Published in final edited form as:
*Nat Neurosci*. 2008 October ; 11(10): 1129–1135. doi:10.1038/nn.2187.

# The uncrowded window of object recognition

**Denis G Pelli** and **Katharine A Tillman**
Department of Psychology and Center for Neural Science, New York University, 6 Washington Place, New York, New York 10003, USA.

## Abstract

It is now emerging that vision is usually limited by object spacing rather than size. The visual system recognizes an object by detecting and then combining its features. 'Crowding' occurs when objects are too close together and features from several objects are combined into a jumbled percept. Here, we review the explosion of studies on crowding—in grating discrimination, letter and face recognition, visual search, selective attention, and reading—and find a universal principle, the Bouma law. The critical spacing required to prevent crowding is equal for all objects, although the effect is weaker between dissimilar objects. Furthermore, critical spacing at the cortex is independent of object position, and critical spacing at the visual field is proportional to object distance from fixation. The region where object spacing exceeds critical spacing is the 'uncrowded window'. Observers cannot recognize objects outside of this window and its size limits the speed of reading and search.

Object recognition means calling a chair a chair, despite variations in style, viewpoint, rendering and surrounding clutter. Crowding is a breakdown of object recognition.

Let us begin by sketching a popular two-step model of object recognition: feature detection and combination. Features are components of images that are detected independently[1–4]. They are typically simple and nonoverlapping. The first step in object recognition is feature detection[4]. Each neuron in the primary visual cortex responds when a feature matches its receptive field. Only the features that drive neurons hard enough are detected[5]. In the second step, the brain combines some of the detected features to recognize the object. This combining step (including 'integration', 'binding', 'segmentation', 'pooling', 'grouping', 'contour integration' and 'selective attention') is still mysterious[3,4,6–11].

Some objects are recognized through a single combining of features over the whole object, whereas other objects require separate combining over each of several regions of the object[12–14]. These distinct regions define object parts. In an object with multiple parts, each part must be recognized before they are all joined together.

The best evidence that features are indivisible elements that we detect and combine is that, even with practice, people combine information across features much less well than within a feature. Searching for a conjunction of several features is usually much harder than searching for a single feature[3]. Despite reading a billion letters over a lifetime, people still recognize letters inefficiently, by detecting and combining many simple features rather than by detecting each letter as a whole[4,15]. Crowding is inappropriate feature combination that spoils object recognition (reviewed in refs. [16,17]).

---

Correspondence should be addressed to D.G.P. (denis.pelli@nyu.edu).

This is an empirical review of crowding in object recognition. Science, in its many styles, creates theory to bind facts into an intelligible whole. This whole, as W.V.O. Quine noted, is a continuum from fact to theory. Broad empirical generalizations, such as those we present here, lie near one end of the continuum, with full explanatory models being present at the other end. Unlike a mature field such as physics, object recognition is an immature topic with only tentative theories, as scientists are still describing the empirical phenomena. A review of the “scattered and diverse” theoretical models of crowding in object recognition finds “a growing consensus” for the two-step account of feature detection and combination[16]. That account does not specify how the crucial combination happens and mostly serves to provide a vocabulary for describing results. This empirical review passes over the details of the diverse models to provide a broad survey of the underlying results, which we find notably consistent. We boil the results down as far as we can, achieving a short synthesis that we call the Bouma law. It binds together most of the facts on crowding and seems to be a useful step toward the computational model of recognition that we all yearn for.

This empirical review includes visual demonstrations that allow the reader to experience the phenomena. The bars in the ‘A in chaff’ demonstration (Fig. 1) represent elementary features. When you look at the demonstration, your brain detects the features and combines them to categorize the letter as A. We cannot yet explain how this process works, but we can easily break it. Fix your eyes on the red minus, far from the A, and the extra features (chaff) make it impossible to recognize the A. When you fixate this far from the A, your brain combines features over too large an area around the A, failing to isolate the relevant features of the A from the nearby junk, and comes up with a jumbled percept instead of a letter. This is crowding. Some well-known illusions are delicate, strongly affected by expectation and only work once. Unlike them, crowding is robust. No matter how many times you move your eyes back and forth from plus to minus, the A quickly comes and slowly goes away every time.

Crowding, unlike overlap masking (ordinary masking by nearby objects that overlap the target), never makes the target disappear[17]. Crowding impairs our ability to identify, count and locate objects, but does not affect detection (Fig. 2). As you can see, the jumbled percept produced by crowding looks like inappropriate combining rather than a failure to detect. The notion that crowding is a breakdown of the second step of object recognition, after feature detection, is consistent with experiments showing that crowding can knock out the observer’s ability to judge target orientation while sparing (or largely sparing) the orientation-specific aftereffect of adapting to that target[18,19]. Finding that we still adapt to stimuli that we cannot identify is evidence for two steps in object recognition, one (feature detection) that is susceptible to adaptation followed by another (feature combination) that is susceptible to crowding.

Crowding is usually specified by the observer’s ‘critical spacing’. Critical spacing is how far (measured center to center) the flanking objects (‘flankers’) must be from the target to allow unimpaired perception of the target. Critical spacing grows in proportion to eccentricity, the distance of the target object from fixation[20]. It has been suggested that critical spacing may reflect the spatial resolution (minimum area) of visual attention, but this is controversial[11,21,22].

Distinguishing crowding from overlap masking is easy, as the critical spacing of crowding is proportional to eccentricity, whereas that of overlap masking is independent of eccentricity[17]. Therefore, crowding dominates in the periphery and overlap masking dominates centrally[17].

Tilted flankers have a long-range effect on the perceived tilt of a foveal target. Unlike crowding and overlap masking, this ‘stochastic recalibration’ affects the orientation threshold, but not the contrast threshold, for orientation discrimination[23].

## The Bouma law

Practically every paper on crowding reports critical spacing. Our story, here, is that (despite the great diversity of models) the results all boil down to a simple law, a generalization of an observation that Herman Bouma reported in 1970, that the critical spacing for identification of small letters is roughly half the eccentricity[20]. We take this observation to its most general form, which we call the Bouma law: for an object that can be identified in isolation, our ability to identify it among similar objects depends solely on the ratio of the object spacing to the observer's critical spacing at that location. The object is crowded whenever the ratio is less than one. For each observer, the critical spacing is independent of what the object is and depends only on where the object is in the visual field and the direction from target object to flanker object. The broad empirical support for this law is unexpected because object recognition is usually assumed to be limited by size, not spacing.

Most studies of crowding have used letters and words as stimuli. (However, a recent special issue of the *Journal of Vision* includes more than twenty articles on crowding, using a wide variety of stimuli. http://www.journalofvision.org/7/2/) Figure 3 demonstrates the critical spacing of the letters in a word[24]. If you try to identify the middle letter in the word 'are', it is easy when you fixate near the word and becomes hard when you fixate far away. This is because, when fixation is too far away, the whole word falls within one critical spacing and features from all of the letters are jumbled together. Some objects, such as words, have parts. The parts of an object crowd each other when they are closer than the critical spacing. Faces, like words, are recognized only if the visual system can isolate their parts: eyes, nose and mouth[14]. Thus, we cannot recognize a face unless we look at or near it (Fig. 4).

The critical spacing is universal, independent of object and size (Fig. 5). The threshold eccentricity for recognition is the same for all objects with the same spacing, even when the objects are as diverse as gratings, letters, animals and furniture. Similarly, the critical spacing of crowding is unaffected by equal motion of the target and flankers[25]. Across different tasks, including discrimination of size, hue, saturation and orientation, the amplitude (maximum threshold elevation) of crowding varies, but the spatial extent of crowding is practically the same[26]. 'Second-order' letters (painted with texture) are more susceptible to crowding than 'first-order' letters (painted with homogeneous ink), but the spatial extent of crowding is the same[27].

The generality of the Bouma law suggests that the critical spacing of crowding is a fundamental parameter of human vision. It is proportional to the distance from fixation (Fig. 6), and depends solely on position and direction in the visual field[17,20]. This proportionality matches the organization of the visual 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) that is independent of eccentricity (see Supplementary Discussion online).

## Size or spacing?

The idea that spacing limits object recognition could not be simpler, but it has been very hard to accept because it displaces a firmly held belief that visibility is limited by size (acuity), not spacing (crowding). For example, an expert reviewer of a related article complained that "the presentation in terms of spacing [instead of size] … made it quite hard for me to understand".

When we view a scene from farther away, both size and spacing decrease. Viewing distance, *per se*, does not matter. What matters is the stimulus at the retina. Some visual tasks are limited by size. The Egyptians (5,000 years ago) and many since have assessed acuity of vision by the ability to distinguish the double star Alcor/Mizar in Ursa Major. Today, to measure a size threshold (acuity) that characterizes a person's vision, we ask the observer to identify a simple object, usually a letter. This measure is unaffected by crowding if done foveally, where critical spacing is only a few minutes of arc, or anywhere on a blank field. Measuring acuity is useful, especially in selecting the best optical correction. However, outside of the optometrist's office, most of us are well corrected (20/20) and, provided that there is enough contrast[28], our ability to see is more limited by object spacing than by size. We can see a bird in the sky without crowding, but most of our visual world is cluttered, and each object that we identify must be isolated from the clutter. When an object is not isolated, it is crowded, and we cannot recognize it. Isolation depends on spacing and not size. To escape crowding, the object spacing must exceed the observer's critical spacing at that location in the observer's visual field (that is, 6 mm at V1).

Critical spacing has profound effects on everyday life. Consider reading. It has long been known that reading consists of a series of eye fixations, 4 per second, rather than a continuous sweep of the eyes across the text[29]. Reading speed is independent of text size over a large 6:1 range, but drops precipitously for sufficiently small text. From ancient to modern times, this has been taken to be a size limit (acuity). Plato complained that he was asked "to read small letters from a distance". This statement shows that he both understood the concept of acuity and thought that it limited reading. In 1985, we said that, "the fairly rapid decline in reading rate for characters smaller than 0.3° is undoubtedly associated with acuity limitations"[30], but we were wrong. Reading speed depends on letter spacing and not size. Measuring with two texts, one widely and one normally spaced, at various viewing distances, it is found that reading speed drops at a particular letter spacing (in deg), independent of letter size[31]. Typographers routinely increase 'tracking' (spacing) to maintain the legibility of text when it is made smaller.

## Spatial extent of crowding

The invariance of critical spacing demonstrated here (Fig. 5) is found when the target and flankers have similar features (for example, black letters flanking a black letter target). These typical cases produce maximum crowding. Flankers that have features that are different than those of the target (for example, white letters flanking a black letter target on a gray background) produce much less crowding or none at all. 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. It seems that the reported reduction of critical spacing may be an artifact of defining critical spacing by a performance criterion. Compared with the effect of target-like flankers, dissimilar flankers may simply have a weaker effect over the same spatial extent (see Supplementary Discussion for more on similarity and effects of salience, grouping, and observer practice).

At present, the simplest account is that the spatial extent of crowding for any given location and direction is independent of the particular target and flanker. That conclusion is tentative because the majority of published studies have not disentangled the amplitude and extent of crowding, but it is supported by all the studies that have done a two-parameter analysis. For the rest of this review, we revert to using 'critical spacing', asking the reader to bear in mind that special cases demand a two-parameter (amplitude and extent) characterization of crowding.

## The uncrowded window

Most of our visual field is crowded most of the time, sparing only a central uncrowded window. This window and the limitation it places on recognition are especially clear in the case of reading. To read text, we must identify letters. The rate at which we read depends on how many letters we take in on each fixation (Fig. 7), which is limited by crowding. The spacing of letters in text is uniform, but the observer's critical spacing increases with distance from fixation. Beyond some eccentricity, the reader's critical spacing exceeds the spacing of the text and the letters crowd each other, spoiling recognition. Peripheral vision, beyond that eccentricity, is crowded. Central vision, within that eccentricity, is uncrowded: the uncrowded window. Inside of the window, letters are uncrowded and we can read them. Outside of the window, letters are crowded and we cannot. To read the letters that now lie outside of the window, we must move our eyes to bring our window to those letters. The number of character positions in a line of text that fit inside the uncrowded window is the uncrowded span[28]. Incidentally, note that letters at the ends of words are much less crowded[24] and have a larger uncrowded window.

Figure 8 demonstrates the uncrowded window by simulating crowding in the periphery. The corruptions outside the uncrowded window are undetectable when you fixate on the center of the window.

It seems that each observer's critical spacing for crowding is the same for all objects. Together, the observer's critical spacing and the spacing of the viewed objects determine the size of the uncrowded window. Inside of the window, we can recognize objects, and outside of it, we cannot[32]. When the spacing is uniform, as in text, then the window will be central, where the critical spacing is smallest. When spacing is not uniform, the window need not be central, and there may be more than one. Many have suggested that a central window (also known as the span of apprehension, visual span, visual attention span, area of focal attention, conspicuity area, association field or number of elements processed per fixation) limits reading or search[10,20,29,32–38] (reviewed in refs. [28,39]), but they usually assumed that the window size is independent of object spacing. Often it has been supposed that the window size is limited by letter or object size (acuity), or sometimes by attention. Until recently, only Woodworth[33] and Bouma[34] claimed that the size of the window is set by spacing (crowding). They made good cases against acuity, but failed to convince their colleagues. Subsequent papers cited them, but persisted in assuming that the window is limited by acuity. However, recent detailed studies of search and reading validate the original claim, showing that the window is where the object spacing exceeds the critical spacing of crowding[28,32,40].

Following the success of the uncrowded window idea in explaining the reading speed of normal adults[28,39], one wonders whether it can help to explain why children and dyslexics read more slowly. Developmental dyslexia is now generally thought to be primarily a phonological deficit[41], but there is evidence that dyslexics have increased crowding[42].

We plotted data (Fig. 9) from all the studies for which we could estimate reading speed as a function of the number of characters in the uncrowded window. For all the normal readers, including both children and adults, reading speed was fairly well predicted (with no degrees of freedom) by the product of span and the standard 4-Hz rate of fixations. The large increase in uncrowded span during childhood contrasts with the small effect of practice on critical spacing (and thus uncrowded span) in adults. This warrants further investigation. Most of the dyslexics had smaller spans than age-matched controls, but they read much more slowly than is predicted by their span: They were all well below the normal line, reading at less than half of their span-predicted speed. This finding is contrary to the hypothesis that most cases of dyslexia are arrested development, with performance similar to that of younger normal individuals matched for reading level[41]. These data indicate that something else (for example, a phonological deficit or longer fixations[43]) must account for the rest of the dyslexic impairment[35]. However, the most notable result is the accuracy of the reading speed prediction for normal readers. The normal development of reading speed seems to be mediated entirely by the uncrowded span[35,43,44].

Crowding also limits the speed of visual search. For searches in the real world (or in *Where's Waldo?*; Supplementary Discussion), where similarity and spacing are variable, it is helpful to trace out an uncrowded neighborhood relative to the target, the area in which you must fixate to see the target without crowding. This is the inverse of the uncrowded window, which is defined relative to the observer's fixation point. The size of the uncrowded neighborhood limits search rate (Supplementary Fig. 1 online).

# DISCUSSION

## Peripheral vision and texture

Typically, only a small portion of the visual field falls in the uncrowded window. Most of our visual field is peripheral and crowded and cannot recognize objects. If we cannot recognize things in this part of our vision, what do we see? We see stuff (unnamed texture) and perceive space (the shape of the scene we are in). With an effort, observers can name and describe texture, but this rarely happens. Texture includes variations of color, depth and motion[8]. Many of the cues to depth (binocular disparity, motion parallax, scale gradients and shape from shading) seem to be immune to crowding. A sense of space is particularly important for mobility, which is greatly impaired by tunnel vision of 20 deg or less[45]. Location of fixation affects perception of texture much less than it affects perception of objects (Supplementary Fig. 2 online).

The Rey Complex Figure Test is widely used to assess the ability of neurological patients to copy a line drawing. Surprisingly, normal observers copying the figure with just their peripheral vision produce drawings that are similar to those produced with free viewing by patients with apperceptive agnosia, a type of object blindness (Fig. 10). These drawings suggest that the crowded peripheral vision of normal observers may be a good model for the central vision of these object-blind patients. Clinically, the excessive feature combination of crowding may account for apperceptive agnosia and strabismic amblyopia[31], whereas insufficient feature combination may correspond to simultanagnosia (see Supplementary Discussion).

## Visual dichotomy

The uncrowded window and crowded surrounding field follow a long tradition of visual dichotomies: direct versus indirect, foveal versus peripheral, focal versus ambient, with versus without scrutiny, attentive versus pre-attentive, sustained versus transient, ventral versus dorsal, what versus where and perception versus action. This history of dichotomies distinguishes two kinds of vision. The first is typically central, acute, serial and 'conscious', and it recognizes and names objects. The second is typically peripheral, indistinct (blurry, vague, fuzzy, uncertain, confused and jumbled), parallel and 'unconscious', and it does not recognize or name objects, but helps to guide movement. Technically, these dichotomies are distinct, but in practice they have been used more or less interchangeably, following the fashions of vision science.

Crowding may be responsible for some of these dichotomies (a very close correspondence between pre-attentive and crowded vision can be seen in Supplementary Fig. 3 online). Similarly, there is a strong association between crowded and unconscious vision. One sign of conscious awareness is reporting what we see, which is much harder when object recognition fails, leaving only unnamed texture. The failure of crowded viewing to produce object names may be why peripheral vision is so rarely described in science and literature. Acuity and other measures have been graphed as a function of eccentricity, but there are very few published descriptions of the everyday experience of crowded viewing (see Supplementary Discussion).

In everyday life, most of the things that we recognize are susceptible to crowding (by surrounding clutter) or self-crowding (among the parts). We see these things through a keyhole, the uncrowded window. Reading and searching speeds are proportional to the size of this window. We talk about and remember the things that we identify. The rest of our visual field is crowded, does not recognize or name things, and is hardly ever mentioned, but it lets us perceive space.

### Attention

Attention is one of the most-studied topics in psychology (PsycInfo lists nearly 4,000 peer-reviewed articles on visual attention). If we take attention to be awareness of the target, it is clearly necessary for most object recognition tasks. Our purpose here is not to review attention as a general factor in object recognition, but rather to focus on a narrower question: the possible connection between attention and crowding (see Supplementary Discussion).

Selective attention is the filtering of a scene by the observer to emphasize a target. It is natural to interpret the critical spacing of crowding as the spatial resolution of selective attention[11]. Although there is evidence supporting this view (see Supplementary Discussion and Supplementary Figs. 3 and 4 online), an alternative interpretation sees crowding and selective attention as independent phenomena that affect object recognition separately. This alternative view is possible because selective attention enhances object recognition without affecting the critical spacing. These two interpretations differ in taking crowding to be either the resolution of attention or independent of attention, yet they agree in supposing that the critical spacing defines the area over which features are combined. Many investigators are trying to establish a link between crowding and attention.

Our ultimate goal is to achieve a computational model of the object recognition process. So far we have said only that features beyond the critical spacing for crowding are ignored. What happens inside of the critical spacing? How are features combined? Psychophysics, physiology and engineering all suggest that the first step is a reduction in the spatial precision of the internal representation of the stimulus through feature pooling (see Supplementary Discussion; a demonstration allows the reader to witness this imprecision, Supplementary Fig. 5 online).

In this empirical review, the various studies of crowding all merge to tell a single story. Although the roles of learning, development, similarity and selective attention in crowding are still being worked out, there is a growing consensus that crowding is the combining of features over an inappropriately large area. Object recognition is usually limited by spacing and not by size. To be identified, simple objects must be separated by at least the observer’s critical spacing, which corresponds to 6 mm at the primary visual cortex. Compound objects, such as words and faces, can crowd themselves. Their parts must be separated by at least the critical spacing. Thus, in our cluttered world, observers can identify objects only in an uncrowded window, determined by the object spacing. When the spacing is uniform, as in text, then the window will be central, where the critical spacing is smallest. These conclusions all spring from the consistent observations that, for each observer, the critical spacing of crowding depends solely on location and direction, which we call the Bouma law.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## References

1. Barlow HB. Summation and inhibition in the frog’s retina. J. Physiol. (Lond.) 1953;119:69–88. [PubMed: 13035718]

2. Robson JG, Graham N. Probability summation and regional variation in contrast sensitivity across the visual field. Vision Res 1981;21:409–418. [PubMed: 7269319]
3. Treisman AM, Gelade G. A feature-integration theory of attention. Cognit. Psychol 1980;12:97–136. [PubMed: 7351125]
4. Pelli DG, Burns CW, Farell B, Moore-Page DC. Feature detection and letter identification. Vision Res 2006;46:4646–4674. [PubMed: 16808957]
5. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. (Lond.) 1962;160:106–154. [PubMed: 14449617]
6. Desimone R, Duncan J. Neural mechanisms of selective visual attention. Annu. Rev. Neurosci 1995;18:193–222. [PubMed: 7605061]
7. Ullman, S. High-Level Vision: Object Recognition and Visual Cognition. Cambridge, Massachusetts: MIT Press; 2000.
8. Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. Compulsory averaging of crowded orientation signals in human vision. Nat. Neurosci 2001;4:739–744. [PubMed: 11426231]
9. Motter BC. Crowding and object integration within the receptive field of V4 neurons. J. Vis 2002;2:274:247a.
10. Ledgeway T, Hess RF, Geisler WS. Grouping local orientation and direction signals to extract spatial contours: empirical tests of "association field" models of contour integration. Vision Res 2005;45:2511–2522. [PubMed: 15890381]
11. Intriligator J, Cavanagh P. The spatial resolution of visual attention. Cognit. Psychol 2001;43:171–216. [PubMed: 11689021]
12. Biederman I. Recognition-by-components: a theory of human image understanding. Psychol. Rev 1987;94:115–147. [PubMed: 3575582]
13. Tanaka JW, Farah MJ. Parts and wholes in face recognition. Q. J. Exp. Psychol. A 1993;46:225–245. [PubMed: 8316637]
14. Martelli M, Majaj NJ, Pelli DG. Are faces processed like words? A diagnostic test for recognition by parts. J. Vis 2005;5:58–70. [PubMed: 15831067]
15. Geisler W, Murray R. Cognitive neuroscience: practice doesn't make perfect. Nature 2003;423:696–697. [PubMed: 12802318]
16. Levi DM. Crowding - an essential bottleneck for object recognition: a mini-review. Vision Res 2008;48:635–354. [PubMed: 18226828]
17. Pelli DG, Palomares M, Majaj NJ. Crowding is unlike ordinary masking: distinguishing feature integration from detection. J. Vis 2004;4:1136–1169. [PubMed: 15669917]
18. He S, Cavanagh P, Intriligator J. Attentional resolution and the locus of visual awareness. Nature 1996;383:334–337. [PubMed: 8848045]
19. Blake R, Tadin D, Sobel KV, Raissian TA, Chong SC. Strength of early visual adaptation depends on visual awareness. Proc. Natl. Acad. Sci. USA 2006;103:4783–4788. [PubMed: 16537384]
20. Bouma H. Interaction effects in parafoveal letter recognition. Nature 1970;226:177–178. [PubMed: 5437004]
21. Freeman J, Pelli DG. An escape from crowding. J. Vis 2007;7:1–14.
22. Reddy L, VanRullen R. Spacing affects some, but not all, visual searches: implications for theories of attention and crowding. J. Vis 2007;7:1–17. [PubMed: 18217818]
23. Solomon JA, Morgan MJ. Stochastic re-calibration: contextual effects on perceived tilt. Proc Biol Sci 2006;273:2681–2686. [PubMed: 17002955]
24. Bouma H. Visual interference in the parafoveal recognition of initial and final letters of words. Vision Res 1973;13:767–782. [PubMed: 4706350]
25. Bex PJ, Dakin SC, Simmers AJ. The shape and size of crowding for moving targets. Vision Res 2003;43:2895–2904. [PubMed: 14568377]
26. van den Berg R, Roerdink JBTM, Cornelissen FW. On the generality of crowding: visual crowding in size, saturation and hue compared to orientation. J. Vis 2007;7:1–11.
27. Chung STL, Li RW, Levi DM. Crowding between first- and second-order letter stimuli in normal foveal and peripheral vision. J. Vis 2007;7:1–13.
28. Pelli DG, et al. Crowding and eccentricity determine reading rate. J. Vis 2007;7:1–36.

29. Huey, EB. The Psychology and Pedagogy of Reading. New York: Macmillan; 1908.

30. Legge GE, Pelli DG, Rubin GS, Schleske MM. Psychophysics of reading. I. Normal vision. Vision Res 1985;25:239–252. [PubMed: 4013091]

31. Levi DM, Song S, Pelli DG. Amblyopic reading is crowded. J. Vis 2007;7:1–17.

32. Vlaskamp BN, Over EA, Hooge IT. Saccadic search performance: the effect of element spacing. Exp. Brain Res 2005;167:246–259. [PubMed: 16078032]

33. Woodworth, RS. Experimental Psychology. New York: Holt; 1938.

34. Bouma, H. Visual search and reading: eye movements and functional visual field: a tutorial review. In: Requin, J., editor. Attention and Performance VII. Hillsdale, New Jersey: Erlbaum; 1978.

35. Bosse ML, Tainturier MJ, Valdois S. Developmental dyslexia: the visual attention span deficit hypothesis. Cognition 2007;104:198–230. [PubMed: 16859667]

36. McConkie GW, Rayner K. The span of the effective stimulus during a fixation in reading. Percept. Psychophys 1975;17:578–586.

37. Engbert R, Nuthmann A, Richter EM, Kliegl R. SWIFT: a dynamical model of saccade generation during reading. Psychol. Rev 2005;112:777–813. [PubMed: 16262468]

38. O'Regan JK. Eye movements and reading. Rev. Oculomot. Res 1990;4:395–453. [PubMed: 7492534]

39. Legge, GE. Psychophysics of Reading in Normal and Low Vision. Mahwah, New Jersey: Erlbaum; 2007.

40. Motter BC, Simoni DA. The roles of cortical image separation and size in active visual search performance. J. Vis 2007;7:1–15. [PubMed: 18217821]

41. Stanovich KE, Siegel LS, Gottardo A. Converging evidence for phonological and surface subtypes of reading disability. J. Educ. Psychol 1997;89:114–127.

42. Bouma H, Legein CP. Foveal and parafoveal recognition of letters and words by dyslexics and by average readers. Neuropsychologia 1977;15:69–80. [PubMed: 831155]

43. 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 2007;47:2521–2530. [PubMed: 17719073]

44. Kwon M, Legge GE, Dubbels BR. Developmental changes in the visual span for reading. Vision Res 2007;47:2889–2900. [PubMed: 17845810]

45. Lévy-Schoen A. Exploration et connaissance de l'espace visuel sans vision périphérique. Trav. Hum 1976;39:63–72.

46. Chung STL, Levi DM, Legge GE. Spatial-frequency and contrast properties of crowding. Vision Res 2001;41:1833–1850. [PubMed: 11369047]

47. Toet A, Levi DM. The two-dimensional shape of spatial interaction zones in the parafovea. Vision Res 1992;32:1349–1357. [PubMed: 1455707]

48. Taylor SE. Eye movements in reading: facts and fallacies. Am. Ed. Res. J 1965;2:187–202.

49. Valdois S, et al. Phonological and visual processing deficits can dissociate in developmental dyslexia: evidence from two case studies. Read. Writing 2003;16:541–572.

50. Martelli M, Di Filippo G, Spinelli D, Zoccolotti P. Crowding, reading and developmental dyslexia. Perception 2006;35

**Figure 1.**

An A in chaff. The bars represent elementary visual features. Fixating close to the bars, at the green plus, makes it easy to recognize the letter A. If you fixate far away, on the red minus, you can still see the features, but you cannot identify the letter. Your visual system is combining over too large an area, including all the features from both the A and the surrounding chaff, which results in a jumbled percept. This is crowding. You can rule out acuity (letter size) as an explanation (for your inability to identify the A) by confirming that you can see the A while fixating the minus if your fingers hide the chaff (for a review, see ref. [17]).

**Figure 2.**

Effects of crowding. While fixating the red minus, can you tell that the clusters differ in letter identity, number and position? Crowding impairs your ability to judge these object properties[20,21]. Using your finger to cover all but the leftmost letter, you can confirm that even this most distant letter is well within your acuity (reprinted from ref. [21]).

**Figure 3.**
Crowding in a word. While fixating the red minus, it is easy to identify the isolated letter on the left, but try to identify the middle letter on the right. It is hard. Fixate the green plus and try again. Now it is easy[24].

**Figure 4.**

Faces are like words. Arnold Schwarzenegger and Elvis Presley are famous, and their faces may be familiar. Fixate on the red minus between them. Can you still recognize the governor and the King? How close to each face do you have to fixate to identify it? As you fixate closer and closer to the face, you will find that you remain unable to recognize it until you are near the cheek. As with words, the parts (eyes, nose and mouth) of faces must be isolated (separated by the observer’s critical spacing) for the whole to be recognized[14].

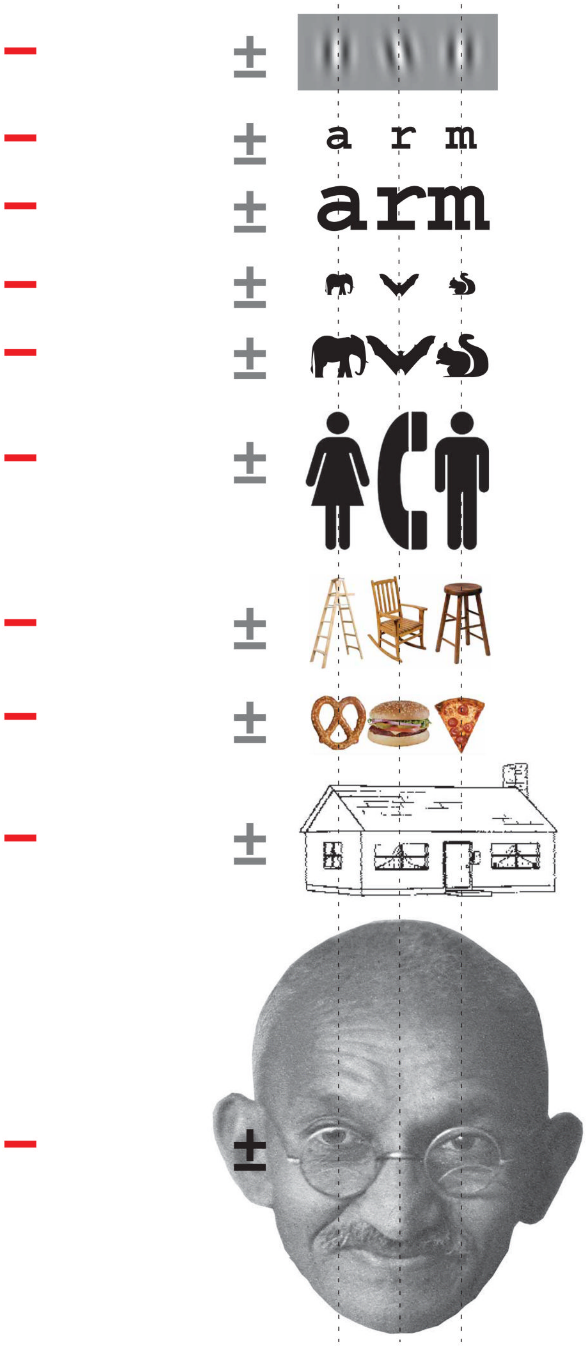

**Figure 5.**

Critical spacing is independent of object and size. Fixating on the red minus, you will be unable to identify the middle object in the first nine rows unless you isolate it by hiding the flanking objects with your fingers (or two pencils). In the last two rows, you will be unable to recognize the single object while fixating on the red minus. Grating patches, similar to those in the top two rows, are often taken to be one-feature objects. In the first row, is the middle grating vertical or tilted? The ± is our estimate of the fixation point where you can just barely identify the target. You can assess the accuracy of this threshold estimate by noting that the task is easy when you fixate to the right of the ± and hard when you fixate to the left. Critical spacing depends solely on position (and direction) in the visual field, which does not vary among rows in this demonstration. Note that halving object size has no effect on critical spacing. Critical spacing is independent of spatial frequency[46] (see Supplementary Sources online).

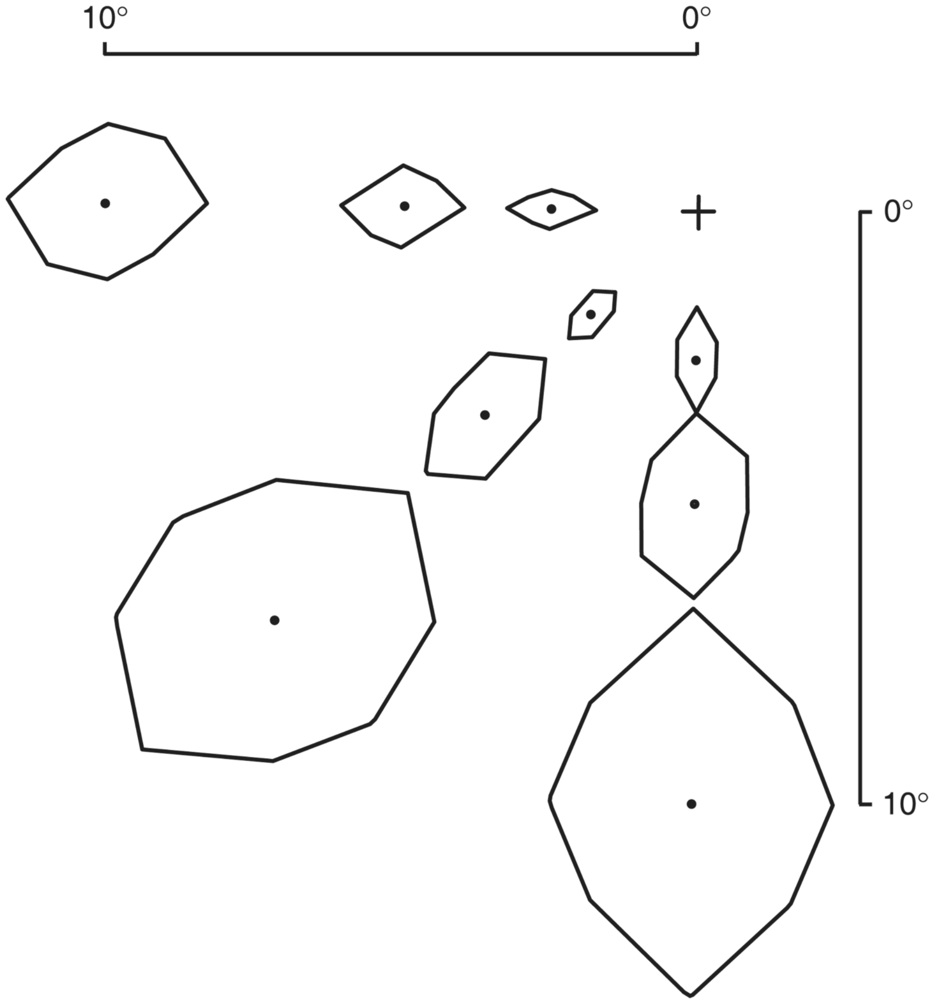

**Figure 6.**

Critical spacing is proportional to eccentricity. The observer fixated on the point indicated by a plus in the upper right and identified the orientation of a target T (right-side up or upside down?) presented (in blocks) at one of the nine locations indicated by the dots. Two flanking Ts were shown symmetrically displaced from the target in opposite directions, −45°, 0°, 45° or 90° relative to horizontal. Each vertex in the roughly elliptical contours represents the measured critical spacing of the pair of flanking letters for 75% correct identification of target orientation. Note that the critical spacing contours are not circles; the direction from target to flanker matters. These were measured with one letter size at each eccentricity. Changing letter size has no effect on the results[28] (figure adapted from ref. [47]).

xuncrowdedx

**Figure 7.**

What is your uncrowded span? Fixate on the o in the center of the word. Your uncrowded span is 3 if you can read ‘row’, 4 for ‘crow’, 5 for ‘crowd’ and a whopping 9 for ‘uncrowded’, which many observers achieve. The variation in the uncrowded span reflects the substantial individual differences in critical spacing reported previously[47]. The Bouma law says that critical spacing is invariant across objects, not subjects (for reviews of uncrowded and visual spans, see refs. [28,32,39]). Image reprinted from ref. [28] and adapted from ref. [33].

**Figure 8.**

The uncrowded window. This figure simulates crowding in reading by substituting letters in the peripheral field. Crowding spoils letter recognition, making reading impossible outside of the uncrowded window. Note that the substitutions are undetectable when you fixate on the center of the circle. As you read this caption, the words are clear and legible near your chosen point of fixation and illegibly crowded beyond that clear region. That central uncrowded field is a window through which we read (figure adapted from ref. [28]).

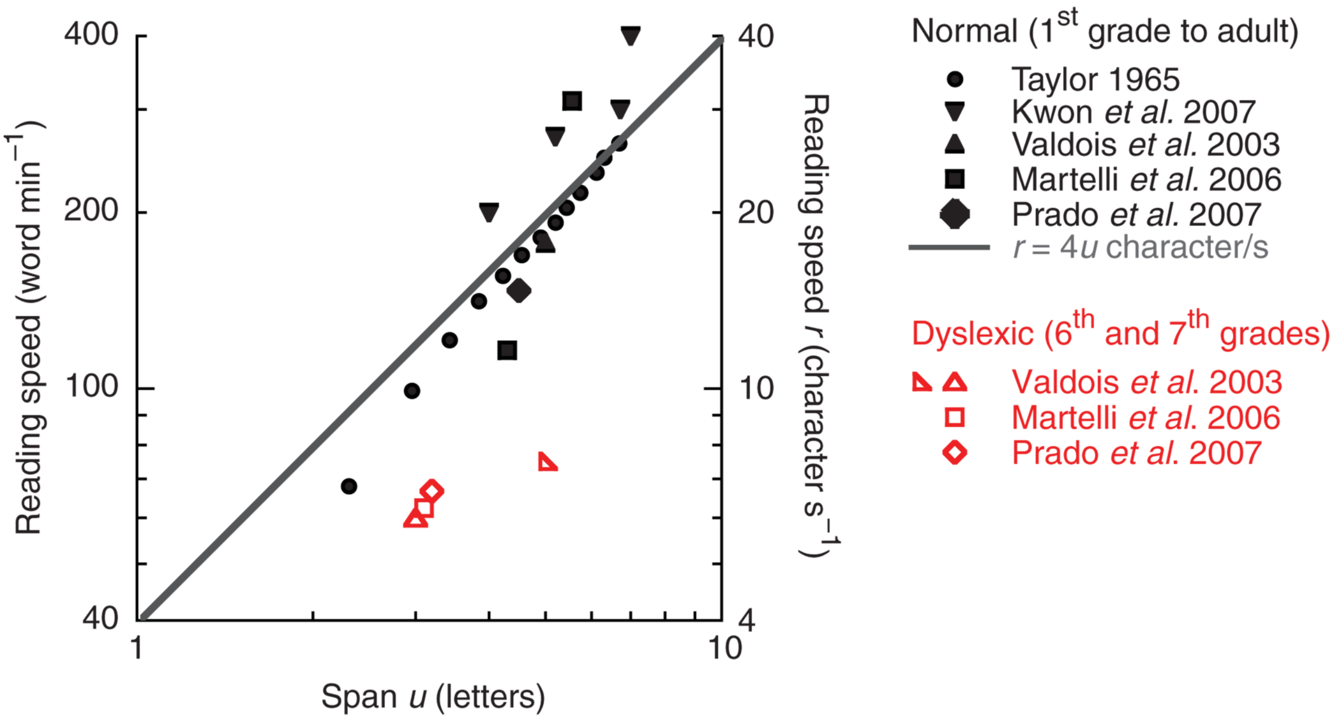

**Figure 9.**

Reading speed versus span. Data are from five studies of normal (black filled symbols) and dyslexic (red empty symbols) readers[44,45,48–50]. The normal readers were of various ages, from 1st grade (age 6) through adult. Reading speed rose monotonically with age. The dyslexic readers were all in the 6th or 7th grades. The vertical scale is reading speed (1 word min$^{-1}$ = 0.1 character s$^{-1}$, assuming an average of five letters and a space for each word). The horizontal scale is letter span, estimated in various ways. Span is the width (in characters) of the uncrowded window. A reader making ρ eye movements per second, advancing an average of $u$ characters per eye movement, reads at a rate $r = \rho u$ character s$^{-1}$. The diagonal line plots this proportionality, assuming 4 eye movements per second ($\rho = 4$ Hz), showing that this simple 4 Hz rule gives a fairly good account of all the data from normal readers (see Supplementary Methods online).

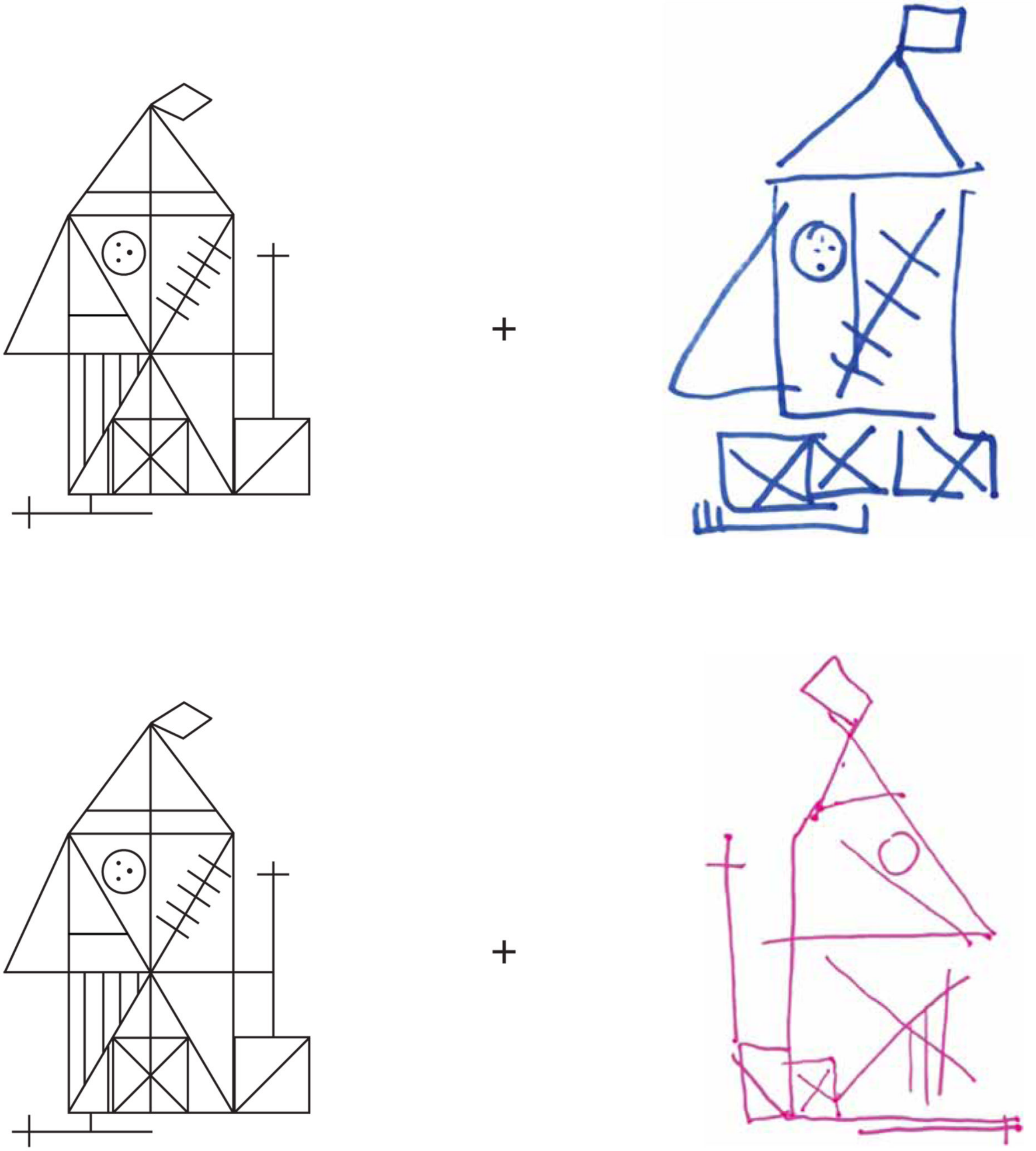

**Figure 10.**

The Rey Complex Figure Test. The original diagram is on the left. The drawings on the right were made by normally sighted graduate students who were asked to copy, from left to right, while fixating on the central + (ignore the left-right reversal, which was the result of ambiguity of the copying instructions). A neurologist who examined these drawings found them to be typical of those produced with unrestricted viewing by patients with apperceptive agnosia. Despite the amateur drawing skill of the students, you can verify that these are reasonably good copies for your peripheral vision by fixating on the central +. Courtesy of M. Martelli (Università di Roma “La Sapienza”).