nature
neuroscience

ONLINE SUPPLEMENT

The uncrowded window of object recognition

Denis G Pelli & Katharine A Tillman
Psychology & Neural Science, NYU
http://psych.nyu.edu/pelli/

This is the online supplement to our Perspective review.

SUPPLEMENTARY DISCUSSION

The Bouma law at the cortex

The known eccentricity-dependence of the *cortical magnification factor* (mm on the cortex per deg of visual angle) produces a logarithmic map of the visual field on the primary visual cortex (V1). The logarithmic transformation of the proportional critical spacing at the visual field results in a fixed critical spacing at the cortex (6 mm at V1), independent of eccentricity.

Let us work this out. Bouma showed that critical spacing $\Delta\varphi$ is $b\varphi$, where b is Bouma's proportionality constant between critical spacing and eccentricity φ (ref. [1]). In V1 and many other areas in the visual cortex, eccentricity φ in the visual field is an exponential function of position d (in mm) on the cortex, $\varphi=\exp[\beta(d+\alpha)]$, where α and β are empirical constants, unique to each cortical area[2]. So position d is $(\log \varphi)/\beta - \alpha$, a logarithmic map. If the target is at eccentricity φ, then a flanker one critical spacing farther from fixation will be at eccentricity $\varphi+b\varphi$. b and β are fixed constants, so the cortical separation $\Delta d = d_{\varphi+b\varphi} - d_{\varphi} = \log(1+b)/\beta$ is a fixed number of mm, independent of target location, in every cortical area that is logarithmically mapped[3, 4]. We take $b = 0.4$, as in **Fig. 5**, so in V1, where $\beta = 0.0577$/mm, Δd is 6 mm.

Spatial extent of crowding

The invariance of critical spacing demonstrated in **Fig. 5** is found when the target and flankers have similar features (e.g., black letters flanking a black letter target). These typical cases produce maximum crowding. Flankers that have features unlike the target (e.g., white letters flanking a black letter target, on a gray background) produce much less crowding or none at all[5-7]. This weaker effect is usually reported as a reduction in critical spacing, but perhaps the spatial extent of crowding is unchanged and the effect is only reduced in amplitude (strength). The reported reduction of critical spacing may be an artifact of defining critical spacing by a performance criterion, as discussed below. Compared to the effect of target-like flankers, dissimilar flankers may simply have a weaker effect over the same spatial extent. Incidentally, the field still lacks an objective definition of similarity to predict crowding. We were surprised to learn that first- and second-order letters crowd each other, with the usual critical spacing, despite having very different features[8]. It might seem that similarity, for the purpose of crowding, would be just the opposite of salience, but in fact salience has little or no effect on crowding[9].

Crowding has usually been characterized by just one number, "critical spacing", i.e., spacing threshold, the spacing required to achieve a criterion level of performance. That single number seems to be enough to characterize crowding when the flanker is similar to the target, but may not adequately describe the weaker crowding produced by dissimilar flankers. Disentangling the amplitude and extent of crowding demands a two-number description. The complete 'psychometric function', plotting proportion correct as a function of spacing, tells us little more than the critical spacing. Proportion correct has a small dynamic range bounded by the floor at chance, when spacing is below critical, and by the ceiling at 100%, when spacing is above critical. To get the whole story, we must replace proportion correct by a better dependent measure: threshold. To measure threshold, one varies a physical parameter of the stimulus to achieve a particular level of performance[10]. Thus, threshold is measured on a physical scale with a wide dynamic range. For example, several studies have measured orientation discrimination thresholds as a function of spacing. These plots show that the weaker crowding produced by less-similar flankers has much less amplitude (maximum threshold elevation) but practically the same spatial extent[6, 7, 11].

Changing the orientation of the flankers from parallel to orthogonal (to the target) halves the amplitude without obvious reduction of extent[6]. Arranging the flankers in a ring to form a closed contour reduces the threshold elevation by a factor of 6 without reducing its spatial extent (defined as spacing for half maximum log threshold elevation)[12].

Crowding diminishes somewhat with practice, but the improvement is specific to the trained strings[13] and does not transfer from 3-letter strings to reading[14]. The benefit of practice has been reported as a reduction in critical spacing, but a two-parameter analysis might reveal, as above, that only amplitude (not extent) is affected.

At present, the simplest account is that the spatial extent of crowding for a particular location and direction is independent of the particular target and flanker. However, that conclusion is tentative because most published studies have not disentangled the amplitude and extent of crowding. Thus, this review focuses on "critical spacing", but the reader should bear in mind that the special cases we just discussed demand a two-parameter (amplitude and extent) characterization of crowding.

Uncrowded neighborhood and search

An object must be in the observer's uncrowded window to be recognized. Inverting the idea, **Supplementary Fig. 1** shows the object's *uncrowded neighborhood* — that is, the area around the object within which you must fixate to see it uncrowded. The uncrowded neighborhood is much like Engel's "conspicuity area", though he did not mention crowding[15]. Only those objects whose uncrowded neighborhoods include the observer's point of fixation are recognizable. If the observer fixates randomly, then the probability that the fixation will land in a particular object's uncrowded neighborhood (and thus that the object can be recognized) is equal to the fraction of the image area occupied by the uncrowded neighborhood. In the popular children's book *Where's Waldo?*[16], your chance of finding Waldo in your first glimpse is proportional to the area of his uncrowded neighborhood[15, 17].

Descriptions of crowded viewing

In 1936, the Gestalt psychologist Wolfgang Metzger described crowded viewing: "Farther out [in the periphery], the structure becomes ever weaker and cruder … The unifying effect of proximity becomes overwhelming. … [D]ifferences … cause an imbalance and restlessness in each intrafigural organization that is difficult to describe and can best be compared with what, in clearly seen objects, is called … texture … You see in that region … no clearly segregated, countable, or, above all, individually identifiable component parts"[18]. In 1976, Jerry Lettvin added, "Things are less distinct as they lie farther from my gaze. It is not as if things there go out of focus … it's as if somehow they lost the quality of 'form'"[19].

Peripheral vision and texture

What do we see when vision is crowded? We see stuff (unnamed texture) and perceive space (the shape of the scene we are in). Location affects perception of texture much less than perception of objects (**Supplementary Fig. 2**).

Clinical conditions

Crowding is the fruitless combining of features over too large an area. Crowding with abnormally large critical spacing may account for several clinical conditions, including apperceptive agnosia and strabismic amblyopia[20, 21].

In principle it would be similarly fruitless to combine over too small an area, getting only a fraction of the object. This matches some clinical descriptions of simultanagnosia: "It often appeared as if he were looking through a peephole which was too narrow to include the entire stimulus"[22].

Supp. Fig. 1. The uncrowded neighborhoods (white polygons) of two objects: water bottle and magazine. You must fixate within its neighborhood to recognize the object. Fixating outside the uncrowded neighborhoods (e.g., on the little girl's face), you cannot recognize either of these objects. The neighborhood size (white polygon) depends on the local density of the clutter around the object and the similarity of the clutter to the object. The polygon is the measured threshold eccentricity for recognition in eight directions. These thresholds are subjective; the observer knew all along what the object was. Train station, Moscow, Russia, 2006. Photo by K.A. Tillman.

Supp. Fig. 2. A forest. This is mostly texture, with very few recognizable objects. Unlike perception of objects, the perception of texture is little affected by the location of fixation[23]. We suggest that one might define "texture" as what one can see without object recognition. Copyright © Ray K. Metzker, Courtesy Laurence Miller Gallery, New York.

Attention

Is attention involved in crowding and object recognition? It depends on what you mean by “attention”. In William James’s broadest view, when looking at the world, attention means awareness. Most object recognition tasks ask the observer to report a target (or its absence). This report (over many trials) typically communicates information that could only have come from the target, which strongly suggests that the observer was aware of it. (The exceptions are the much-discussed special cases of stimuli that may affect reports without entering consciousness, e.g., blindsight and subliminal priming.) If awareness of the target is attention, then attention is a near-essential part of most object recognition tasks, but this general fact is not related to crowding in particular.

More specifically, there has been great interest in selective attention in visual search and texture segregation. *Selective attention* refers to the observer's ability to filter the visual scene, emphasizing some areas (or things) and ignoring others. Helmholtz said, “A human being cannot attend to more than one object at a time, … [but,] in spite of the vagueness of the broad field of view, the eye is capable of taking in at a rapid glance the main features of the whole surroundings, and of noting immediately the sudden appearances of new objects … to divert our attention to any new or extraordinary phenomenon that may arise out toward the periphery of the field”[24]. A century later, Julesz and Bergen[25] express the same idea: “Thus preattentive vision serves as an ‘early warning system’ by pointing out those loci of texton [i.e. feature] differences that should be attended to”[25, p. 1619]. Treisman’s Feature Integration Theory (FIT) goes on to suppose that observers have a *focused attention* process (i.e., selective attention) that can be directed to any area or object in the visual field, and that only this process can correctly integrate (i.e., combine) features for object recognition[26]. Most of the two thousand or so papers on visual search interpret their results as characterizing the limits of selective attention[27]. In this framework, it is natural to interpret the critical spacing for crowding as being the spatial resolution of attention[28]. That interpretation is parsimonious in attributing the crowding phenomenon to FIT’s mechanism of focused attention.

Wolfe and Bennett[29] describe shapeless “preattentive object files” of “unbound” [not combined] features. This is similar to the “bag of features” idea in machine learning[30], and seems to be a good description of crowding. Indeed, the “attentive” vs. “preattentive” dichotomy seems to correspond to uncrowded vs. crowded vision. In particular, Treisman and Gelade[26] show that finding an R among P’s and Q’s is serial, “requiring focused attention”. Similarly, Julesz & Bergen[25, p. 1621] say that “element-by-element scrutiny, called ‘focal attention’, is required to find the T’s embedded in the L’s.” And Wolfe and Bennett[29] report serial search for a target plus sign, +, consisting of a green vertical and a red horizontal bar among distractor pluses that each consist of a red vertical and a green horizontal bar. **Supplementary Fig. 3** shows minimal versions of these stimuli. In each of the top three rows, fixating on the minus, you can easily identify the isolated target to the left, but you cannot identify the same target on the right, where it is flanked by distractors. This is crowding. While fixating the minus, no amount of willpower will rescue the crowded target. However, moving fixation closer, to the small grey plus, does restore your ability to identify the target. Thus, for this fixed spacing, reducing eccentricity changes vision from “preattentive” to “attentive”, even though the observer can concentrate on the target throughout. The bottom row shows self-crowding. While fixating the minus, you will find that the isolated target ‘10’ on the right is indistinguishable from the isolated distractor ‘S’ on the left. Julesz says, “The ‘S’ and ‘10’ shaped elements … in isolation appear quite different, [but are] preattentively indistinguishable”[25, p. 1626]. These demonstrations are consistent with the finding that target eccentricity strongly affects performance of searches for multi-feature targets[3, 17, 33].

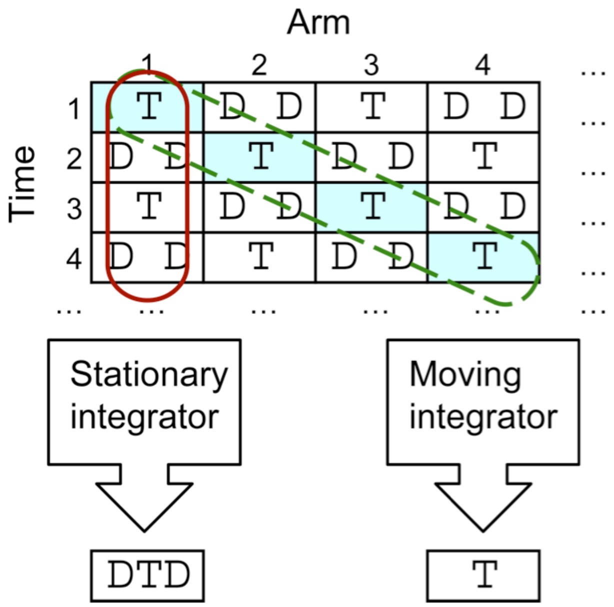
Supp. Fig. 3. Is crowded vision “preattentive”? These are minimal versions of stimuli that are known to require serial search[29, 31, 32]. In the top three rows, fixating the central minus, you will easily identify the isolated target at the left, but you will fail to identify the same now-crowded (middle) target on the right. The bottom row shows self-crowding: Fixating on the minus you will find that the left and right targets are indistinguishable. In every case, you can readily identify the target if you fixate the nearby small grey plus. Thus, reducing eccentricity changes your vision from “preattentive” to “attentive” even though you are concentrating on the target throughout.

One could say that these crowding effects are independent of attention. Or, as above, one could take the critical spacing of crowding to be the resolution (minimum area) of attention, treating the two as one.

While it is clear that directing the observer's attention to the object can enhance recognition, it is not clear that this focused attention is essential (despite FIT's claim that it is). Motter and Simoni recorded eye position and analyzed the trial-by-trial results as the probability of target recognition as a function of spacing (to nearest distractor) as a fraction of target eccentricity[3, 17]. They concluded that each target recognition is limited by the critical spacing of crowding, without invoking selective attention. Similarly, using a target embedded at a known or unknown location in a ring of distractors, all at the same eccentricity, several investigators have looked for and failed to find an effect of selective attention on the critical spacing for crowding[7, 34, 35]. This suggests another interpretation, equally parsimonious: Perhaps crowding may be understood on its own, independent of selective attention. Selective attention does enhance object recognition, but since it does not affect critical spacing, perhaps it is a separate factor in object recognition, independent of crowding. These interpretations differ in taking crowding to be either the resolution of attention or independent of attention, yet they agree in describing crowding as the combining of features over an inappropriately large area.

Non-retinotopic crowding. A popular crowding demo, http://visionlab.harvard.edu/Members/Patrick/CrowdingMovies/, by Cavanagh and Holcombe, is taken by many to be evidence for the attentional-resolution account of crowding. Our view is that the demo shows just what one would expect given the Bex et al.[36] finding that the critical spacing of crowding is unchanged when measured in a moving reference frame. The Cavanagh and Holcombe demo alternately displays a target and flanking distractors[37]. Even though the target and distractors are never displayed simultaneously, the target is crowded unless the alternation rate is very slow (few Hz). They called that flickering linear array of letters an "arm", and arranged 16 such arms radially, like the spokes in a wheel, around the central fixation point. The arms are numbered 1 to 16. All the odd-numbered arms (i.e., every other arm) are in sync, all showing the target at the same time. The even-numbered arms are in sync with each other but out-of-sync with the odd-numbered arms, displaying the distractors when the odd arms display the target. Again, unless the alternation rate is very low, the target is crowded. However, the authors then add an attentional guide, brightening one sector (containing arm *t*) at time *t*, and ask the observer to maintain central fixation while mentally tracking the bright sector as it goes around. The brightened arm is just a target letter, in every frame, and observers are much better able to identify it with the guide than without it. Cavanagh and Holcombe call this "non-retinotopic crowding". This phenomenon needs more than a strictly spatial account, since the result is spatiotemporal, "suggesting [that] the crowding is specific to the flankers, if any, that move with the target and not to the letters that surround each target locally in retinotopic coordinates."

It seems that, without a guide, the observer is using a stationary integrator (continuous oval in **Supplementary Fig. 4**). With a guide, the observer seems to be using a moving integrator (dashed oval). It matters little how many frames the integrator integrates provided it's more than one. It is well known that "translation *per se* across the retina has little effect on temporal summation" of an object[38]. Apparently the moving attentional cue helps the observer select a moving, instead of a stationary, integrator. Since motion *per se* does not seem to affect crowding[36], the cueing here seems to allow the observer to select an uncrowded (moving) representation of the stimulus. A rotating bright bar is a powerful motion stimulus. We would expect it to push the observer towards seeing motion, even if attention were not involved.



Time \ Arm	1	2	3	4	…
1	T	D D	T	D D	…
2	D D	T	D D	T	…
3	T	D D	T	D D	…
4	D D	T	D D	T	…
…	…	…	…	…	…

Supp. Fig. 4. Non-retinotopic crowding. Analysis of the Cavanagh and Holcombe demo.[37] http://visionlab.harvard.edu/Members/Patrick/CrowdingMovies/ ***Their demo (not shown):*** Sixteen radial arms are presented, each with a target letter in the middle, flanked by distractors on each side. The display counterphases targets and distractors. In one frame (presented at odd times 1, 3, …), every odd arm (1, 3, …) shows only the target without distractors, while the even arms (2, 4, …) show only the distractors without the targets. In the other frame (presented at even times), even arms show targets while odd arms show the distractors. Thus, when a guide (a brightened radial sector) moves from arm to arm in phase with the alternation (arm 1 at time 1, arm 2 at time 2, …), it contains only the target letter and no distractors. Subjects fixate the center of the circular array and report the orientation of the target letter. When attention is directed to one fixed location, there is substantial interference from the distractors. However, when following the guide, crowding is much reduced, suggesting that distractors only crowd the target if they remain with it over time. ***Our space-time diagram:*** The table shows the content of each arm at each time, which is either the target letter (T) or the distractors (D D). The shading represents the attentional cue: the brightening of arm *t* at time *t*.

Feature pooling

It seems that combining features to recognize objects carries the risk of crowding. Above, we reviewed evidence that the critical spacing is the spatial extent over which features are combined, but we have said little about what goes on within this combining area. Some of the complaints about crowding — especially the impaired judgment of position and shape — seem to stem from uncertainty (confusion) about feature position. One may complain about this uncertainty, but we all benefit from the positional invariance of recognition. Although the relative positions of features vary among fonts and handwritings, for a letter to be read, it must be assigned to the same category, e.g., "a", regardless of its rendering.

How are features combined? Three lines of investigation (psychophysics, physiology, and engineering) converge on *maximum pooling* as a key step. In maximum pooling, many feature detectors with similar receptive fields, differing only in position, all respond to the stimulus independently, but only the maximum detector response, regardless of detector position, is passed on. This immediate loss of precision of feature position is an important aspect of psychophysical and physiological models and engineering solutions for object recognition[39-41].

Psychophysically, when attempting to identify something, human observers act as though they are always considering many possible positions, like the ideal observer for an uncertain signal, which does maximum pooling[42-45]. **Supplementary Fig. 5** allows you to witness this vagueness of feature position. Physiologically, in the primary visual cortex, complex cell responses are position invariant and do not summate, consistent with maximum pooling[46, 47]. In practical engineering, some of the most successful machine classifiers of handwritten digits (and other objects) use maximum pooling to tolerate "deformations and shifts in position"[48-50]. In all these cases, invariance of object recognition is achieved by maximum pooling, which results in uncertainty of feature location, which, in turn, limits the precision of judgments of object position and shape.

SUPPLEMENTARY SOURCES

Figure 4 images. The Elvis image is used by permission, Elvis Presley Enterprises, Inc. The Arnold image is publicly available from the State of California website. **Figure 5 images.** The gratings were created in MATLAB. The letters are in the Courier font. The animal silhouettes are in our Animals font, which is available for research purposes. The men, women, and telephone signs are from aiga.com[51]. The ladder is licensed from and copyright Stockbyte. The rocking chair is copyright 2008 Jupiter Images Corporation. In the following credits, we use the convention (Photographer/Name of collection/Source). The stool (C Squared Studios /Photodisc/GettyImages), pretzel (Steve Wisbauer/Photodisc/Getty Images), hamburger (Ryan McVay/Photodisc/Getty Images), and pizza (Raimund Koch/Riser/Getty Images) are from Getty Images. The house is courtesy of Snodgrass and Vanderwart[52]. The image of Gandhi is copyright Vithalbhai Jhaveri/GandhiServe.

SUPPLEMENTARY METHODS

Figure 9 methods. Taylor[53] tested thousands of normal students in 1st grade through college. We plot one point per grade, 1-12, plus college. Reading speed (vertical scale): Subjects read age-appropriate paragraphs. Carver corrected the speeds for text difficulty [54, Table 2.1]. Span (horizontal scale): Taylor measured eye movements. We plot Taylor's "span of recognition", the average length of forward saccades (the product of words per saccade, from Taylor, and characters per word, for text at each difficulty level, from Carver's Eq. 2.2). **Kwon et al.**[55] tested normal 3rd, 5th, and 7th graders, and adults (4 points). Reading speed: Subjects read sentences displayed one at a time. Span: Kwon et al. measured 'visual span profiles', which trace out the subject's accuracy for identifying a triplet of random letters as a function of position in the visual field. We plot the number of letter positions in the visual span profile (for 1 deg letters) for which the triplet accuracy is at least 80%. **Valdois et al.**[56] tested two dyslexic subjects (in the 6th and 7th grades), classified as a surface dyslexic (△) and a phonological dyslexic (◺), and age-matched controls (▲). Reading speed: Valdois provided reading speeds for ordinary text. Span: They measured accuracy versus letter position for reporting a string of 5 briefly-presented random letters (their Fig. 1). We plot the number of positions (out of 5) at which the subject got at least 80% correct. **Martelli et al.**[57] tested normal and dyslexic 6th graders. Normal adult data were provided separately. Reading speed: The normal and dyslexic children read ordinary text printed on paper. The adults read 8-letter nouns in rapid serial visual presentation. Span: They measured the critical spacing required to identify the central letter in a triplet of three random letters with 90% accuracy as a function of eccentricity. They then calculated Bouma's factor b (proportionality constant between critical spacing and eccentricity). We plot the uncrowded span $u = 1+2/b$ [1]. **Prado et al.**[58] tested dyslexics and age-matched controls. (They did not report the students' grade level, but average age was 11 years, which is typical for the 6th grade.) Reading speed: They measured eye movements as subjects read short passages. We plot rate as the number of words in the passage divided by the product of the total number of fixations and the mean fixation duration (their Table 2). Span: We plot the average number of letters reported correctly from a string of 5 briefly presented letters (their Table 1).

For reviews of reading, see refs [59-63]. *For further reading, see refs* [64-67].

Supp. Fig. 5. Experience the vagueness of feature position predicted by maximum pooling. Viewing the page from 17 cm away (though distance hardly matters), fixate the plus. The letter (3.3° at an eccentricity of 46°) is too small to recognize, and looks like "a jumble of lines or an unorganized heap of marks"[68]. Optical blur is noticeable, but does not prevent you from seeing the lines. The feature position errors are so large that you see only a jumble of floating features. One observer said, "I see something that appears to be composed of straight lines about half an inch high. Could be a drawing. Could be a letter or letters. I cannot see clearly what it is. At the moment it looks like a capital Y, but it's indefinable. The lines are not precise. They appear to be shimmering, fading in and out. Very unstable figure". Such confusion of feature position is predicted by maximum pooling. We think that maximum pooling not only contributes to crowding, but also limits acuity, as shown here. There is no crowding here because there is only one simple letter, which consists of only one part, though it has many features. The uncertainty of feature position seems to be a fixed fraction of the combining area. For example, the just noticeable difference in position of a grating patch (not shown) is independent of the spatial frequency of the grating and is about 4% of whichever is larger: the extent of the grating or 50% of the eccentricity[69]. We think that this uncertainty contributes to crowding, but its spatial extent is much too small to be the main cause of crowding.

ACKNOWLEDGMENTS

Thanks to Lyuba Azbel, Diana Balmori, Nate Blanco, Marie-line Bosse, Herman Bouma, Marisa Carrasco, Patrick Cavanagh, Rama Chakravarthy, Susana Chung, Peter Enneson, Hannes Famira, Jeremy Freeman, Hugh Hardy, Lloyd Kaufman, MiYoung Kwon, Michael Landy, Yann LeCun, Dennis Levi, Larry Maloney, Flavia Mancini, Marialuisa Martelli, Tony Movshon, Cesar Pelli, Rafael Pelli, Elizabeth Segal, Josh Solomon, Lothar Spillman, Hans Strasburger, Jordan Suchow, Sylviane Valdois, David Williams, Jeremy Wolfe, and, especially, Bart Farell, Rob Fergus, David Heeger, Brad Motter, Jamie Radner, Bob Shapley and Nick Wade for helpful comments. Jamie Radner, through his comments on so many of our 160 drafts, is de facto third author of the abstract and introduction. Nick Wade told us about acuity testing in ancient times[70]. Supported by US National Eye Institute grant R01-EY04432 to Denis Pelli.

1. Pelli, D.G., Tillman, K.A., Freeman, J., Su, M., Berger, T.D. & Majaj, N.J. Crowding and eccentricity determine reading rate. *Journal of Vision* **7(2)20**, 1-36 (2007). [Pubmed]
2. Larsson, J. & Heeger, D.J. Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci* **26**, 13128-13142 (2006).
3. Motter, B.C. & Simoni, D.A. The roles of cortical image separation and size in active visual search performance. *Journal of Vision* **7(2)6**, 1-15 (2007).
4. Levi, D.M., Klein, S.A. & Aitsebaomo, A.P. Vernier acuity, crowding and cortical magnification. *Vision Res* **25**, 963-977 (1985).
5. Kooi, F.L., Toet, A., Tripathy, S.P. & Levi, D.M. The effect of similarity and duration on spatial interaction in peripheral vision. *Spat Vis* **8**, 255-279 (1994).
6. Andriessen, J.J. & Bouma, H. Eccentric vision: adverse interactions between line segments. *Vision Res* **16**, 71-78 (1976).
7. Wilkinson, F., Wilson, H.R. & Ellemberg, D. Lateral interactions in peripherally viewed texture arrays. *J Opt Soc Am A* **14**, 2057-2068 (1997).
8. Chung, S.T.L., Li, R.W. & Levi, D.M. Crowding between first- and second-order letter stimuli in normal foveal and peripheral vision. *Journal of Vision* **7(2)10**, 1-13 (2007).
9. Felisbert, F.M., Solomon, J.A. & Morgan, M.J. The role of target salience in crowding. *Perception* **34**, 823-833 (2005).
10. Farell, B. & Pelli, D.G. Psychophysical methods, or how to measure threshold, and why. In *A Practical Guide to Vision Research* (ed. J.G. Robson & R.H.S. Carpenter) 129-136 (Oxford University Press, New York, 1999).
11. Levi, D.M., Hariharan, S. & Klein, S.A. Suppressive and facilitatory spatial interactions in peripheral vision: peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision* **2(2)3**, 167-177 (2002).
12. Livne, T. & Sagi, D. Configuration influence on crowding. *Journal of Vision* **7(2)4**, 1-12 (2007).
13. Huckauf, A. & Nazir, T.A. How odgcrnwi becomes crowding: Stimulus-specific learning reduces crowding. *Journal of Vision* **7(2)18**, 1-12 (2007).
14. Chung, S.T.L. Learning to identify crowded letters: Does it improve reading speed? *Vision Res* **47**, 3150-3159 (2007).
15. Engel, F.L. Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Res* **17**, 95-108 (1977).
16. Handford, M. *Where's Waldo?* (Little Brown, Boston, 1987).
17. Vlaskamp, B.N., Over, E.A. & Hooge, I.T. Saccadic search performance: the effect of element spacing. *Exp Brain Res* **167**, 246-259 (2005).
18. Metzger, W. *Laws of seeing* (MIT Press, Cambridge, MA, 2006).
19. Lettvin, J.Y. On seeing sidelong. *N. Y. Acad. Sci.* **16**, 10-20 (1976).
20. Levi, D.M., Song, S. & Pelli, D.G. Amblyopic reading is crowded. *Journal of Vision* **7(2)21**, 1-17 (2007). [Pubmed]
21. Song, S., Levi, D.M. & Pelli, D.G. Can “equivalent eccentricity” account for two measures of performance in the amblyopic visual system? *ARVO 2008 Annual Meeting* April 27-May 1 in Fort Lauderdale, Florida (2008).
22. Trobe, J.R. & Bauer, R.M. Seeing but not recognizing. *Surv Ophthalmol* **30**, 328-336 (1986).
23. Braun, J. & Sagi, D. Texture-based tasks are little affected by second tasks requiring peripheral or central attentive fixation. *Perception* **20**, 483-500 (1991).
24. Helmholtz, H.v. *Treatise on Physiological Optics (Dover ed.), vol. 1, sec. 10, pg. 93 and vol. 3, sec. 28, pg. 205* (Dover Publications, Mineola, NY, 1867/2005).
25. Julesz, B. & Bergen, J.R. Textons, the fundamental elements in preattentive vision and perception of textures. *Bell System Technical Journal* **62**, 1619-1645 (1983).
26. Treisman, A.M. & Gelade, G. A feature-integration theory of attention. *Cognit Psychol* **12**, 97-136 (1980).
27. Wolfe, J.M. Visual search. In *Attention* (ed. H. Pashler) 13-73 (Psychology Press, New York, 1998).
28. Intriligator, J. & Cavanagh, P. The spatial resolution of visual attention. *Cognit Psychol* **43**, 171-216 (2001).
29. Wolfe, J.M. & Bennett, S.C. Preattentive object files: shapeless bundles of basic features. *Vision Res* **37**, 25-43 (1997).
30. Breiman, L. Bagging predictors. *Machine Learning* **24**, 123-140 (1996).
31. Treisman, A. Features and objects: the fourteenth Bartlett memorial lecture. *Q J Exp Psychol A* **40**, 201-237 (1988).
32. Julesz, B. A brief outline of the texton theory of human vision. *Trends in Neurosciences* **7**, 41-45 (1984).
33. Carrasco, M., Evert, D.L., Chang, I. & Katz, S.M. The eccentricity effect: target eccentricity affects performance on conjunction searches. *Percept Psychophys* **57**, 1241-1261 (1995).
34. Scolari, M., Kohnen, A., Barton, B. & Awh, E. Spatial attention, preview, and popout: which factors influence critical spacing in crowded displays? *Journal of Vision* **7(2)7**, 7 1-23 (2007).
35. Nazir, T.A. Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Res* **32**, 771-777 (1992).
36. Bex, P.J., Dakin, S.C. & Simmers, A.J. The shape and size of crowding for moving targets. *Vision Res* **43**, 2895-2904 (2003).
37. Cavanagh, P. & Holcombe, A.O. Non-retinotopic crowding. *Journal of Vision* **7(9)**, 338a (2007).
38. Burr, D.C. Temporal summation of moving images by the human visual system. *Proc R Soc Lond [Biol]* **211**, 321-339 (1981).
39. Baldassi, S., Megna, N. & Burr, D.C. Visual clutter causes high-magnitude errors. *PLoS Biol* **4**, e56 (2006).

40. Movshon, J.A., Thompson, I.D. & Tolhurst, D.J. Receptive field organization of complex cells in the cat's striate cortex. *J Physiol (Lond)* **283**, 79-99 (1978).
41. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86**, 2278-2324 (1998).
42. Pelli, D.G. Uncertainty explains many aspects of visual contrast detection and discrimination. *J Opt Soc Am A* **2**, 1508-1532 (1985). [Pubmed]
43. Tjan, B.S. & Nandy, A.S. Classification images with uncertainty. *Journal of Vision* **6(4)8**, 387-413 (2006).
44. Shaw, M.L. Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology* **14**, 353-409 (1982).
45. Hess, R.F., McIlhagga, W. & Field, D.J. Contour integration in strabismic amblyopia: the sufficiency of an explanation based on positional uncertainty. *Vision Res* **37**, 3145-3161 (1997).
46. Hubel, D.H. & Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* **160**, 106-154 (1962).
47. Lampl, I., Ferster, D., Poggio, T. & Riesenhuber, M. Intracellular measurements of spatial integration and the MAX operation in complex cells of the cat primary visual cortex. *J Neurophysiol* **92**, 2704-2713 (2004).
48. Fukushima, K. & Miyake, S. Neocognitron - A new algorithm for pattern-recognition tolerant of deformations and shifts in position. *Pattern Recognition* **15**, 455-469 (1982).
49. Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M. & Poggio, T. Robust object recognition with cortex-like mechanisms. *IEEE Trans Pattern Anal Mach Intell* **29**, 411-426 (2007).
50. Ranzato, M.A., Huang, F.-J., Boureau, Y.-L. & LeCun, Y. Unsupervised learning of invariant feature hierarchies with applications to object recognition. *Proc. Computer Vision and Pattern Recognition Conference (CVPR'07)*, 1-8 (2007).
51. AIGA. *Symbol signs* (Hastings House, New York, 1981).
52. Snodgrass, J.G. & Vanderwart, M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J Exp Psychol [Hum Learn]* **6**, 174-215 (1980).
53. Taylor, S.E. Eye movements in reading: Facts and fallacies. *American Educational Research Journal* **2(4)**, 187-202 (1965).
54. Carver, R.P. *Reading rate: a review of research and theory* (Academic Press, San Diego, 1990).
55. Kwon, M., Legge, G.E. & Dubbels, B.R. Developmental changes in the visual span for reading. *Vision Res* **47**, 2889-2900 (2007).
56. Valdois, S., Bosse, M.-L., Ans, B., Carbonnel, S., Zorman, M., David, D. & Pellat, J. Phonological and visual processing deficits can dissociate in developmental dyslexia: Evidence from two case studies. *Reading and Writing* **16**, 541-572 (2003).
57. Martelli, M., Di Filippo, G., Spinelli, D. & Zoccolotti, P. Crowding, reading, and developmental dyslexia. *Perception* **35, supp**, 174 (2006).
58. Prado, C., Dubois, M. & Valdois, S. The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span. *Vision Res* **47**, 2521-2530 (2007).
59. Huey, E.B. *The psychology and pedagogy of reading* (Macmillan, New York, 1908).
60. Snowling, M.J. From language to reading and dyslexia. *Dyslexia* **7**, 37-46 (2001).
61. Legge, G.E. *Psychophysics of reading in normal and low vision* (Erlbaum, Mahwah, NJ, 2007).
62. Pelli, D.G. & Tillman, K.A. Parts, wholes, and context in reading: a triple dissociation. *PLoS ONE* **2**, e680 (2007). [Pubmed]
63. Hutzler, F. & Wimmer, H. Eye movements of dyslexic children when reading in a regular orthography. *Brain Lang* **89**, 235-242 (2004).
64. Plato. *Republic, 2, line 368*^{d} (Hackett Publishing Co., Indianapolis, 1992).
65. Quine, W.V.O. Two dogmas of empiricism. *Philosophical Review* **60**, 20-43 (1951). [Reprinted in Quine, W. V. O. *From a logical point of view*. (Harvard University Press, Cambridge, MA, 1980)]
66. Rey, A. L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives de Psychologie* **28**, 286-340 (1941).
67. Faye, E. *Clinical low vision* (Little Brown, Boston, 1984).
68. Zigler, M.J., Cook, B., Miller, D. & Wemple, L. The perception of form in peripheral vision. *American Journal of Psychology* **42**, 246-259 (1930).
69. Levi, D.M. & Tripathy, S.P. Localization of a peripheral patch: the role of blur and spatial frequency. *Vision Res* **36**, 3785-3803 (1996).
70. Wade, N.J. Image, eye, and retina. *J Opt Soc Am A Opt Image Sci Vis* **24**, 1229-1249 (2007).