An enhanced Bouma model fits a hundred people's visual crowding

Jan W. Kurzawski[1], Augustin Burchell[3], Darshan Thapa[2], Najib J. Majaj[2], Jonathan A. Winawer[1,2], 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, meridional asymmetry

ABSTRACT

Crowding is the failure to recognize an object due to surrounding clutter. Its strength varies across the visual field and individuals. To characterize the statistics of crowding—ultimately to relate psychophysics of crowding to physiology—we measured radial crowding distance and acuity of 105 observers along the four cardinal meridians of the visual field. Fitting the well-known Bouma law — crowding distance depends linearly on radial eccentricity — explains 52% of the variance in log crowding distance, cross-validated. Our enhanced Bourma model, with factors for observer, meridian, and target kind, explains 72% of the variance, again cross-validated. The meridional factors confirm previously reported asymmetries. We find a 0.62 horizontal:vertical advantage, a 0.92 lower:upper advantage, and a 0.82 right:left advantage. Crowding distance and acuity have a correlation of 0.41 at the fovea, which drops to 0.23 at ±5 deg along the horizontal midline. Acuity and crowding represent the size and spacing limits of perception. Since they are dissociated in clinical populations (Song et al., 2014; Strappini et al., 2017) and shown here to be only moderately correlated in our sample of mostly university students, clinical testing to predict real-world performance should consider measuring both. In sum, enhancing the Bouma law with terms for meridian, observer, and target kind provides an excellent fit to our 105-person survey of crowding.

INTRODUCTION

Crowding is the failure to recognize an object due to surrounding clutter (Bouma, 1970, 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, Weymouth, 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). 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; Toet & Levi, 1992), and increases with target-flanker similarity (Andriessen & Bouma, 1976; Kooi et al., 1994; Leat et al., 1999; Nazir, 1992) as well as the number of distractors (Felisbert et al., 2005; Grainger et al., 2010). For review see (Herzog et al., 2015; Levi, 2008; Pelli & Tillman, 2008; Strasburger, 2020; 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, Heath, et al., 1963; Liu & Arditi, 2000; Malania et al., 2007; Pelli et al., 2016; Siderov et al., 2013; Toet & Levi, 1992).

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 105 observers to characterize the statistics of crowding, looking at parameters across individuals. The comparison with fMRI will be reported separately (Kurzawski et al., 2021 VSS abstract). Here we report only the psychophysics. We tested with letters, which, after little or no training, provide many possible targets needed for quick testing (Pelli & Robson, 1991). Long term, we are also 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 (Liu et al., 2009; Petrov & Meleshkevich, 2011; Wallis & Bex, 2012). Better performance along the horizontal than vertical midline and upper than lower meridian is also observed with other visual stimuli and psychophysical tasks (Abrams et al., 2012; Cameron et al., 2002; Carrasco et al., 2004; Carrasco et al., 2001; Fuller et al., 2008; Himmelberg et al., 2020).

Crowding distance is a potentially valuable biomarker for several reasons. Crowding severely limits what we see and 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 (vss abstract). We measured crowding in 105 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 all letter-based (Anstis, 1974; Bouma, 1970, 1973; Ehlers, 1936; Ehlers, 1953; Stuart & Burian, 1962). 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., 2019), 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). Here, we use our data from 105 observers to study the relationship between acuity and crowding in the fovea and periphery.

In our 105 participants, we measured radial crowding distance at 13 locations in the visual field and acuity at 3 locations. We used the Sloan font in the periphery and the Pelli font in the fovea, plus a few Pelli thresholds in the periphery. As far as we know, Pelli is still the only font skinny enough to measure crowding distance in the fovea. We assessed crowding's variation with eccentricity along the four cardinal meridians and across individuals. We find that the mean Bouma factor is conserved across radial eccentricity, but varies across meridians, producing several asymmetries. On the other hand, we find stronger correlation on fixed-eccentricity rings than along the meridians.

METHODS

Crowding database

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 radial crowding distance with the Sloan and Pelli fonts, with a total of 105 observers (78 in Dataset A, 17 in Dataset B, and 10 in Dataset C). With Sloan we measured crowding at 12 different peripheral locations in the visual field: 3 eccentricities (2.5, 5, 10 deg) along the four cardinal meridians (upper, lower, left and right). With Pelli we measured crowding at the fovea (twice) and 5 deg eccentricity on the left and right meridians Not all observers were tested at all locations. With gaze tracking, we also tested 10 experienced observers at 5 deg eccentricity on all cardinal meridians (Dataset C). See Fig. 1 and Table 1 for details. 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). We report our results as Bouma factor (slope of crowding distance vs eccentricity) 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 (flanker symmetry matters too, as explained in Methods). Fig. 1 shows the spatial coverage of our testing.
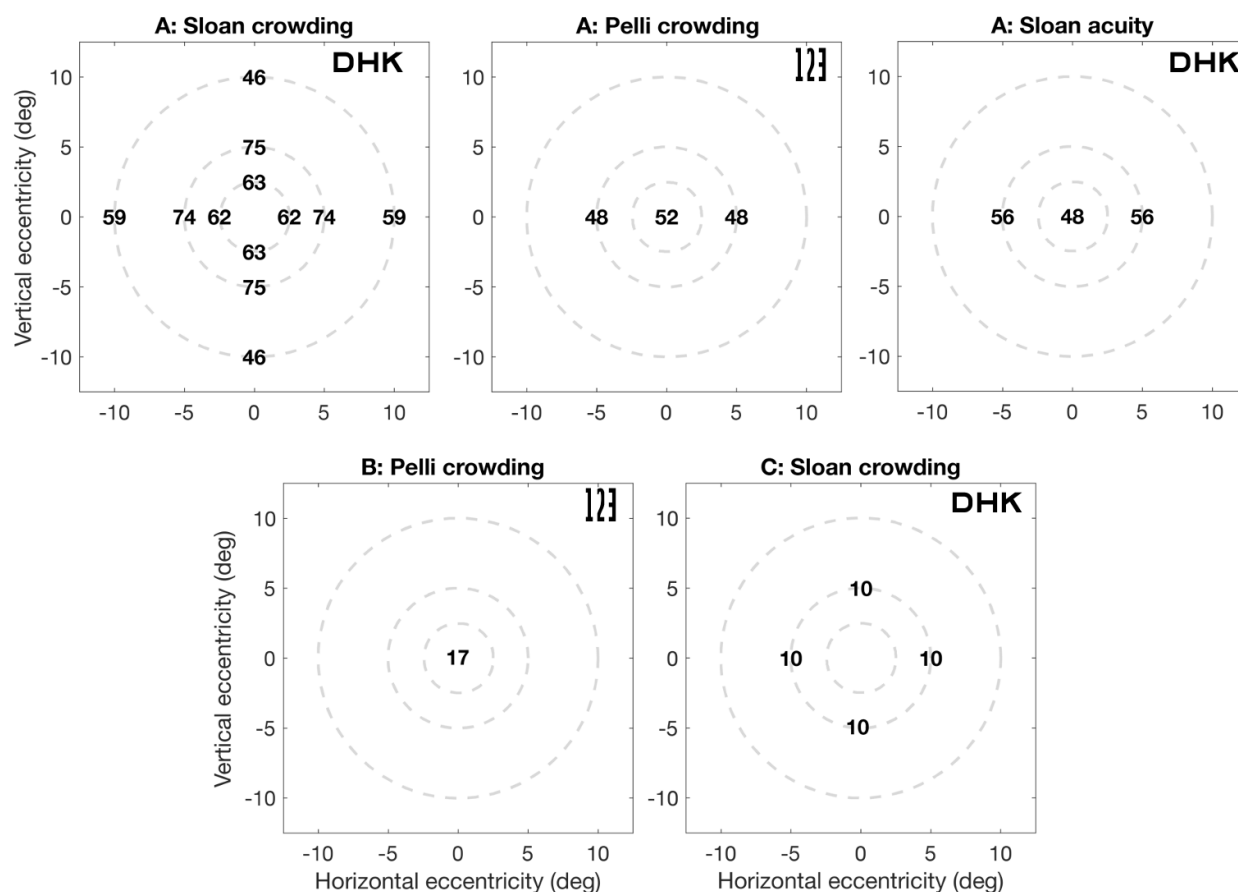
Figure 1 – Map of testing. Each plot title indicates the Dataset (A, B, or C), font (Sloan or Pelli), and task (crowding or acuity). Each plotted number represents the number of observers tested with that font and task at that location. We tested 78 people for Dataset A, 17 in Dataset B and another 10 for Dataset C, but not everyone in Dataset A was tested at all locations. Peripheral crowding was radial; foveal crowding was horizontal. Examples of Sloan letters and Pelli digits are shown as insets.

Participants

Our three datasets are described in Table 1, and their spatial coverage of the visual field is plotted in Fig. 1. The main study (Datasets A & B) tested 78 observers, mostly New York University undergraduate students. Our followup (Dataset C) tested 10 experienced observers with gaze tracking. They were NYU graduate students and postdocs, most of whom were experienced psychophysical observers. Each observer had normal or corrected-to-normal vision. All experiments were conducted in accordance with the Declaration of Helsinki and was approved by New York University's ethics committee on activities involving human observers.

Dataset	N	Measure	Font	Radial eccs. (deg)	Cardinal meridians	N	Thresholds per observer	Gaze tracking	Recruited	Notes
A	78	Crowding	Sloan	2.5, 5, 10	All	78	12	–	Mostly NYU undergraduates.	Number of thresholds varied between observers.
		Crowding	Pelli	0 (twice)	–	52	2	–		–
		Crowding	Pelli	5	Right, left	48	2	–		–
		Acuity	Sloan	0 (twice)	–	48	2	–		–
		Acuity	Sloan	5	Right, left	56	2	–		–
B	17	Crowding	Pelli	0 (twice)	–	17	2	–	Mostly NYU undergraduates	–
C	10	Crowding	Sloan	5	All	10	4	Yes	NYU graduate students and postdocs experienced in psychophysics.	–

Table 1. Our three datasets, each with an independent set of participants. In the periphery, we measured crowding distance radially, using log-spaced flankers. In the fovea, we measured crowding distance horizontally. Peripheral thresholds were measured once (35 trials). Foveal thresholds were measured twice.

Apparatus

Each testing session was completed on either an Apple iMac 27" (screen 3360 x 2100 pixels, 60.2 x 34.1 cm) or one of several MacBook Pro 15" laptops (2880 x 1800 pixels, 33.1 x 20.6 cm). Both computers have 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 three 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 250 cm. For ±2.5 and ±5 deg eccentricity the distance was 40 cm, and for ±10 deg eccentricity the

viewing distance was 25 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 (https://github.com/denispelli/CriticalSpacing) (Pelli et al., 2016) implemented in MATLAB 2016 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 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 radially, meaning that the centers of all three letters lay on a straight line from fixation.

Each testing session included several blocks, and was about an hour long. Most blocks measured two thresholds, interleaved, usually two crowding thresholds at mirror symmetric eccentricities.

Each block was restricted to just the vertical or horizontal midline. To minimize the temptation to look away from fixation toward an expected target location, we randomly interleaved conditions measuring thresholds at symmetric locations with respect to fixation, either horizontally (±x, 0) or vertically (0, ±y). A sample stimulus sequence for two trials appears in Fig. 2A. 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 pressing the corresponding key on the keyboard. For long viewing distances we use a wireless keyboard. Observers are instructed to return their eyes to fixation before clicking their response. The computer waits indefinitely for each response. The next trial began immediately after the observer's response.

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, $s=(\varphi_0+\varphi)b$. 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 because psychophysical crowding distance scales with eccentricity and cortical magnification scales inversely with eccentricity (Kurzawski, J., Winawer J., Pelli D., (2021). Conservation across individuals of cortical crowding distance in human V4, Submitted to annual meeting of Vision Sciences Society, St. Pete Beach, FL, May 21-26, 2021.). Therefore we designed our flanker placement so that the target is midway between them on the cortical surface. Thus, 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. More precisely, set φ_0=0.15 deg (we'll see later that this is the intercept of the Bouma law), then, given a target at eccentricity φ, we choose the two flanker eccentricities φ_{in} and φ_{out} so that

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

and report the inner spacing $s'=\varphi-\varphi_{in}$. Nearly all published papers on crowding have used linear flanker spacing. Datasets A, B, and C used log flanker spacing.

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 beta 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. And we set the "finger error" probability delta to 1%, so that QUEST can better cope with an occasional typing mistake. At the end of each block, QUEST estimates threshold (crowding distance in deg,

Fig. 2B). 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 width (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, but we typically randomly interleave two conditions in a block, to get a threshold for each condition, while keeping the observer uncertain as to which condition is present on each trial. For peripheral trials, to minimize the urge to "peek" away from fixation, we always interleaved testing at two eccentricities that are symmetrically arranged about the fixation point, e.g. ±5 deg on the horizontal or vertical midline. 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 don't explain the fixed number of trials per condition, we don't think they count, and this would affect only the last few trials.)
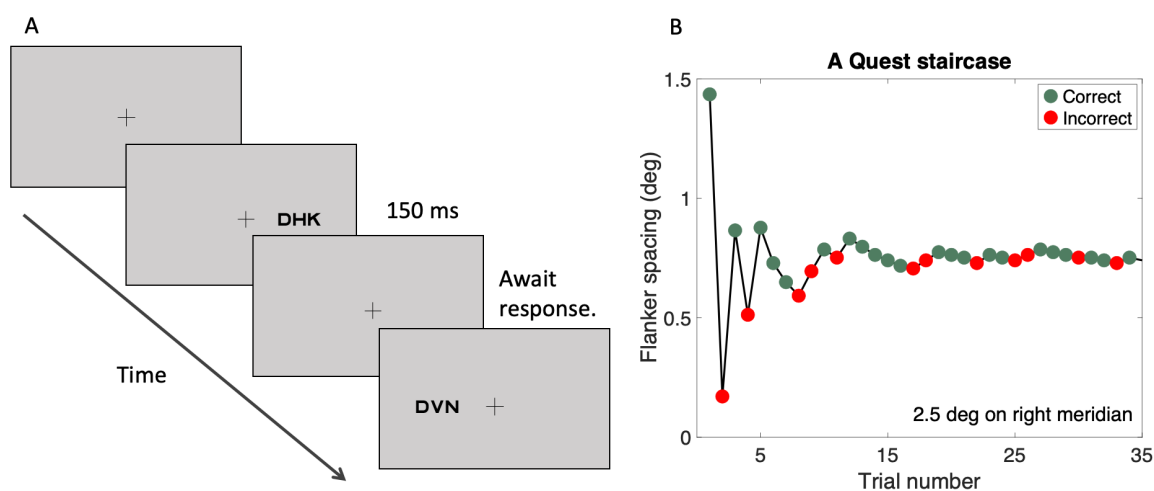
Figure 2 – Stimulus and procedure. A) The image sequence for two peripheral trials. While gazing at the fixation mark, which is always present, the observer presses a key, which begins the next trial. The target presentation is accompanied by a low-pitched purr. Then the observer identifies the target by pressing a key, hears a brief beep of reward if correct and silence otherwise, followed by the next target presentation. In this example, we randomly interleave two conditions, to measure thresholds at ±5 deg on the horizontal midline. By chance, the first trial shown is at +5 deg (i.e. right meridian), and the second is at -5 deg (left meridian). The actual target sequence is random, so the observer does not know which condition comes next. **B)** The staircase sequence of spacings tested on 35 successive trials of one condition, 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.

Analysis

For each peripheral location, we estimate the Bouma factor by $b \approx s/\varphi$, where s is the crowding distance in deg and φ is the radial eccentricity in deg. (This is based on the peripheral approximation Eq. 4, which neglects the small intercept φ_0 of the Bouma law Eq. 3.) The distribution of log crowding distance deviated from normality at 11 of 13 locations in the visual field, as assessed by the Jarque-Bera test. The number of tested observers varied between locations in the visual field, so to compare the Bouma factor across the visual field we used permutation tests, performed as follows. First, we calculated the difference between mean Bouma factors in the selected locations of the visual field (e.g. upper and lower meridians). We refer to this as the observed difference. Second, we shuffle the labels between the datasets and estimate the difference again (N = 1000). We refer to these values as the random differences. In doing so, we create a distribution of differences with 1000 entries. We then check how many times the absolute value of our random difference was higher than the absolute value of observed difference and divide this count by the number of permutations to estimate the *p*-value.

To report the mean Bouma factor we bootstrap the values at each location of the visual field (N=1000). With every iteration we sample the data with replacement to create a distribution of means. We use this distribution to report the mean and confidence intervals.

Model fitting

Fitting the Bouma law. Table 3 lists the Bouma law (Eq. 3) and several enhanced versions (Eqs. 7, 8, 9). 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.

Fitting the linear S approximation. The third column of Supplementary Table 2 lists simple model approximations, only valid in the periphery, that are linear in S and thus can be solved by linear regression. For peripheral data, this gives practically the same fit as the nonlinear fit with the model equation. We use a regression model with independent predictors for

participant, meridian, and target kind (i.e. font). We construct a design matrix having one constant column; 78 predictor columns, one for each participant; 4 predictor columns, one for each meridian; and two predictor columns, one for each font. All columns consist of zeros and ones, where a value of one in each row indicates that the measurement was performed with a specific participant, meridian, and font. We then estimate the participant, meridional, and font factors by solving the model, as in Eq. 1.

$$c = \vec{S} \setminus X \qquad (1)$$

where c is a vector of estimated parameters, $\vec{S}$ is a vector of log crowding distances, and X is the design matrix. Eq. 2 calculates variance explained by the model,

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

where $\vec{S}$ is the vector of log crowding distances, $\vec{S'} = Xc,$ and $<\vec{S}>$ is mean log crowding distance, a scalar. This linear regression (Eq. 1) finds the parameters that minimize the RMS error in predicting the log crowding distances.

Cross validation

First, we divide data into randomly partitioned subsets equal in size. As the length of our data is equal to 78 subjects times the number of thresholds, we set the size of the subset to the number of observers. In that way for each model we keep the size of the subset the same and only change the number of cross-validation steps necessary to predict the whole dataset. Therefore the number of subsets was set equal to the number of thresholds per observer (12, 14, or 15, depending on the model). 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. Model coefficients are fit to minimize squared error in predicted log crowding distance. Variance explained R^2 is calculated by Eq. 2.

Comparing crowding distance across studies

As shown in Table 2, we want to compare crowding distance (and Bouma factor) across studies that used various numbers of response choices (e.g. 9 Sloan letters vs. 2 orientations of a tumbling T) and various threshold criteria (e.g. 70% vs. 75% correct). Including the present one, we know of four studies that compare crowding distance across meridians.

Study	Target	Choices	Prob. of guessing γ	Proportion correct at threshold P	After correction for guessing $P*$	ΔB re Toet & Levi	b ratio re Toet & Levi
Datasets A, B, and C	Sloan or Pelli letter	9	1/9	0.70	0.66	0.09	1.23
(Greenwood et al., 2017)	Tumbling clock	4	1/4	0.80	0.73	0.12	1.32
(Toet & Levi, 1992)	Tumbling T	2	1/2	0.75	0.50	0	1
(Grainger et al., 2010)	Courier New letter or symbol	9	1/9	0.75	0.72	0.11	1.29

Table 2 – Comparing crowding distance (and Bouma factor) across studies. 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 compute the corresponding threshold criterion *P** to be applied after correction for guessing. The relative log threshold *ΔB* is computed in Fig. 3A. The *b* ratio is $10^{\Delta B}$

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*,

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

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, which we set to 2.3, based on fitting psychometric functions to hundreds of trials at several spacings by two experienced observers.

Fig. 3A 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. The various guessing rates γ, make the psychometric functions hard to compare. So we "correct for guessing",

$$P^* = (P - \gamma)/(1 - \gamma)$$

This is a popular transformation of psychometric data, usually justified by assuming that the guessing rate can be modeled 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

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

Fig. 3B 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^*. The Toet and Levi 1992 study came first, so we treat it as the reference. 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 Toet and Levi criterion would report.
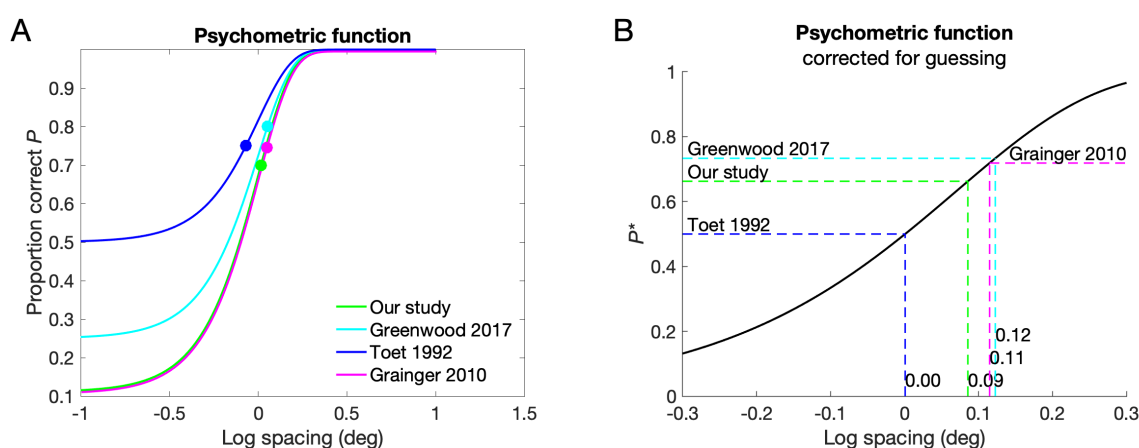
Figure 3 – Effect of criterion on threshold log crowding distance. A. The four studies used various numbers of response choices and threshold criteria (Grainger et al., 2010; Greenwood et al., 2017; Toet & Levi, 1992). **B.** Relative to the Toet & Levi parameters, the horizontal intercept indicates how much each study's parameters increase the reported threshold. These numbers appear in the ΔB column of Table 2.

Since this analysis produced nearly the same factor, about 1.3, for three of the four studies, we have not applied the correction. It would make more sense to shift the Toet and Levi estimates of Bouma factor up by a factor of 1.3:1. We ask the reader to imagine that.

Effect of symmetric placement of flankers with regard to either linear or log eccentricity

When measuring radial crowding, the target lies between two flankers on a radial line from fixation. Bouma spaced the flankers equally, and most investigators followed suit. However, we space 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 (Eq. 3) 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 see ("*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.

Effect of target kind (including font)

Pelli and Tillman (Pelli & Tillman, 2008) observed that crowding distance is remarkably similar for a wide range of target and flanker, but we will see later, comparing Bouma factors for many targets, that crowding distance for one letter font can be twice that for another (Table 4).

Minimizing fixation errors in our results

This is a mapping project, measuring crowding distance at many eccentricities. It is important that the eccentricities be accurate, but the target eccentricity depends partly on the observer's point of gaze, which is not directly under our control. We ask the observer to fixate on the fixation mark. When the target is foveal, most observers follow instructions to keep their gaze at fixation. However, when the target is peripheral, it is tempting to move one's gaze to the target — we call it "peeking" — and this is hard for untrained observers to resist. When the target is presented, an eye position away from fixation, closer to (or farther from) the target, would reduce (or increase) its eccentricity from what we intended, and

reduce (or increase) crowding. We did five things to minimize the presence of fixation errors in the statistics that we report.

1. *Fixation*. We instructed observers to re-fixate the fixation mark at the end of each trial before giving their keypress response, which initiates the next trial. For peripheral targets, the fixation mark was continuously present for the whole block, from the beginning of the first trial to the end of the last. (When the target was foveal, fixation was blanked 500 ms before target onset, and reappeared 500 ms after target offset.)

2. *Brief*. The target duration is brief, 150 ms, not enough time for a saccade to reach the target, so there is no harm in the observer making a saccade to the target in response to the target's presentation. When only doing one thing, observers reliably fixate. But it takes practice to stay on fixation while anticipating a peripheral target, so we worry about *anticipatory* eye movements toward an anticipated target location.

3. *Symmetric balance of temptation.* We tried to minimize the temptation to saccade to the anticipated target locations by randomly interleaving conditions with opposite eccentricities, symmetric about fixation: left and right or up and down. We suppose that equal expectations to either side of fixation tend to produce equal and opposite attractions, which might cancel.

4. *Excluding thresholds with poor QUEST convergence, i.e. for which the overall proportion correct was less than 50% or over 90%.* In our testing, with good fixation, QUEST converges quickly so that the set of 35 trials tends to have an overall proportion correct close to 70%. In the Supplement ("Estimating the effect of peeking") we calculate the effect of fixation errors ("peeking") on the proportion correct. To minimize the use of thresholds contaminated by fixation errors, we discarded any crowding distance estimate for which the proportion correct over the 35 trials (of the various spacings selected by QUEST) was far from 70%, i.e. less than 50% or over 90%. (Our guessing rate was 1/9=11%.) This excluded 5.6% of the thresholds. At 2.5 deg eccentricity we excluded 1/63 from the right and 2/63 from the left meridian. At 5 deg eccentricity we excluded 3/74 from the left meridian, 4/75 on the lower meridian, and 5/75 on the upper meridian. At 10 deg eccentricity we excluded 3/59 from the horizontal midline, 10/46 thresholds from the upper meridian, and 16/46 from the lower meridian. For the Pelli font we excluded 1/48 thresholds from the left meridian at 5 deg. We did not exclude any foveal thresholds because fixation is usually reliable for foveal targets.

5. *Gaze tracking.* In Dataset C, to validate our findings, we repeated the radial crowding measurements at 5 deg eccentricity on all four meridians on ten experienced observers while tracking their gaze with the Pupil Labs "Pupil Core" gaze tracker (Kassner et al., 2014). Eye position was sampled at 200 Hz, i.e. 30 times during the 150 ms target presentation. If the participant's gaze was more than 1 deg away from the central cross at any time during the target presentation, the trial was discarded, and he or she was alerted. To make up for the lost trial, we shuffled in a new trial of the same condition to the random list of conditions to be tested in that block. [With hindsight, we regret discarding the "bad" trials, instead of just putting them aside, as they would be useful for other analyses.]

Digitizing data from other papers

We used WebPlotDigitizer (Rohatgi, 2020), a web and desktop application to digitize data plotted in other documents. Notably, we were able to extract precise quantitative estimates of crowding's variation across observers and the visual field, as reported in Toet & Levi, 1992 (Fig. 6) (Toet & Levi, 1992) and Greenwood, 2017 (Fig. S1) (Greenwood et al., 2017). These radial crowding estimates are compared to the current study and other literature in Table 3.

RESULTS

Extending the Bouma law

The most salient feature of our data is the linear growth of crowding distance *s* with eccentricity φ, the Bouma law, which can be written as

$$\hat{s} = (\varphi_0 + \varphi)\, b \tag{3}$$

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). As shown in Table 3, the Bouma law (Eq. 3), with just two degrees of freedom, accounts for 52% of the variance in our measured log crowding distances.

The eccentricity offset φ_0 (in deg) is typically very small, e.g. 0.15 deg (Pelli et al., 2016) and thus negligible, $\varphi \gg \varphi_0$, in the periphery, i.e. eccentricity φ greater than 1 deg. All our non-foveal data are at eccentricities of at least 2.5 deg, so we can model them by neglecting φ_0 to write:

$$\hat{s} \approx \varphi\, b \qquad \text{for } \varphi \geq 1 \text{ deg} \tag{4}$$

Taking the logarithm (base 10) of both sides, we have

$$\log \hat{s} \approx \log \varphi + \log b \qquad \text{for } \varphi \geq 1 \text{ deg} \tag{5}$$

or

$$\hat{S} \approx \Phi + B \qquad \text{for } \varphi \geq 1 \text{ deg} \tag{6}$$

where $\hat{S}$ is the log of predicted crowding distance $\hat{s}$, Φ is log of radial eccentricity φ, and *B* is log of Bouma factor *b*. The Eq. 6 model explains 33% of the variance in our peripheral data (See Supplementary Table 2). This value is substantially lower than the 52% variance explained reported above primarily because it excludes the foveal measurements, not because the peripheral approximation is inaccurate.

Model	Eq.	Equation	R^2	Degrees of freedom
Bouma law	3	$\hat{s} = (\varphi_0 + \varphi)\, b$	52%	2
× meridional factor	7	$\hat{s} = \varphi_0 b + \varphi\, b_\theta$ where b = geom. mean b_θ	58%	5
× observer factor	8	$\hat{s} = (\varphi_0 b + \varphi\, b_\theta)\, o_i$	72%	92
× target-kind factor	9	$\hat{s} = (\varphi_0 b + \varphi\, b_\theta)\, o_i\, t_{kind}$	72%	93

Table 3 – How well four models predict log crowding distance measured at 13 visual locations (12 peripheral, 1 foveal) in 78 observers (Dataset A in Table 1). We begin with the Bouma law (Eq. 3). Successive models try to account for more variance by adding factors that depend on the meridian θ, observer i, and font. Most of our crowding thresholds were collected with the Sloan font. The target-kind (or font) parameter extends our model to accommodate the Pelli font as well. The R^2 column shows cross-validated variance accounted for in predicting log crowding distance (using Sloan and Pelli fonts) over the whole visual field (16 thresholds per observer). The values of all the fitted parameters appear in Supplementary Table 1. Supplementary Table 2 shows a shortcut that fits a simplified model to just the peripheral data.

Eq. 7 accounts for more variance by replacing b with b_θ, which depends on meridian θ,

$$\hat{s} = \varphi_0 b + \varphi\, b_\theta \tag{7}$$

where "b" now represents the geometric mean of b_θ, i.e. $b = \exp(\text{mean}(\ln b_\theta))$. Note that the meridian θ is undefined at the fovea ($\varphi = 0$), so $\varphi_0 b_\theta$ would be undefined there.

Eq. 8 accounts for yet more variance by adding a factor o_i for each observer i,

where $\prod_i o_i = 1$,

$$\hat{s} = (\varphi_0 b + \varphi\, b_\theta)\, o_i \tag{8}$$

Eq. 9 extends our model to accommodate any dependence on the target kinds (e.g. Sloan vs. Pelli), by adding a factor t_{kind} for each font (Sloan or Pelli),

$$\hat{s} = (\varphi_0 b + \varphi\, b_\theta)\, o_i\, t_{kind} \tag{9}$$

where t_{Sloan} = 1. The fit gives t_{Pelli} = 1.05. Over the whole visual field, the Eq. 9 model accounts for 72% of the variance, which, unsurprisingly, is no better than Eq. 8 since Sloan and Pelli have the same Bouma factor, so there is no font difference to account for. We will see below that the Bouma factor depends strongly on target kind, including letter font, so we anticipate that the target-kind factor will help explain datasets that have more diversity in target kind.

Except for target-kind, each factor that we added to the model increased the variance explained. The models are all cross validated, so the additional variance explained is not a necessary consequence of the increase in degrees of freedom. If the additional parameters were overfitting the training data, then we would not expect to find more variance accounted for in the left-out test data. While a number of methods exist to test for significant differences between such models, we find it is more informative to instead ask what, if any, systematic effects in the data correspond to the factors in the model. We do so in the next several sections.

Bouma factor vs meridian

In the previous section we found that adding an independent factor for each meridian to our model increases the proportion of variance accounted for. Here we show why. We first consider just the Sloan-font data, because we used Sloan for most of our crowding measurements. Fig. 3 shows the Bouma factor across the four meridians. One notable effect is that the Bouma factor varies systematically with meridian (Fig. 4A).
A second notable effect is that Bouma factor shows little dependence on eccentricity (Fig. 4B). This is expected from the Bouma law.
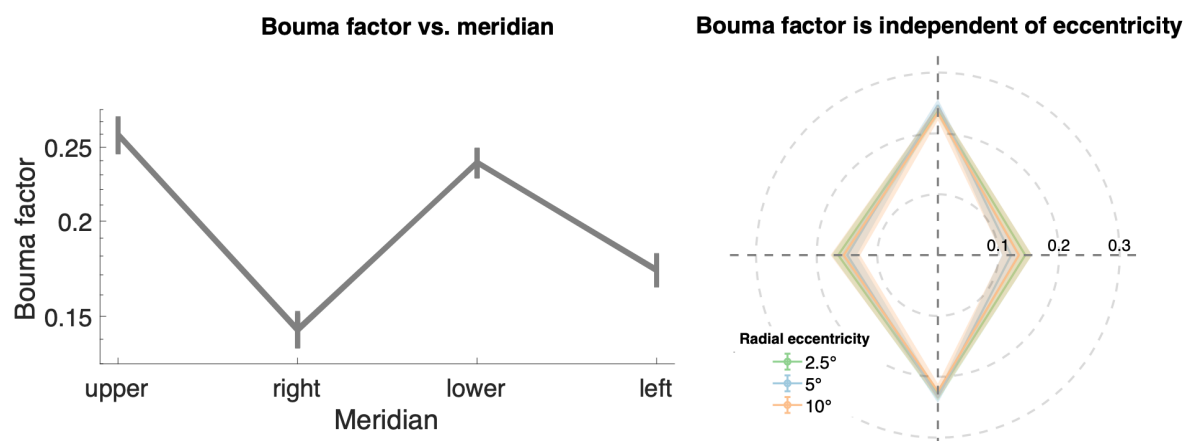
Figure 4 – Bouma factor vs. meridian. Radial crowding with the Sloan font (in Dataset A in Table 1). These plots complement Table 3, showing Bouma factor vs. meridian. The data are plotted on a cartesian axis in the main figure for visibility, and on a polar axis in Panel B. We find conservation of the Bouma factor across eccentricity but not across meridians. It is smallest in the right meridian and largest in the upper meridian. In the periphery, the Bouma factor is, to a good approximation, the crowding distance divided by the target's radial eccentricity (Eq. 4). In both plots each point represents the mean derived from bootstrap ($n = 1000$) and the error bars are 68% confidence intervals. To calculate the significance we used permutation tests. Data were included from all peripheral eccentricities ($p=0.04$ for lower vs. upper meridian, $p=0.03$ for right vs left meridians). Additionally, we performed Wilcoxon signed rank tests between left vs. right meridian and upper vs. lower meridian including data from all eccentricities or averaging data across eccentricities. Both tests yielded significant results.

To assess the generality of the meridian effects, we compared our results to several prior reports that measured crowding distances on more than one meridian. The studies and their results are summarized below in Table 2 and Fig. 4.

Row	Study	N	Gaze tracking	Target	Geometric mean b [95% confidence interval]				Scalar to compare w Toet & Levi
					Upper	Lower	Right	Left	
1	Dataset A	78	–	Sloan letter	0.26 [0.24, 0.29]	0.24 [0.21,0.25]	0.14 [0.12,0.16]	0.17 [0.16,0.19]	1.23
2	Dataset A	48	–	Pelli digit in periphery	–	–	0.14 [0.10, 0.18]	0.17 [0.13 0.21]	1.23
3	Dataset C	10	Yes	Sloan letter	0.39 [0.34, 0.43]	0.27 [0.23, 0.30]	0.17 [0.12, 0.23]	0.24 [0.17, 0.32]	1.23
4	(Greenwood et al., 2017)	12	Yes	Tumbling clock	0.58 [0.49, 0.68]	0.42 [0.35, 0.49]	0.28 [0.24, 0.33]	0.31 [0.25, 0.36]	1.32
5	(Toet & Levi, 1992)	6	–	Tumbling T	–	0.34 [0.27, 0.40]	–	0.24 [0.17, 0.30]	1
6	(Grainger et al., 2010)	29	–	Courier New letter	–	–	0.22 est. [0.18,0.27]	0.33 est. [0.26, 0.36]	1.29
7	(Grainger et al., 2010)	29	–	Courier New symbol	–	–	0.33 est. [0.28 0.38]	0.42 est. [0.37, 0.47]	1.29

Table 4 – Bouma factor vs meridian, for various targets. We compare our results with those of three studies that estimated crowding distance on more than one of the four cardinal meridians. The last column is copied from Table 2. We estimated the confidence intervals of Bouma factors in the Grainger et al. study based on the significance they reported for differences.
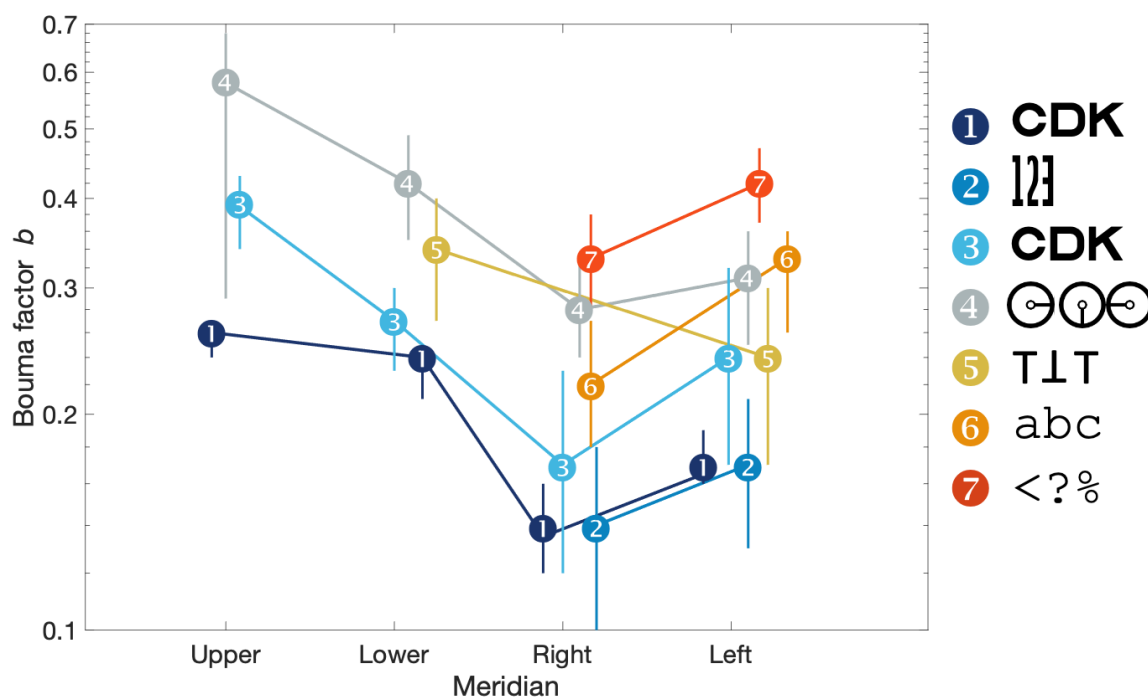
Figure 5 – Bouma factor vs. meridian, for various targets. Each numbered point corresponds to a numbered row of data in Table 4. To minimize occlusion, some points have been shifted left or right. Each point is a geometric mean derived from a bootstrap of *b;* the error bars indicate the 95% confidence interval. The legend shows examples of the objects that were used to measure crowding distance. The horizontal positions of the points are offset slightly for visibility.

In Table 2, only three rows — 1 and 3 (both Sloan) and 4 (Tumbling clock) — report the Bouma factor for all four cardinal meridians. Two of those rows, 3 are 4, used gaze tracking, so fixation is reliable on the trials retained for analysis. Fig. 4 shows that although the Bouma factor is higher for the clock than for Sloan, by a factor of 1.5 (ratio of geometric means), the two curves are otherwise similar, showing roughly the same dependence on meridian. The 1.5:1 difference is not an artifact of number of choices or threshold criterion, because Table 2 shows that the different threshold criteria predict a Bouma ratio of 1.07. Fixation errors can reduce the estimated Bouma factor by a factor of roughly 1.5:1 (compare row 3, without fixation errors, to row 1, with fixation errors), but lines 3 and 4 both used gaze tracking to exclude fixation errors. Thus, this seems to be a real 1.5:1 difference in Bouma factor between target kinds, precisely what the target-kind factor t_{kind} is meant to account for (Eq. 9). What we know about crowding suggests that the tumbling clocks are more similar to each than are the 9 Sloan letters are.

Pelli font

The Pelli and Sloan fonts have the same Bouma factors at the points where we measured both, ±5 deg on the horizontal midline. Pelli has higher SD across observers so it has larger confidence intervals, so its right advantage, identical to Sloan's, is not significantly different from 1. The Pelli font was originally designed to study crowding in the fovea and has a height-to-width ratio of 5:1 and a character-width-to-stroke-width ratio of 2 (Pelli et al., 2016). The Sloan font was designed to test acuity (usually in the fovea) and has a 1:1 height-to-width ratio and a character-width-to-stroke-width ratio of 5 (Sloan et al., 1952).

Asymmetries

Three asymmetries are listed in Table 5 and plotted in Fig. 5. For our Sloan data (row 1 and plot symbol 1), there is an advantage of 38% for the horizontal midline, 18% for right meridian, and 8% for lower meridian (all $p < 0.05$).

Row	Study	N	Gaze tracking	Target	Bouma ratio [95% CI]			Bouma factor on horizontal midline	SD of log Bouma factor on horizontal midline across observers
					Horizontal: vertical	Right:left	Lower:upper		
1	Dataset A	78	–	Sloan letter	0.62 [0.55,0.68]	0.82 [0.79,0.90]	0.92 [0.89,0.99]	0.16	0.19
2	Dataset A	48	–	Pelli digit in periphery	–	0.82 [0.60,1]	–	0.16	0.35
3	Dataset C	10	Yes	Sloan letter	0.62 [0.48, 0.76]	0.70 [0.61, 0.83]	0.69 [0.66 0.73]	0.22	0.18
4	(Greenwood et al., 2017)	12	Yes	Tumbling clock	0.59 [0.55,0.66]	0.91 [0.82,1.06]	0.71 [0.60,0.83]	0.31	0.13
5	(Toet & Levi, 1992)	6	–	Tumbling T	0.73 [0.43,0.91]	–	–	0.25	0.17
6	(Grainger et al., 2010)	29	–	Courier New letter	–	0.67 est. [0.56,0.76]	–	0.28	–
7	(Grainger et al., 2010)	29	–	Courier New symbol	–	0.80 est. [0.70,0.90]	–	0.37	–
Row	Study	N	Gaze tracking	Target	–			Crowding distance (deg)	SD of log crowding distance across observers (deg)
					–	–	–		
8	Datasets A and B	69	–	Pelli digit in fovea	–	–	–	0.09	0.27

Table 5 – Bouma factor and its asymmetries, compared with the literature. As in Table 4, we compare our results with those of all three studies that estimated crowding distance on more than one of the four cardinal meridians. Each row lists the target type, geometric average ± 95% CI of Bouma ratio across meridians (or midlines), geometric average Bouma factor calculated from Eq. 4, and the SD of log Bouma factor on the horizontal midline. Lacking access to the Grainger et al. data, we estimated their Bouma ratio SE by assuming they had the same SD as we do in row 1. By analogy to Toet and Levi's Tumbling T, we refer to any target that varied only in orientation as "Tumbling". Not included in the table, because they reported only proportion correct, not crowding distance, (He et al., 1996) were first to show the lower advantage over upper meridian. In their remote study of 793 observers, (Li et al., 2020) measured radial crowding distance for a tumbling circle with a gap at ±5 deg on the horizontal midline; however,

they reported only the average of the left and right crowding distances, so we cannot estimate the asymmetry from their published results.
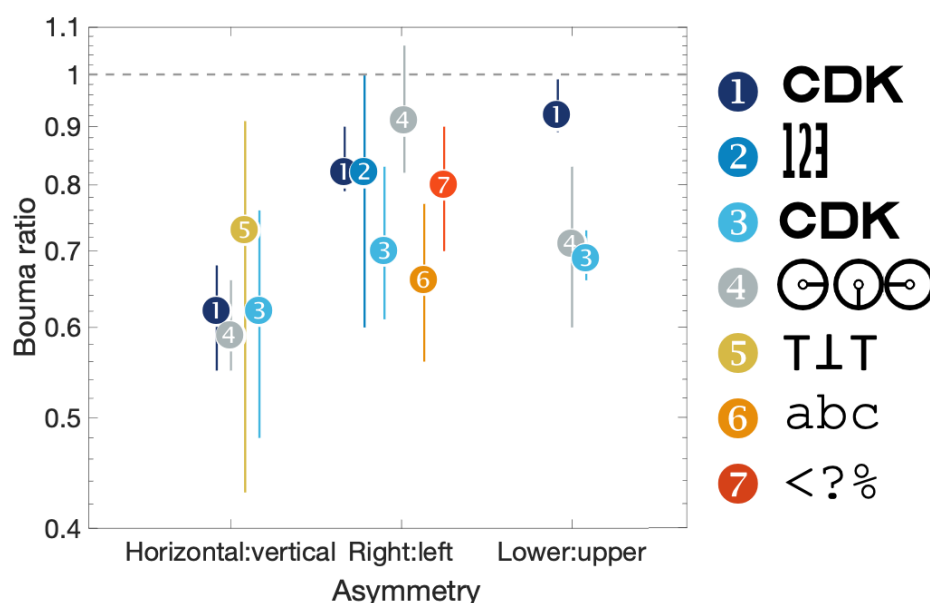
Figure 6 – Plot of the three asymmetries, expressed as Bouma ratios. Each numbered point corresponds to a numbered row of data in Table 5, based on several studies, including this one. The horizontal dashed line at a ratio of 1 represents equality of the two Bouma factors, no asymmetry. All but one of the points are significantly below 1. Each point is a mean derived from bootstrap *b;* the error bars indicate the 95% confidence interval. The legend shows examples of the objects that were used to measure crowding distance. All three estimates of the horizontal:vertical advantage are consistent with 0.6. Estimates of the right:left advantage cluster at about 0.66 (2, 6) and 0.9 (1, 3, 4) . The two estimates of the lower:upper advantage are inconsistent: 0.73 (4) and 0.92 (1). Thus the clusters range from about 0.6 to about 0.9.

As shown in Table 3 and Fig. 5, all the asymmetry estimates show an advantage for horizontal over vertical midline, right over left meridian, and lower over upper meridian. However, in row 4 (Greenwood et al., 2017) the right advantage for tumbling clocks is not significant.

Our Sloan results (row 1), based on 78 observers, are the biggest study and, consequently, have the smallest error bars.

Excluding fixation errors

Target eccentricity is a key parameter in crowding. As explained in Methods, we took several steps to minimize the inclusion of trials with fixation errors, since they present the target at a wrong eccentricity, different from what our analysis assumes. In all of our peripheral data, we interleaved conditions with opposite eccentricities to minimize the temptation to look at the

anticipated target location, since both locations are equally probable and equidistant on opposite sides of fixation. Here we consider all three datasets from Table 1.

Dataset A. 780 thresholds were collected on naive observers with no gaze tracking. We excluded thresholds for which the overall proportion correct was less than 50% or over 90%. Most of our figures and analyses are based on this dataset.

Dataset B. No thresholds were excluded as the whole dataset consisted of foveal crowding.

Dataset C. To confirm and validate our findings, we repeated the experiment with 10 experienced observers, while tracking their gaze (Pupil Labs "Pupil Core" gaze tracker). We measured radial crowding at 5 deg eccentricity along all four meridians. Participants were asked to repeat the trial if they were fixating more than 1 deg away from the fixation cross during stimulus presentation. In total, we discarded and redid 6.2% of all trials with gaze tracking.

Errors in fixation, or "peeking", could favor one possible target location at the expense of the other, and thus lead to asymmetric estimates of crowding distance, so we analyzed the two datasets separately. Dataset C recruited experienced observers and used gaze tracking to retain only trials with good fixation, allowing us to measure the asymmetries without fixation errors. At each meridian, the estimated Bouma factor is higher in Dataset C than A. The biggest difference is in the upper meridian, where the Bouma factor is 0.26 in Dataset A and and 0.39 in C. We replicated all three asymmetries: Crowding distance is smaller at the horizontal midline, right meridian, and lower meridian.

Dataset	Exclusion criterion	Font	Meridian				*N*	Thresholds before exclusion	Fraction of data excluded
			Upper	Lower	Right	Left			
A	Exclude thresholds with poor QUEST convergence.	Sloan	0.26 [0.24, 0.29]	0.24 [0.21, 0.25]	0.14 [0.12, 0.16]	0.17 [0.16, 0.19]	78	780	5.6% of thresholds
C	Use gaze tracking to exclude any trial with fixation error of more than 1 deg.	Sloan	0.39 [0.34, 0.43]	0.27 [0.23, 0.30]	0.17 [0.12, 0.23]	0.24 [0.17, 0.32]	10	40	6.2% of trials

Table 6. The Bouma factor [and 95% confidence interval] for each cardinal meridian in Datasets A and C. Each cell provides the geometric mean Bouma factor and its 95% confidence interval. Dataset A excluded thresholds with poor QUEST convergence, i.e. for which the overall proportion correct was less than 50% or over 90%. For Dataset C, we recruited experienced observers and used gaze tracking to repeat trials with bad fixation. 6.2% of all trials in Dataset C were repeated to replace bad fixations detected by gaze tracking.
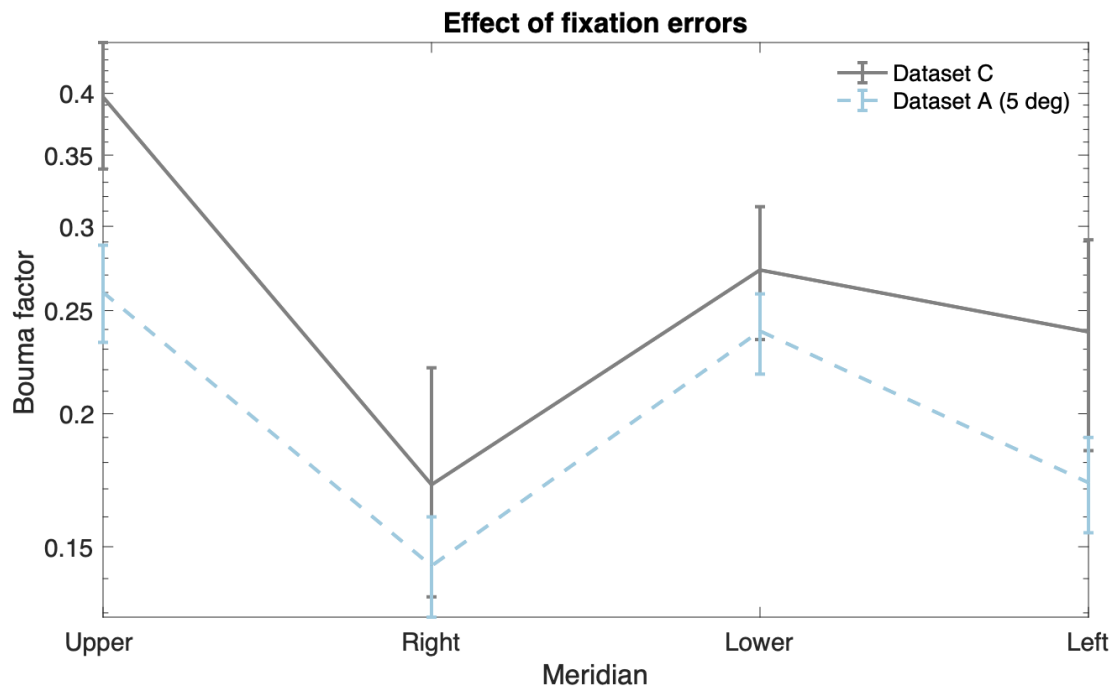
Figure 7 – Effect of fixation errors. Bouma factor for radial crowding with (Dataset A, dashed line) or without (Dataset C, solid line) fixation errors. Dataset C used gaze tracking to discard trials with bad fixation. Measured with the Sloan font at 5 deg eccentricity on all meridians.

Power analysis

Since our 78 observers revealed an asymmetry that is not commonly reported, we did a power analysis to calculate how many participants are needed to reveal each asymmetry, i.e. detect that the b ratio is not 1.00 with significance $p \leq 0.05$. Table 7 shows the results.

Suppose n observers, with Bouma factors b_i and b'_i, for $i = 1 \ldots n$, on two meridians (or midlines). Let $B_i = \log_{10} b_i$ and $B'_i = \log_{10} b'_i$ be the log Bouma factors. Across all observers, the mean, standard deviation, and standard error of the difference are

$$M = \text{mean } B_i - B'_i \tag{10}$$

$$SD = \text{stdev } B_i - B'_i \tag{11}$$

$$SE = SD / \sqrt{n} \tag{12}$$

The Student statistic t is

$$t = M / SE \tag{13}$$

The difference between meridians is significant if the absolute value of t exceeds a criterion of about 2:

$$|t| \geq 2 \tag{14}$$

Substituting from above for t and SE, and solving for n, we get

$$n \geq 4\,SD / M \tag{15}$$

For each asymmetry, we get M and SD from our data, and we compute what would have been the minimum number n of observers to achieve significance.

Asymmetry	*b* ratio	Observers Min.	Observers Actual	*B* difference Mean	*B* difference SD
Horizontal: vertical	0.62	3	74	-0.23	0.19
Right: left	0.82	13	74	-0.06	0.20
Lower: upper	0.92	42	74	-0.02	0.21

Table 7 – Power analysis. Based on our measurements of radial crowding distance with the Sloan font. For each asymmetry, the b ratio is the ratio of geometric mean Bouma factors across observers. The minimum number of observers for significance of the difference is computed by Eq. 15. The last two columns are the mean and SD across observers of the difference in log Bouma factor B between two meridians (or midlines) in the visual field. (The horizontal midline

factor is the geometric mean of the factors for right and left meridians; the vertical midline factor is the geometric mean of the factors for upper and lower meridians.)

Individual differences

Table 8 reports that individual differences, expressed as SD of log Bouma factor across observers, are twice as big for Pelli as for Sloan. We don't know why Pelli has higher SD than Sloan, or how to interpret the second mode in its histogram (Fig. 8). We suppose that fixation was reliable for the foveal data, and the peripheral histogram has a similarly bimodal shape. Perhaps there is a bimodality in the degree of similarity, and thus degree of crowding, between the various pairings of the symbols 1 to 9. If so, then changing or selecting digit shapes to achieve more consistent similarity might eliminate a hump and greatly reduce the SD.

Log Bouma factor across observers for horizontal midline
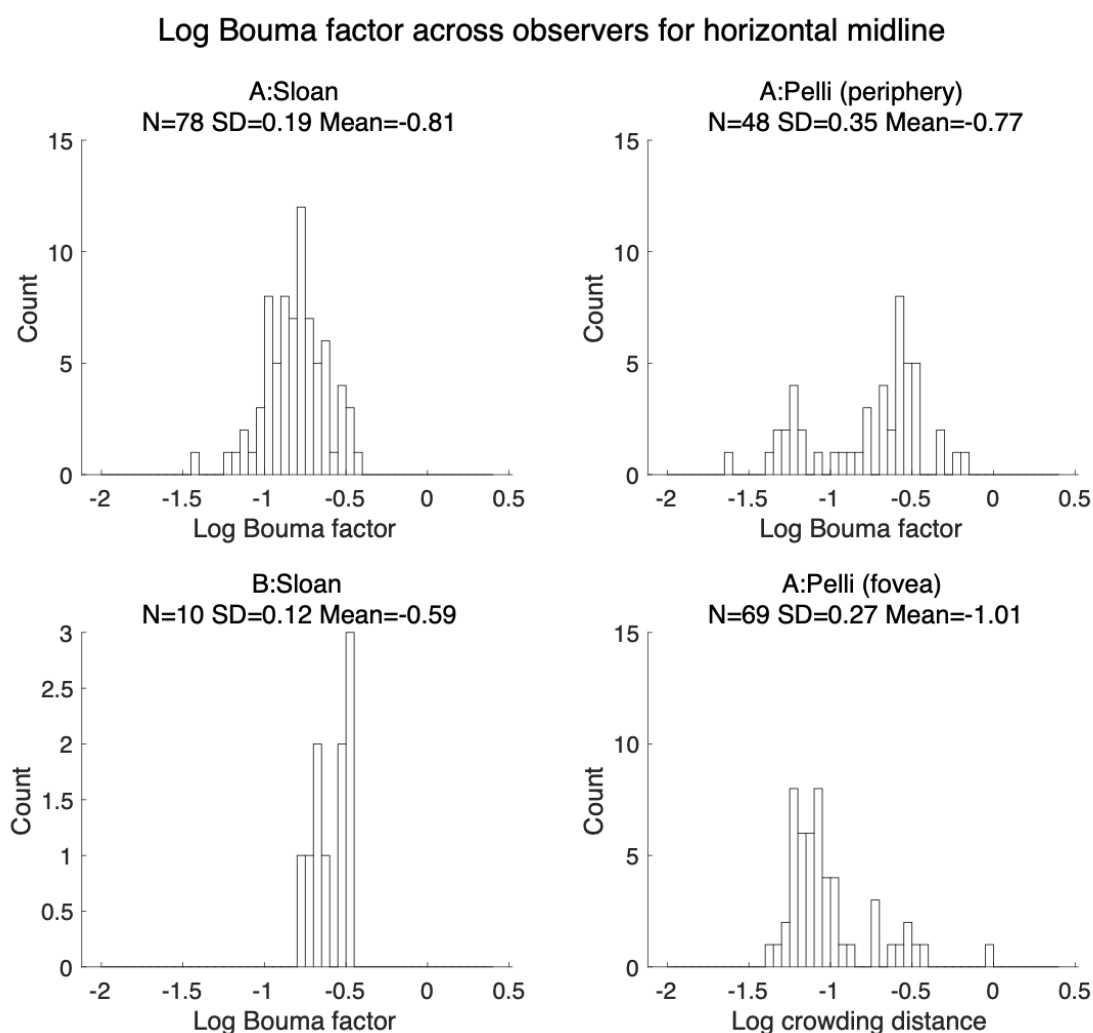
Figure 8 – Histograms of crowding for Sloan and Pelli fonts. (Periphery) Histogram of log radial Bouma factor *B* at radial eccentricities 2.5 and 5 deg, on all meridians, for Sloan and Pelli fonts. **(Fovea)** Histogram of log horizontal crowding distance for Pelli font.

Spatial correlation

This is the first attempt to measure and map the spatial correlations of crowding. Some correlations are surprisingly high and seem important to know. Most are still raw measurements awaiting an interpretive story.

Using our results from 88 observers, we correlated log crowding distance within observers across locations in the visual field. Each row of Table 9 presents the average correlation between two subsets of our data, always excluding the case of correlating a measurement with itself. We measured radial crowding distance at each visual field location once per observer, except in the fovea, where we interleaved two identical conditions to get two independent measurements of horizontal crowding distance. The most salient observation is that the correlations are always high when comparing locations with the same radial eccentricity ($0.52 \leq R \leq 0.83$). The correlations are always lower when the radial eccentricities differ ($0.04 \leq R \leq 0.51$). For the mixed case (last row) the correlation value is intermediate ($R = 0.43$).

|	Region 1	Region 2	Number of correlations	*R*, average correlation of log crowding distance
Same radial eccentricity	Fovea, Pelli font	Fovea, Pelli font, repeated measurement	1	0.83
	Right meridian	Symmetric point in left meridian	3	0.66
	Lower meridian	Symmetric point in upper meridian	3	0.52
	Ring (at 2.5°, 5°, or 10° radial ecc.)	Same ring	18	0.52
Different radial eccentricity	Any meridian	Same meridian	12	0.51
	Ring (at 2.5°, 5°, or 10° radial ecc.)	Other rings	48	0.42
	Fovea, Pelli font	Right meridian	3	0.27
	Fovea, Pelli font	Left meridian	3	0.13
	Fovea, Pelli font	Lower meridian	3	0.33
	Fovea, Pelli font	Upper meridian	3	0.17
Mixture of same and different radial eccentricities	Any meridian	Other meridians	54	0.43

Table 9 – Correlation of crowding across the visual field. The average correlation of log crowding distance between paired locations, one from each of the two specified regions. The green rows are correlations between points at the same radial eccentricity; the red rows are correlations between points at different radial eccentricities; and the yellow row includes both. For each row, we report the average correlation and the number of paired locations whose correlations were averaged. Note that all of the green *R* values are higher than all of the red, and the yellow is in between. The reported average may include many correlations, but always excludes the case of correlating a measurement with itself. The crowding distance at each point is radial, except at the fovea, where it is horizontal. Unless stated otherwise, each region excludes the fovea, and the font is Sloan. The last four red rows report the correlation between fovea and ±5 deg on horizontal midline, computed from testing with Pelli-Pelli or Pelli-Sloan fonts, to assess the effect of the font difference. One can wonder how different the foveal correlations would be if we measured all locations with Pelli font, as we did in the fovea. We did measure crowding with both Sloan and Pelli at ±5 on the horizontal midline. The foveal Pelli correlation with ±5° Pelli is 0.35 and 0.30 and with ±5° Sloan is 0.16 and 0.04.

Unsurprisingly, the highest correlation in Table 9 is the 0.83 correlation of test vs retest, which is for log crowding distance in the fovea with the Pelli font. Correlations in the periphery are also high when the radial eccentricity is matched (green rows in Table 7). For example, with the Sloan font, the correlation is 0.66 between corresponding points in the right and left meridians (e.g. 10° right vs 10° left), and 0.52 for upper and lower. The data points comprising these correlations are shown in Fig. 9. The plots demonstrate the correlations, the biases (right crowding distance less than left, lower less than upper), and the increase in crowding distance with eccentricity. The 0.66 right vs left correlation is surprisingly high, given that the locations are processed in different brain hemispheres.

Correlation between opposite points
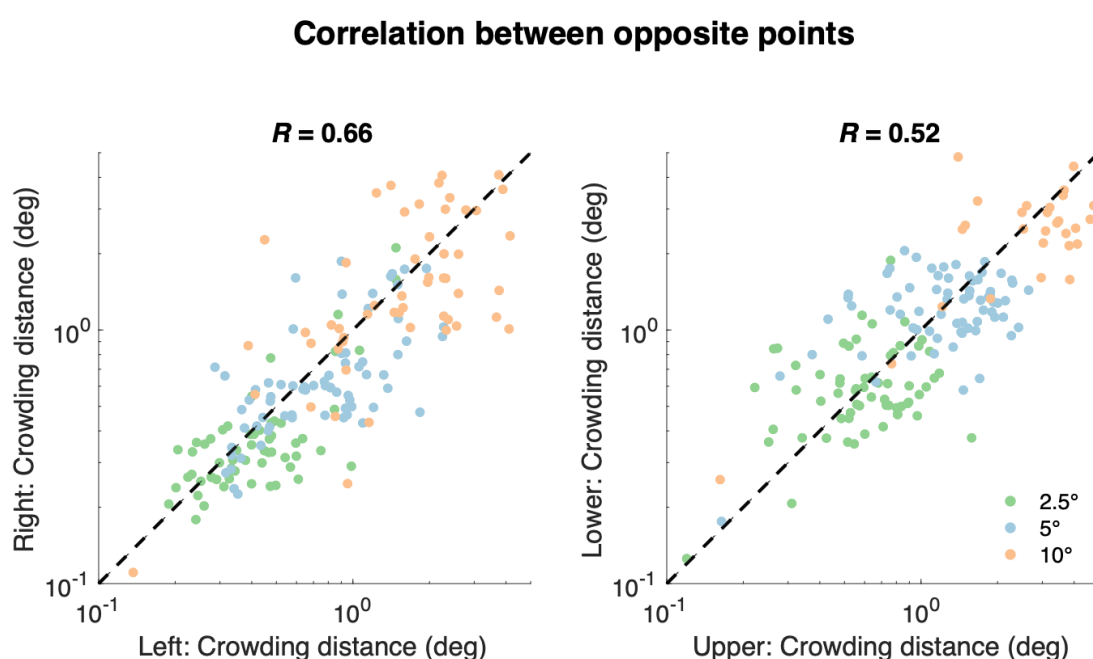
Figure 9 – Correlation of Sloan log radial crowding distance between locations symmetric with respect to fixation. A) Scatter diagram of crowding distance estimates along the right and left meridians. **B)** Scatter diagram of crowding measured along the lower and upper meridians. Data from all peripheral eccentricities. The black dashed line represents equality. The reported R is the average of three R's, one for each radial eccentricity.

Also surprising, the foveal correlation along the horizontal midline (Fig. 10) is not symmetric about the fovea and is non-monotonic with a minimum of 0.04 at -5 deg, rising roughly linearly on both sides to reach 0.23 at -10 and 0.35 at +10 deg. Along the vertical midline the correlation is high and constant (0.35) starting from -10 deg to 2.5 deg and drops for more peripheral locations in the upper meridian to about 0.1
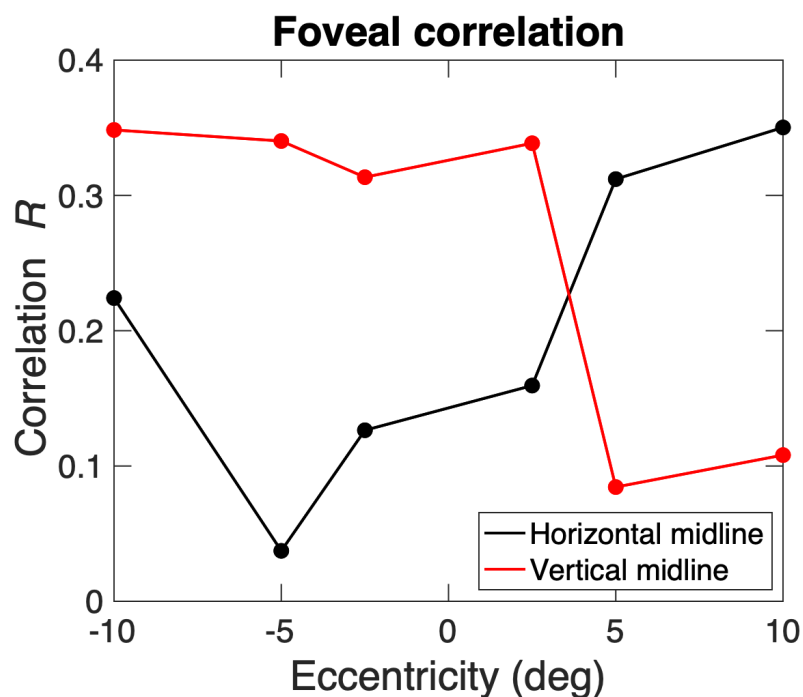
Figure 10 – Foveal correlation along the midlines. Correlation of foveal log crowding distance (Pelli) with peripheral log crowding distance (Sloan) along the horizontal and vertical midlines. Negative eccentricity corresponds to the left or lower meridian.

Crowding versus acuity

We included acuity in our measurements to study its link to crowding. The geometric mean Sloan acuity was 0.07 deg in the fovea and 0.23 deg at both +5 and -5 deg eccentricity along the horizontal midline. Correlating acuity and crowding distance separately at each location, we find 0.41 correlation in the fovea (Spearman's R; Fig. 11A). This correlation significantly drops when measured at ±5 degrees eccentricity, whether crowding is measured with Sloan (R=0.07; Fig. 11B) or Pelli (R=0.23; Fig. 11C).
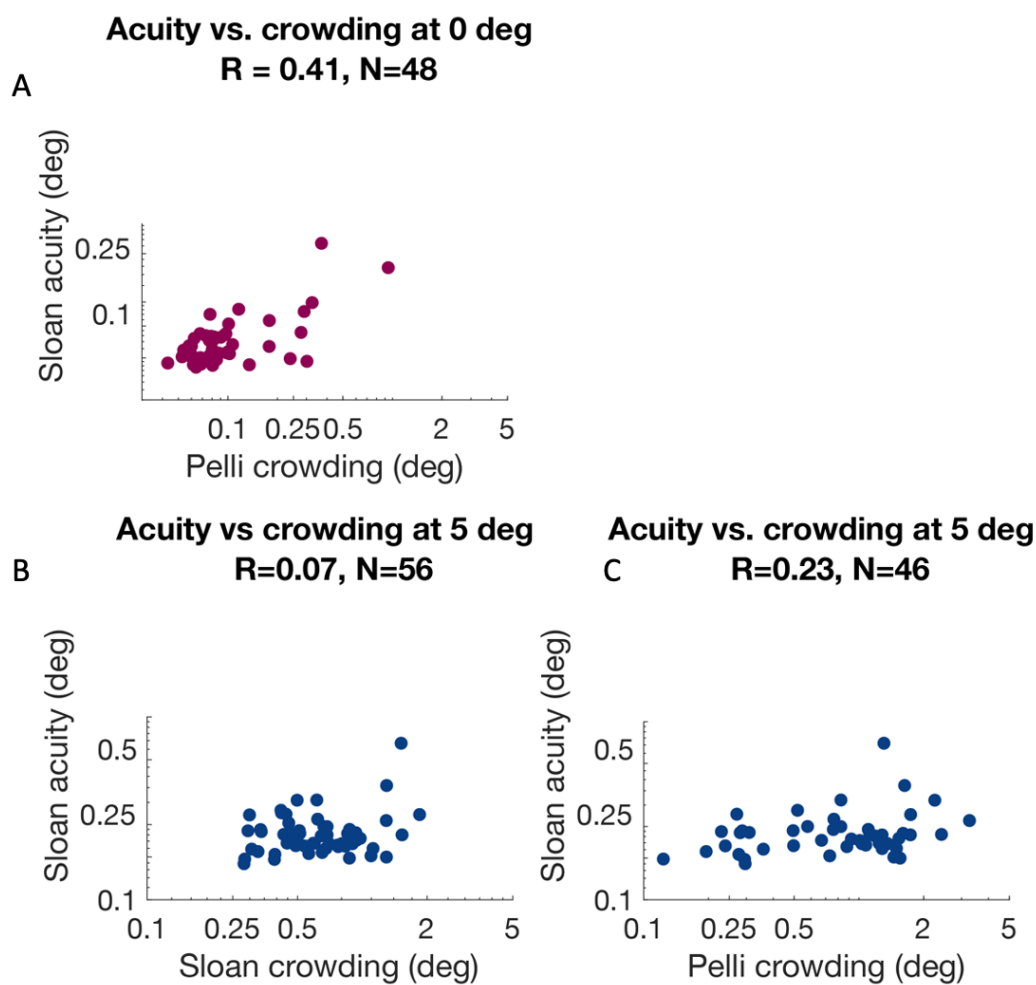
Figure 11 – Acuity vs. crowding. A. Sloan Acuity vs Pelli crowding at 0 deg eccentricity. **B.** Sloan Acuity vs Sloan crowding at 5 deg (geometric mean between left and right meridian , **C.** Sloan Acuity vs Pelli crowding at 5 deg (geometric mean between left and right meridian).

DISCUSSION

We present crowding and acuity measurements across the visual field. Across 105 observers, we replicate the Bouma law (crowding distance grows linearly with radial eccentricity) and three previously reported visual asymmetries. Taking the geometric mean across observers, the Bouma factor (for the Sloan font) varies across meridians, ranging from 0.14 in the right-horizontal meridian to 0.26 in the upper-vertical meridian. Comparing the Bouma factor across meridians, we replicate the previously observed horizontal advantage over the vertical midline, lower advantage over upper meridian, and right advantage over left meridian. By adding three new factors, we enhance the Bouma law to explain 72% of the variance (cross-validated) in log crowding distance over the whole visual field (out to 10 deg radial eccentricity). We find half as much correlation between acuity and crowding in the periphery (23%) as in the fovea (41%).

Right advantage in crowding

Consistent with past reports, we find a crowding advantage of the right over left meridian. The asymmetry is present with both Sloan and Pelli fonts and persists when gaze tracking is used to eliminate trials with fixation errors. Our power calculation (Table 7) indicates that reliably detecting the right-left asymmetry with the Sloan font requires at least 13 participants. (White et al., 2020) recently reported the right advantage from measurements on 22 observers of proportion correct, instead of crowding distance, for letter trigrams. And Greenwood et al. (2017), tested 12 observers using tumbling "clocks", and reported a right advantage in crowding area (proportional to product of radial and tangential crowding distances) (see their Fig 2A) (Greenwood et al., 2017).

Beyond crowding, other right advantages have been reported. For native readers of left-to-right written languages, like English, the right meridian is known to outperform the left in word recognition (Mishkin & Gorgays, 1952), and (Hagenbeek & Van Strien, 2002) reported an upper-right quadrant advantage for letter naming. Worrall and Coles (1976) examined letter recognition across the visual field and found a significant right hemifield advantage only along the right horizontal midline (Worrall & Coles, 1976).

Bouma factor

In their review, Pelli and Tillman (2008) suggested that the Bouma factor varies more across experimental designs than stimulus kinds (Pelli & Tillman, 2008). Bouma originally estimated the factor to be 0.5 (Bouma, 1970), which he later revised to 0.4 (Andriessen & Bouma, 1976), while several recent studies report values of 0.3-0.4 (see our Table 3). In our measurements, the geometric mean Bouma factor for Sloan letters was 0.19, measured with log-symmetric flankers, which we think corresponds to about 0.26 with traditional linear-symmetric flanker spacing.

Individual differences

In our model of log crowding distance we find that including an observer factor for each observer greatly increases the variance explained by the model, from 52% to 72%. This supports previous claims that crowding varies greatly across participants (Greenwood et al., 2017; Petrov & Meleshkevich, 2011; Toet & Levi, 1992). Fig. 8 shows a histogram of 78 individual Bouma factors spanning a 34:1 range.

Conservation of crowding distance on the surface of cortical area hV4

Our previous work links crowding with cortical architecture. Because crowding distance scales with eccentricity, while cortical magnification of visual areas scales inversely with eccentricity, cortical crowding distance is conserved across eccentricity (Pelli, 2008). We previously reported that crowding distance is also approximately conserved across orientation of the direction of crowding in hV4 but not V1, V2, and V3 (Zhou et al., VSS, 2018). In our recent work (Kurzawski, J., Winawer J., Pelli D., (2021). Conservation across individuals of cortical crowding distance in human V4, Submitted to annual meeting of Vision Sciences Society, St. Pete Beach, FL, May 21-26, 2021) we assessed conservation across individuals and found that cortical crowding distance is conserved in area hV4 (but not V1, V2, or V3) despite individual variations in map size (Benson et al., 2021) and crowding distance.

Acuity vs. crowding

Past studies reported that the BOLD activity in V1 is less for crowded than uncrowded stimuli (Millin et al., 2014) and that crowding distance correlates with the population receptive field sizes in V2 (Freeman & Simoncelli, 2011; He et al., 2019). Because these studies linked crowding to properties of early visual areas, and these early areas are also important for acuity, this might seem to support the suggestion that crowding and acuity are determined by the same optical and neural factors (Flom, Weymouth, et al., 1963; Levi et al., 1985). However, we find that crowding and acuity are 41% correlated at the fovea and only 23% at ±5 deg eccentricity. The low peripheral correlation challenges the conjecture of common cause. Crowding distance and acuity both grow linearly with eccentricity, but they have very different X intercepts — differing more than tenfold: 0.24 deg for crowding vs. 3.0 deg for acuity, based on our data here, which are similar to published values, e.g. 0.15 deg for crowding vs 2.7 deg for acuity (Pelli et al., 2016; Song et al., 2014) — and they are dissociated in two clinical populations (Song et al., 2014; Strappini et al., 2017).

CONCLUSIONS

Enhancing the Bouma law with terms for meridian, observer, and target kind provides an excellent fit to our hundred-person survey of crowding and acuity.

1. Fitting the well-known Bouma law — crowding distance depends linearly on radial eccentricity — explains 52% of the variance in log crowding distance, cross-validated.

2. Our enhanced Bouma model, with factors for observer, meridian, and target kind, explains 72% of the variance, cross-validated.

3. The meridional factors confirm previously reported asymmetries. We find a 0.62 horizontal:vertical advantage, a 0.82 right:left advantage, and a 0.92 lower:upper advantage. A power test indicates that 3 observers are needed to detect the horizontal advantage, 13 for the right advantage, and 42 for the lower advantage.

4. Bouma factors reported in the literature differ by more than two-fold.

5. Correcting for guessing to model the effects of guessing rate and threshold criterion accounts for very little of the difference in reported Bouma factors between most of the studies considered, but indicates that, because Toet & Levi used only 2 response choices (vertical or upside-down T), their crowding distances should be multiplied by 1.3 for comparison with three other studies.

6. Comparing crowding distance measured with and without fixation errors indicates that fixation errors by naive observers can reduce the measured Bouma factor by 40% and increase by 60% the SD of log Bouma factor.

7. In datasets free of fixation errors, the variation of log Bouma factor across observers has an SD of 0.19 in our data (Sloan font) and 0.13 in Greenwood's (tumbling clocks).

8. Comparing crowding distance for two different kinds of object (Sloan letter vs tumbling clock) measured with gaze tracking in different labs, we find a 2.4:1 ratio in Bouma factor that cannot be attributed to fixation errors, guessing rate, or threshold criterion. Sloan letters crowd each other more than tumbling clocks do, presumably because they are more similar.

9. Measured across observers, there are strong spatial correlations of log crowding distance. R is 0.66 for left-right symmetric points on the horizontal midline, despite

being processed in different hemispheres. The average correlation is 0.52 for different points with the same radial eccentricity (on any cardinal meridian), but only 0.42 with different radial eccentricities. The average correlation is 0.51 for different points on the same meridian (at any eccentricity), but only 0.43 with different meridians.

10. Crowding distance and acuity have a correlation of 0.41 at the fovea, which drops to 0.23 at ±5 deg along the horizontal midline.

11. Acuity and crowding represent the size and spacing limits of perception. Since they are dissociated in clinical populations (Song et al., 2014; Strappini et al., 2017) and shown here to be only moderately correlated in our sample of mostly university students, clinical testing to predict real-world performance should consider measuring both.

ACKNOWLEDGEMENTS

SUPPLEMENTARY MATERIAL

Extending the Bouma law

Model	Eq.	φ_0 (deg)	b	b_{upper}	b_{lower}	b_{right}	b_{left}	o_i	t_{Sloan}	t_{Pelli}
Bouma law	3	0.10	0.15	–	–	–	–	–	–	–
× meridional factor	7	0.09		0.22	0.20	0.11	0.14	–	–	–
× observer factor	8	0.09		0.22	0.20	0.11	0.14	1.52 1.70 1.28 1.23 0.49 1.50 1.20 0.98 1.22 0.40 0.42 0.68 0.48 0.90 1.35 2.08 1.49 1.69 1.18 0.99 1.07 1.11 0.88 0.67 0.82 0.54 1.95 0.62 0.78 1.20 1.24 1.39 0.88 0.82 1.05 1.66 1.63 0.73 1.53 1.59 1.48 0.04 1.67 1.18 1.37 1.06 0.66 1.69 0.92 1.52 0.88 1.44 1.02 0.37 2.26 0.81 1.18 1.21 1.10 1.36 1.26 0.76 1.93 1.70 0.92 0.71 0.83 1.07 1.34 0.73 1.31 0.51 0.73 0.68 0.76 1.13 0.88 1.01	–	–
× target-kind factor	8	0.09		0.22	0.21	0.11	0.13	1.51 1.70 1.28 1.23 0.49 1.50 1.20 0.99 1.22 0.40 0.42 0.68 0.50 0.89 1.35 2.07 1.48 1.68 1.19 1.00 1.08 1.11 0.89 0.67 0.82 0.54 1.95 0.61 0.78 1.21 1.25 1.40 0.87 0.81 1.06 1.65 1.62 0.72 1.53 1.58 1.48 0.05 1.66 1.19 1.37 1.05 0.66 1.68 0.92 1.52 0.87 1.44 1.02 0.39 2.24 0.80 1.18 1.21 1.10 1.37 1.25 0.75 1.91 1.70 0.91 0.70 0.82 1.06 1.34 0.73 1.31 0.52 0.73 0.67 0.75 1.12 0.87 1.01	1	1.05

Table S1 – Parameter values of the four model fits reported in Table 3.

Model	Eq.	Equation	Whole field		Peripheral approximation	Peripheral field			
			Sloan & Pelli fonts			Sloan font		Sloan & Pelli fonts	
			R^2	Deg. of f.		R^2	Deg. of f.	R^2	Deg. of f.
Bouma law	3	$\hat{s} = (\varphi_0 + \varphi)\, b$	52%	2	$\hat{S} \approx \Phi + B$	40%	1	33%	1
× meridional factor	7	$\hat{s} = \varphi_0 b + \varphi\, b_\theta$ where b = geom. mean b_θ	58%	5	$\hat{S} \approx \Phi + B_\theta$	50%	4	42%	4
× observer factor	8	$\hat{s} = (\varphi_0 b + \varphi\, b_\theta)\, o_i$	72%	92	$\hat{S} \approx \Phi + B_\theta + O_i$	69%	130	61%	92
× target-kind factor	9	$\hat{s} = (\varphi_0 b + \varphi\, b_\theta)\, o_i t_{kind}$	72%	93	$\hat{S} \approx \Phi + B_\theta + O_i + T_{kind}$	69%	131	61%	93

Table S2 – Comparing how well four models predict log crowding distance measured at 13 visual locations (12 peripheral, 1 foveal) in 78 observers. The green columns (an exact copy of Table) model the whole visual field, and that is what our abstract reports. The yellow columns present a handy shortcut for modelers. Uppercase variables are the $\log_{10}$ of corresponding lowercase variables, e.g. $B = \log_{10} b$. R^2 represents variance explained after model cross-validation. The first (green) R^2 column models crowding distance (using Sloan and Pelli fonts) over the whole visual field (16 thresholds per observer); the second (yellow) just Sloan over the periphery (12 thresholds per observer); the third (aslo yellow) Sloan and Pelli over the periphery (14 thresholds per observer). All our fitting minimizes error in log crowding distance so fitting green Eqs. 3, 7, 8, and 9 requires nonlinear minimization, while the yellow approximations can be fit by linear regression.

Estimating the effect of peeking

We wondered how sensitive our estimate of crowding distance is to the fraction of trials in which the observer is fixating near a possible target location instead of at the fixation mark. We call it "peeking", rather than "cheating", as our participants are generally trying to follow our instructions, but it takes practice to learn to control gaze, and in Datasets A and B we mostly tested new observers. To keep our model simple, we define a "*peek*" as fixating closer to a possible peripheral target position than to the fixation mark during the target presentation (regardless of which of the possible locations the target is actually displayed at).

In our peripheral testing, the fixation mark is midway between the two possible target locations. Suppose that when the target is presented, the observer's gaze is away from fixation, closer to one of the two possible target locations. Then, with 50% probability, if the target lands at the nearby location (possibly anticipated) its radial eccentricity will be reduced to half or less its nominal value, and so will the crowding distance, because it's proportional, so the task will be much easier. Otherwise, with 50% probability, the target will land at the symmetric point on the other side of the fixation mark, and its eccentricity and crowding distance will be increased from nominal by at least a factor of 1.5, making the task much harder.

We assume that with accurate fixation, the proportion correct P is a Weibull function of the log spacing X,

$$P(X) = \gamma + (1 - \gamma)[1 - \exp(-10^{\beta(X-S+0.0156)})]$$
(S1)

with a threshold log crowding distance S, where γ is the guessing rate (the reciprocal of the number of response choices), which we set to $\gamma = 1/9 = 0.11$, and β is a steepness parameter, which we set to $\beta = 2.3$, based on fitting psychometric functions to hundreds of trials at several spacings by experienced observers who fixate reliably. The offset 0.0156 makes P equal our threshold criterion value of 0.7 when X equals threshold S, $P(S)=0.7$.

Let ε be the probability of "peeking". When not peeking, on a fraction $1 - \varepsilon$ of the trials, the observer fixates correctly and identifies the target with probability P from Eq. S1. On the remaining trials, the fraction ε, the observer peeks, fixating one of the two possible target locations, and gets the trial right if the target lands there (at zero eccentricity), and just

guesses if the target lands at the other possible location (at double the nominal eccentricity). Thus, for this observer who fixates the target on fraction ε of the trials, we predict a proportion correct P_{near} for the occasionally-fixated target location, and P_{far} for the symmetrically opposite location that is never fixated:

$$P_{near} = (1-\varepsilon)P + \varepsilon \qquad\qquad\qquad\qquad (S2)$$

$$P_{far} = (1-\varepsilon)P + \varepsilon\gamma \qquad\qquad\qquad\qquad (S3)$$

As the peeking rate approaches 1, P_{near} approaches 1 and P_{far} approaches the guessing rate γ. We have 9 letters, so $\gamma = 1/9$. Thus, Fig. S3A shows that as ε grows from zero to 1, P' shrinks the function P down to the constant 0.56. The measured threshold (i.e. crowding distance) is the spacing at which proportion correct is 0.70 (horizontal dashed line). Because our 0.70 threshold criterion is not far from the 0.56 correct of 100% peeking, modest amounts of peeking have little effect on threshold. Indeed, if we were to use a threshold criterion equal to the 0.56 of pure peeking, then threshold (the spacing at which P reaches criterion) would be independent of the probability ε of peeking, provided it's less than 100%.

We use MATLAB's `interp1` function to find the thresholds, i.e. log spacing S at which $P=0.70$, S_{near} at which $P_{near} = 0.70$ and S_{far} at which $P_{far} = 0.70$. Fig. S3B plots the error in log crowding distance (relative to S) as a function of the peeking rate ε. This reveals that while a peeking rate over 33% is catastrophic (because it reduces the maximum proportion correct lower than our threshold criterion, leaving threshold undefined), a small peeking rate has only a small effect. This is clearer in Fig. S3C, where we plot the estimated asymmetry in log crowding distance $S_{far} - S_{near}$ as a function of peeking rate ε. A 9% peeking rate produces a 0.05 asymmetric in estimated log crowding distance, in our symmetric model.
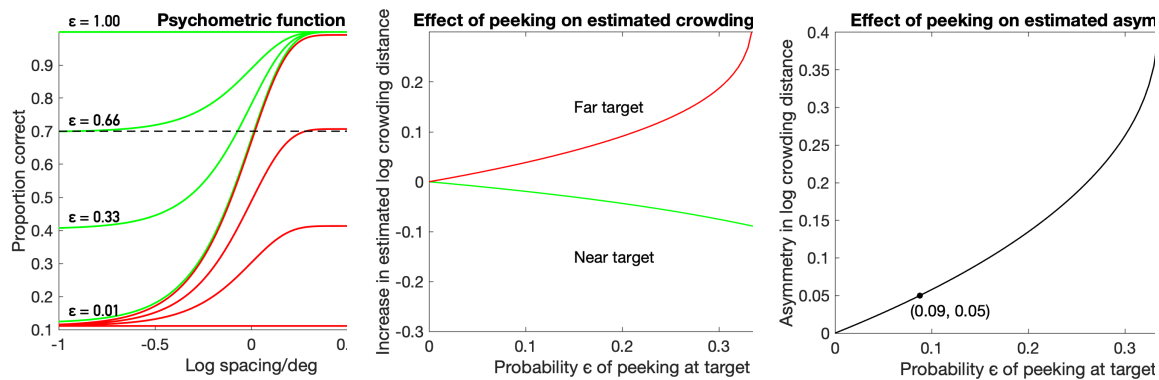
Figure S1. Effect of peeking. In a peripheral trial, the two target locations are equally probable and equidistant from the fixation mark. Let ε be the probability of "peeking", i.e. looking directly at a particular one of the two possible locations. **A.** The predicted psychometric functions P_{near} (green, Eq. S2) and P_{far} (red, Eq. S3) for recognition of the near and far targets with several probabilities of peeking, from 1% to 100%. The horizontal dashed line is the threshold criterion (70%). When $\varepsilon = 0$, we get the original psychometric function $P_{near} = P_{far} = P$, and there is no error in threshold (i.e. spacing at which $P=0.70$). Once ε becomes too large (larger than 0.33 in this example), the far threshold becomes undefined because $P_{far} < 0.70$ at all spacings. **B.** Error in estimated log crowding distance as a function of the probability ε of peeking. When $\varepsilon = 0.13$, the far error in log crowding distance is a mere $S_{far} - S = 0.05$, and the near error is even smaller: -0.02. **C.** The asymmetry $S_{near} - S_{far}$ in measured log crowding distance caused by peeking.

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 (Eq. 3) 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

$$\log(\varphi_0+\varphi) = [\log(\varphi_0+\varphi_{in}) + \log(\varphi_0+\varphi_{out})] / 2, \tag{S4}$$

where $\varphi_0=0.15$ deg, and report the spacing $s = \varphi-\varphi_{in}$. We can rearrange

$$\log(\varphi_0+\varphi_{out}) = 2 \log(\varphi_0+\varphi) - \log(\varphi_0+\varphi_{in}) \tag{S5}$$

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:

$$\Delta\Phi = \log(\varphi_0+\varphi_{out}) - \log(\varphi_0+\varphi_{in}) \tag{S6}$$

If we are in the periphery, i.e. $\varphi_{in} > 1$, then $\varphi_0=0.15$ is negligible, and we can simplify,

$$\Delta\Phi \approx \log(\varphi_{out} / \varphi_{in}) \tag{S7}$$

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:

$$\Delta\Phi \approx \log (\varphi+s)/(\varphi-s) \tag{S8}$$

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:

$$\Delta\Phi \approx 2 \log \varphi/(\varphi-s') \tag{S9}$$

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

$$\log (\varphi+s)/(\varphi-s) = 2 \log \varphi/(\varphi-s') \tag{S10}$$

Solve for s',

$$s' = \varphi - \varphi \; [(\varphi-s)/(\varphi+s)]^{0.5} \tag{S11}$$

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

$$b' = 1 - [(1-b)/(1+b)]^{0.5} \tag{S12}$$

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

$$b' \approx 0.845 \; b \tag{S13}$$

Correpondence between linear and log flanker spacing

- $b' = 1\text{-sqrt}[(1\text{-}b)/(1\text{+}b)]$
- $b' = 0.845\ b$

Error rms 0.010, max 0.014

(x-axis: Bouma factor b; y-axis: Bouma factor b')
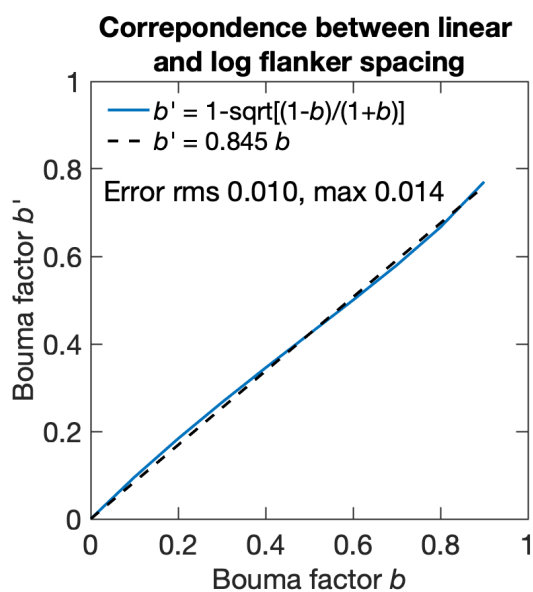
Figure S2. Log-symmetric spacing $s'=\varphi b'$ for same crowding effect as each linearly symmetric spacing $s = \varphi b$.

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

REFERENCES

Abrams, J., Nizam, A., & Carrasco, M. (2012, Jan 1). 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, Jan). 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, Aug 3). Visual crowding effect in the parvocellular and magnocellular visual pathways. *J Vis, 20*(8), 6. https://doi.org/10.1167/jov.20.8.6

Benson, N. C., Yoon, J. M. D., Forenzo, D., Kay, K. N., Engel, S. A., & Winawer, J. (2021). Variability of the Surface Area of the V1, V2, and V3 Maps in a Large Sample of Human Observers. *bioRxiv*, 2020.2012.2030.424856. https://doi.org/10.1101/2020.12.30.424856

Bi, T., Cai, P., Zhou, T., & Fang, F. (2009, Oct 14). 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, Apr 11). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177-178. https://doi.org/10.1038/226177a0

Bouma, H. (1973, Apr). 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

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

Cameron, E. L., Tai, J. C., & Carrasco, M. (2002, Apr). Covert attention affects the psychometric function of contrast sensitivity. *Vision Res, 42*(8), 949-967. https://doi.org/10.1016/s0042-6989(02)00039-1

Carrasco, M., Giordano, A. M., & McElree, B. (2004, Jun). Temporal performance fields: visual and attentional factors. *Vision Res, 44*(12), 1351-1365. https://doi.org/10.1016/j.visres.2003.11.026

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

Coates, D. R., Levi, D. M., Touch, P., & Sabesan, R. (2018, Jun 15). Foveal Crowding Resolved. *Sci Rep, 8*(1), 9177. https://doi.org/10.1038/s41598-018-27480-4

Ehlers, H. (1936, 1936/03/01). V: THE MOVEMENTS OF THE EYES DURING READING. *Acta Ophthalmologica, 14*(1-2), 56-63. https://doi.org/https://doi.org/10.1111/j.1755-3768.1936.tb07306.x

Ehlers, H. (1953, Apr). 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, 1990/10/01). Differences in the legibility of letters at contrast threshold using the Pelli-Robson chart. *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

Flom, M. C., Heath, G. G., & Takahashi, E. (1963, Nov 15). Contour Interaction and Visual Resolution: Contralateral Effects. *Science, 142*(3594), 979-980. https://doi.org/10.1126/science.142.3594.979

Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963, Sep). Visual Resolution and Contour Interaction. *J Opt Soc Am, 53*, 1026-1032. https://doi.org/10.1364/josa.53.001026

Fortenbaugh, F. C., Silver, M. A., & Robertson, L. C. (2015, Feb 12). 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, Aug 14). Metamers of the ventral stream. *Nat Neurosci, 14*(9), 1195-1201. https://doi.org/10.1038/nn.2889

Fuller, S., Rodriguez, R. Z., & Carrasco, M. (2008, Jan 24). Apparent contrast differs across the vertical meridian: visual and attentional factors. *J Vis, 8*(1), 16 11-16. https://doi.org/10.1167/8.1.16

Grainger, J., Tydgat, I., & Issele, J. (2010, Jun). 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, Apr 25). 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

Hagenbeek, R. E., & Van Strien, J. W. (2002, Jun). Left-right and upper-lower visual field asymmetries for face matching, letter naming, and lexical decision. *Brain Cogn, 49*(1), 34-44. https://doi.org/10.1006/brcg.2001.1481

He, D., Wang, Y., & Fang, F. (2019, Jul 8). The Critical Role of V2 Population Receptive Fields in Visual Orientation Crowding. *Curr Biol, 29*(13), 2229-2236 e2223. https://doi.org/10.1016/j.cub.2019.05.068

He, S., Cavanagh, P., & Intriligator, J. (1996, Sep 26). 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., Winawer, J., & Carrasco, M. (2020, Sep 2). Stimulus-dependent contrast sensitivity asymmetries around the visual field. *J Vis, 20*(9), 18. https://doi.org/10.1167/jov.20.9.18

Kassner, M., Patera, W., & Bulling, A. (2014). *Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction* Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, Seattle, Washington. https://doi.org/10.1145/2638728.2641695

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

Kwon, M., Bao, P., Millin, R., & Tjan, B. S. (2014, Nov 15). Radial-tangential anisotropy of crowding in the early visual areas. *J Neurophysiol, 112*(10), 2413-2422. https://doi.org/10.1152/jn.00476.2014

Latham, K., & Whitaker, D. (1996, 1996/01/01/). 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, Feb). 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, Feb). 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, Dec 15). 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, Jan 22). 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., Jiang, Y., Sun, X., & He, S. (2009, Jan 27). 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, Nov 26). 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, Jan 29). Grouping of contextual elements that affect vernier thresholds. *J Vis, 7*(2), 1 1-7. https://doi.org/10.1167/7.2.1

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005, Feb 4). 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, Dec). Visual crowding in V1. *Cereb Cortex, 24*(12), 3107-3115. https://doi.org/10.1093/cercor/bht159

Mishkin, M., & Gorgays, D. G. (1952, Jan). 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, Sep 20). 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, Apr). 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

Pelli, D., & Robson, J. (1991, 01/01). 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, 01/19). A clinical test for visual crowding. *F1000Research, 5*. https://doi.org/10.12688/f1000research.7835.1

Pelli, D. G. (2008, Aug). 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, Dec 30). 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

Pelli, D. G., & Tillman, K. A. (2008, Oct). 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, Oct 26). 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, May 25). Asymmetries and idiosyncratic hot spots in crowding. *Vision Res, 51*(10), 1117-1123. https://doi.org/10.1016/j.visres.2011.03.001

[Record #409 is using a reference type undefined in this output style.]

Rosen, S., Chakravarthi, R., & Pelli, D. (2014, 08/01). 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, Oct 14). Capabilities and Limitations of Peripheral Vision. *Annu Rev Vis Sci, 2*, 437-457. https://doi.org/10.1146/annurev-vision-082114-035733

Siderov, J., Waugh, S. J., & Bedell, H. E. (2013, Jan 25). Foveal contour interaction for low contrast acuity targets. *Vision Res, 77*, 10-13. https://doi.org/10.1016/j.visres.2012.11.008

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, May 5). 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

Strappini, F., Pelli, D. G., Di Pace, E., & Martelli, M. (2017, Apr). Agnosic vision is like peripheral vision, which is limited by crowding. *Cortex, 89*, 135-155. https://doi.org/10.1016/j.cortex.2017.01.012

Strasburger, H. (2020, May-Jun). 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, Jun). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Percept Psychophys, 49*(6), 495-508. https://doi.org/10.3758/bf03212183

Stuart, J. A., & Burian, H. M. (1962, Mar). 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, Jul). 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

Tyler, C. W., & Likova, L. T. (2007, Jul 30). Crowding: a neuroanalytic approach. *J Vis, 7*(2), 16 11-19. https://doi.org/10.1167/7.2.16

Wallis, T. S., & Bex, P. J. (2012, Jul 13). 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, Feb). 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, Oct). 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, Apr). 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, Apr). 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, 1976/01/01). Visual field differences in recognizing letters. *Perception & Psychophysics, 20*(1), 21-24. https://doi.org/10.3758/BF03198698

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