waugh, S. J., & Bedell, H. E. (2013). Foveal contour interaction for low contrast acuity targets. *Vision Res*, *77*, 10-13. https://doi.org/10.1016/j.visres.2012.11.008

Silva, M. F., Brascamp, J. W., Ferreira, S., Castelo-Branco, M., Dumoulin, S. O., & Harvey, B. M. (2018). Radial asymmetries in population receptive field size and cortical magnification factor in early visual cortex. *Neuroimage*, *167*, 41-52. https://doi.org/10.1016/j.neuroimage.2017.11.021

Sloan, L. L., Rowland, W. M., & Altman, A. (1952). Comparison of three types of test target for the measurement of visual acuity. *Q Rev Ophthalmol*, *8*(1), 4-16.

Song, S., Levi, D. M., & Pelli, D. G. (2014). A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *J Vis*, *14*(5), 3. https://doi.org/10.1167/14.5.3

Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *J Vis*, *5*(11), 1024-1037. https://doi.org/10.1167/5.11.8

Strasburger, H. (2020). Seven Myths on Crowding and Peripheral Vision. *Iperception*, *11*(3), 2041669520913052. https://doi.org/10.1177/2041669520913052

Strasburger, H., Harvey, L. O., Jr., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Percept Psychophys*, *49*(6), 495-508. https://doi.org/10.3758/bf03212183

Strasburger, H., & Malania, M. (2013). Source confusion is a major cause of crowding. *J Vis*, *13*(1), 24. https://doi.org/10.1167/13.1.24

Strasburger, H., Rentschler, I., & Juttner, M. (2011). Peripheral vision and pattern recognition: a review. *J Vis*, *11*(5), 13. https://doi.org/10.1167/11.5.13

Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty. Its relationship to visual acuity in normal and amblyopic eyes. *Am J Ophthalmol*, *53*, 471-477. https://www.ncbi.nlm.nih.gov/pubmed/13917936

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Res*, *32*(7), 1349-1357. https://doi.org/10.1016/0042-6989(92)90227-a

Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Res*, *42*(20), 2357-2369. https://doi.org/10.1016/s0042-6989(02)00197-9

Tyler, C. W., & Likova, L. T. (2007). Crowding: a neuroanalytic approach. *J Vis*, *7*(2), 16 11-19. https://doi.org/10.1167/7.2.16

Veríssimo, I. S., Hölsken, S., & Olivers, C. N. L. (2021). Individual differences in crowding predict visual search performance. *Journal of Vision*, *21*(5), 29-29. https://doi.org/10.1167/jov.21.5.29

Wallis, T. S., & Bex, P. J. (2012). Image correlates of crowding in natural scenes. *J Vis*, *12*(7). https://doi.org/10.1167/12.7.6

Watson, A. B., & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys*, *33*(2), 113-120. https://doi.org/10.3758/bf03202828

Waugh, S., Pelli, D., Álvaro, L., & Formankiewicz, M. (2018). Crowding distance in healthy children. *Journal of Vision*, *18*(10), 855-855. https://doi.org/10.1167/18.10.855

Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Res*, *15*, 1137-1141. https://doi.org/10.1016/0042-6989(75)90012-7

Westheimer, G., Shimamura, K., & McKee, S. P. (1976). Interference with line-orientation sensitivity. *J Opt Soc Am*, *66*(4), 332-338. https://doi.org/10.1364/josa.66.000332

White, A., Tang, K., & Yeatman, J. (2020). The field of view for word recognition: crowding and hemifield asymmetries. *Journal of Vision*, *20*(11), 911-911. https://doi.org/10.1167/jov.20.11.911

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn Sci*, *15*(4), 160-168. https://doi.org/10.1016/j.tics.2011.02.005

Worrall, N., & Coles, P. (1976). Visual field differences in recognizing letters. *Perception & Psychophysics*, *20*(1), 21-24. https://doi.org/10.3758/BF03198698

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *J Vis*, *10*(10), 16. https://doi.org/10.1167/10.10.16

Zhou, J., Benson, N., Pelli, D., & Winawer, J. (2017). Conservation of crowding distance in human V4. *Journal of Vision*, *17*(15), 19-19. https://doi.org/10.1167/17.15.19