

Published in final edited form as:

Trends Cogn Sci. 2011 April ; 15(4): 160–168. doi:10.1016/j.tics.2011.02.005.

Visual Crowding: a fundamental limit on conscious perception and object recognition

David Whitney¹ and Dennis M. Levi²

¹Department of Psychology, University of California, 3210 Tolman Hall, Berkeley, CA 94720-1650; Helen Wills Neuroscience Institute, University of California, 3210F Tolman Hall, Berkeley, CA 94720-3192

²School of Optometry, University of California, 351 Minor Addition, Berkeley, CA 94720-2020; Helen Wills Neuroscience Institute, University of California, 3210F Tolman Hall, Berkeley, CA 94720-3192

Abstract

Crowding, the inability to recognize objects in clutter, sets a fundamental limit on conscious visual perception and object recognition throughout most of the visual field. Despite how widespread and essential it is to object recognition, reading, and visually guided action, a solid operational definition of what crowding is has only recently become clear. The goal of this review is to provide a broad-based synthesis of the most recent findings in this area, to define what crowding is and is not, and to set the stage for future work that will extend crowding well beyond low-level vision. Here we define five diagnostic criteria for what counts as crowding, and further describe factors that both escape and break crowding. All of these lead to the conclusion that crowding occurs at multiple stages in the visual hierarchy.

Dispelling the illusion

With regular flicks of the eye, we establish and maintain the illusion of a continuous high-resolution representation of our visual environment. This compelling illusion is easy to dispel by trying to describe the details of objects in your peripheral visual field—scrutinizing or trying to count objects in the visual periphery is impossible. This partly reflects the well-known decline in visual acuity in peripheral vision. However, the most widespread impediment to reading and object recognition in the periphery is the mysterious process known as *crowding*—the deleterious effect of clutter on peripheral object recognition. Objects that can be easily identified in isolation seem indistinct and jumbled in clutter (Fig. 1).

Crowding is an essential bottleneck, setting limits on object perception, eye and hand movements, visual search, reading and perhaps other functions in peripheral, amblyopic and developing vision. Crowding impairs not only discrimination of object features and contours but the ability to recognize and respond appropriately to objects in clutter. Thus, studying

© 2011 Elsevier Ltd. All rights reserved.

Corresponding Author: Levi, D.M. (dlevi@berkeley.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

DML and DW contributed equally to this work.

crowding may lead to a better understanding of the processes involved in object recognition. Crowding also has important clinical implications for patients with macular degeneration, amblyopia and dyslexia.

Interest in crowding has significantly increased in the past few years, yielding a more sophisticated understanding of the phenomenon itself as well as of the processes involved in object recognition and reading. Two reviews [1,2] provide overviews of much of the relevant literature published at the time. Levi[1] concluded that “Crowding is an enigma wrapped in a paradox and shrouded in a conundrum. Despite a great deal of new (and old) work, we do not yet have a full understanding of crowding.” Since then, new approaches, models and findings have provided new insights into the mysteries of crowding, suggesting that crowding occurs at multiple stages in the visual hierarchy.

Operationally defining crowding

The significance of crowding is clear from phenomenological demonstrations of its power and ubiquity in natural scenes (Fig. 1). Ultimately, however, characterizing and understanding the mechanism(s) of crowding requires more than a phenomenological description. Recent work has established that there are several diagnostic criteria for crowding, and using these as converging evidence can help studies individuate and distinguish crowding from other effects, such as masking, lateral interaction and surround suppression. All of these share the characteristic of making a target more difficult to see, but each is distinct and most likely reflects different neural processes.

Diagnostic Criteria

i) Crowding impairs identification, not detection—When objects are crowded they do not simply disappear, as might be expected if crowding was a disruptive process that suppressed their signals. Fig. 1B confirms that crowding has little or even no effect on detection of a feature or object [3-5].

ii) Crowded objects are perceived as having high contrast but are indistinct or jumbled together—This can be easily confirmed from inspection of Fig. 1B.

iii) Critical Spacing—Is Bouma’s rule a law?—Crowding depends on the eccentricity of a target object and how densely spaced the surrounding objects are (Fig. 2). At a given eccentricity, identification of a crowded target improves as the distance between the target and flankers increases. In his highly influential report, Bouma stated “for complete visual isolation of a letter presented at an eccentricity of ϕ° , it follows that no other letters should be present (roughly) within $0.5 \phi^\circ$ distance”[6], giving rise to the notion of a “critical spacing” that is proportional to eccentricity. Bouma’s proportionality constant, b varies across studies, depending upon both how it is measured and computed, but it is widely reported to be $\approx 0.4-0.5$. Thus an object at an eccentricity of 10 deg may be crowded by other objects as much as 5 degrees away. Bouma’s proportionality constant, or Bouma’s rule, is sometimes conferred the status of a “law,” but this is controversial (see Box 1).

iv) Anisotropies—Crowding in peripheral vision is not isotropic. There is a very substantial ($\approx 2:1$) radial-tangential anisotropy[7], such that radially positioned flankers are more effective than tangentially positioned ones. For example, in the vertical meridian, vertically arranged flankers are more potent than horizontally arranged flankers, whereas in the horizontal meridian, horizontally arranged flankers are more potent than vertically arranged ones. In diagonal locations of the four quadrants, crowding is significantly stronger (critical spacing is larger) when the target and distractors are horizontally rather than vertically arranged[8] (this does not occur on the vertical meridian because the radial-

tangential anisotropy dominates). In addition, crowding is also stronger in the upper field than in the lower field[9]. Finally, a recent report suggests that crowding is also stronger when the distractors and target are within the same visual field than in separate visual fields, despite equated retinal distance [10].

v) Asymmetries—Peripheral crowding is asymmetric. Bouma[6] noted that two flankers (one on each side of the target letter) were much more potent than one, and that crowding was stronger with a single flanker at an eccentric locus greater than the target compared to a single flanker at an eccentric locus nearer to the fovea (at the same angular separation from the target). This inner/outer asymmetry occurs for recognition of letters (e.g., Ref[11]), Gabor patch orientation[12], and face recognition[13]. While this asymmetry might be readily explained on the basis of cortical geometry (the far flanker is actually closer to the target than the near flanker after mapping to cortical space; [14]) we know of no similar explanation for the large radial/tangential anisotropy.

vi) Temporal tuning—While crowding is usually thought of as a spatial phenomenon, it also occurs over time. For example, moving objects are crowded more by flankers that slightly lead than trail the target [15], and the temporal relationship between the target and flankers modulates crowding [16-18]. This temporal tuning of crowding may be a diagnostic feature, or at least a distinguishing characteristic. That the crowding effect is strong for nearly simultaneous flankers and target[17,18] distinguishes it from object substitution and metacontrast masking [19]. That the full-width at half max of the temporal tuning function is approximately 150 msec and is not strongly skewed toward backward masking (that is, targets presented before flankers are about as crowded as targets presented after flankers [17,18]) could further distinguish crowding from backward pattern masking. Whether there is an independent mechanism of temporal crowding remains unclear, but the effects of spatial crowding are correlated with those of temporal crowding across subjects [20], supporting the possibility that crowding involves spatiotemporal and perhaps attentional mechanisms (e.g., [16,21]).

Crowding and appearance

As discussed earlier, crowded objects do not simply disappear. On the contrary, crowding changes the appearance of the crowded zone, which is important, because it can help distinguish among the main models for crowding (masking, pooling and substitution—discussed below). As examples, Tyler & Likova noted their impression of a crowded letter as a “gray, or inchoate, smudge between the two outer letters, including the inner parts of those letters” [22] (see Figs. 1 and 2). Greenwood, Bex & Dakin[23] elegantly showed that crowded objects appear to take on characteristics of the flankers, a finding consistent with the “jumbled” percept that accompanies crowding. They conclude, in agreement with Levi & Carney [24] that crowding is a regularization process that simplifies the appearance of the peripheral array by promoting consistent appearance among adjacent objects.

Greenwood et al’s [23] results are consistent with a number of earlier studies suggesting that information about crowded objects is not lost. In particular, Parkes et al. [25] found that the orientation signals from the target and flankers in a cluttered peripheral display were pooled rather than being lost through masking. They concluded that crowding reflects compulsory averaging of signals (but see Refs[26-28]) and that crowding is the term we use to define texture perception “when we do not wish it to occur”.

Under conditions of crowding, orientation perception is characterized by strong perceptual assimilation (e.g., the flanker orientation captures the target) near the target and perceptual repulsion (i.e. “anti-crowding”) farther from the target[27,29-31]. Assimilation could

regularize perception of the peripheral array, whereas repulsion could highlight salient differences among visual signals (making different stimuli “pop out”). Assimilation and repulsion reflect opponent influences on orientation perception, and a recent study [30] suggests that the switch from assimilation (crowding) to repulsion depends on cortical distance. Whether this is specific to orientation, or more general is not yet known.

What information survives crowding?

The type of information that gets through to conscious perception under conditions of crowding provides important clues about the nature of crowding.

Detection of a feature

Observers can easily detect the appearance of a feature under conditions that render identification or discrimination of a change in the feature impossible [3-5,24].

Aftereffects following adaptation to a crowded target (Fig. 3A-B)

While a target may be crowded and indiscriminable, adapting to it can produce a variety of aftereffects including orientation [9] and motion [32,33]. Whether these aftereffects are modulated by crowding is debated [34], but their existence shows that crowding does not destroy feature-level information.

Statistical properties

As mentioned earlier, Parkes et al[25] elegantly demonstrated that while observers are unable to correctly report the orientation of an individual patch under conditions of crowding, they can reliably report the average, ensemble orientation, suggesting that the local orientation signals are combined rather than lost (Fig. 3C). This led to the widely held notion that crowded signals undergo a form of compulsory pooling or averaging of signals, a finding that has now been demonstrated under a variety of different conditions (e.g., [24,35-37]) and forms the basis of the “faulty integration” theory.

Some target identity information

A well-established effect of crowding is substitution-like effects—in crowded displays, observers frequently mistakenly report a flanker rather than the target. Whether this reflects positional uncertainty (the observer confuses the position of the flanker with that of the target) or simply the fact that the observer has to report something, is not clear (i.e., if they are unable to see the crowded target, they simply report what they could see, the flanker). However, when required to report all the letters in a crowded display (i.e., give a full report of target and flanker letters), the proportion of correct “target” responses is much higher when the correct order (position) is not required [38,39]. For example, given a crowded display ‘BTH’, an observer may respond ‘BHT’. Clearly some information about the target is preserved (perhaps even semantic information [40]), but the location information is lost.

What information breaks crowding?

Under certain circumstances, crowding may be reduced or released completely.

- i. Ungrouping of target and flankers. In peripheral vision, there is a predilection to perceptually group features into a Gestalt.

Target-flanker grouping. When targets and flankers are similar they are likely to group, and when they are dissimilar they ungroup and the target pops out (Fig. 4). Thus, crowding is reduced when targets and flankers are dissimilar in shape and size [41,42], orientation [4,43,44], polarity[16,41], spatial frequency [45], depth

[41] color [41,46-48], synesthetic color to some degree [49], motion [50] and “order” (i.e., first-order vs. second order[51]). Temporal grouping also modulates crowding. Crowding is maximal when targets and flankers are presented nearly simultaneously; presenting targets before or after the flankers (by ~150 ms) is sufficient to break crowding [17,18].

Flanker-flanker grouping. In multi-element flankers, when the flankers group separately from the target, crowding may be reduced [24,28,52-56]. Thus when target and flanker look like a regular texture, it is difficult to make judgments about the target and crowding is strong, whereas when the target appears distinct from the flankers, crowding is weak or absent[56].

- ii. Object-centered or “holistic” crowding (Fig 4D-E). Crowding can occur between configural, high-level representations of objects. Inverting face flankers can release face crowding. Specifically, it is harder to recognize an upright target face when it is surrounded by a crowd of nearby upright faces than by a crowd of inverted faces[13,57] (Fig. 4). The inversion effect in crowding also occurs between Mooney (two-tone) faces[13], stimuli that require holistic processing[58]. These object-centered crowding effects adhere to all of the diagnostic criteria for crowding, and are not due to masking, similarity effects, or grouping of low-level features. Likewise, holistic crowding—crowding between upright face representations—can be distinguished from within-face or facial feature crowding[13,59]. In fact, these object-centered crowding effects demonstrate compound crowding within the same stimulus—crowding between the whole upright faces, and between the low level features that comprise each face; this suggests that crowding operates at multiple stages. Recent work demonstrates that object-centered crowding effects also occur for letter-like stimuli[35], raising the possibility that object-centered, holistic crowding might occur independently at different levels of visual processing.
- iii. Attention. Several recent studies have demonstrated that cueing a crowded target location reduces the effects of crowding [21,39,60,61]. Dakin and colleagues[62] also found that devoting attention to the target region ameliorated the effect of crowding. The only explanation for these findings, as well as Cavanagh & Holcombe’s demonstration of attentionally-gated crowding effects (*Vision Sciences Society Annual Meeting*, 2007), is that attention can modulate the critical spacing in crowded arrays (see also Box 1).
- iv. Masked flankers. Crowding may be “released” when the flankers are masked. However, Chakavarthy & Cavanagh[63] showed that this release only occurred with noise and metacontrast masks but not with object substitution masks. They argue that noise and metacontrast masks act early in the visual processing cascade, degrading the features, whereas object substitution masks do not interfere with feature encoding but act much later, by replacing the representation of the stimulus.
- v. Suppression of flankers from visual awareness. Wallis & Bex[64] used “adaptation-induced blindness” to render flankers perceptually invisible and used a dual report paradigm to obtain a trial-by-trial assessment of awareness and crowding. Target identification was dependent on the number of flanking letters perceived on a given trial, independent of the number that were physically present, and they concluded that crowding is “released” when the flankers are suppressed from visual awareness.

Where in the brain does crowding take place?

There has been a great deal more psychophysical than neurophysiological work on crowding. More than anything, the neurophysiological studies of crowding have proven that it is surprisingly difficult to isolate the neural mechanism(s) of crowding per se (Box 2). Nevertheless, the psychophysical studies help narrow down the level(s) at which crowding occurs, and will help guide the design of more stringent future neurophysiological experiments.

Although there is great diversity in the stimuli employed (ranging from oriented bars, Gabor patches and shapes to letters, words, and faces among others—for reviews, see [1,2]), most studies of crowding implicitly (if not explicitly) argue that crowding is a unitary phenomenon, occurring at a single circumscribed level of visual processing, or perhaps in a particular visual area. Early work demonstrated that crowding works dichoptically (target to one eye, distractors to the other eye) [65,66], suggesting that crowding arises in the cortex. Since then, various researchers have suggested that the site of crowding might be V1 (e.g., [67]), V2 (e.g., Freeman and Simoncelli (*Vision Sciences Society Annual Meeting*, 2010), V3 (e.g., [22,68]), V4 (e.g., [10,69] but c.f., Ref [70]), or even later in visual processing[32,57]. The evidence that crowding occurs in each of these and other visual areas is mixed, and difficulties comparing across studies is compounded by the fact that crowding can occur selectively between different kinds of stimuli (see Box 1 and the preceding section). Even within a stimulus type (e.g., orientation, motion, or faces), crowding is modulated by stimulus similarity, context, and attention (Fig. 4 and Box 1), as discussed earlier. Concluding that any single visual area could explain this range of effects is therefore tenuous, at best. Neurophysiological studies have provided relatively little additional evidence about the neural mechanism(s) of visual crowding (Box 2).

Models of crowding

There is no shortage of ideas about crowding, but few are computational or make specific quantitative predictions. The large number of different models may be distilled down to three basic classes: i) masking, ii) pooling (either pooling of low level features or pooling by attention) and iii) substitution. Within each class many different architectures and algorithms have been proposed. These extant models are largely descriptive, and have been reviewed in detail elsewhere[1]. There are few quantitative models of crowding, and most are quite recent. Wilkinson et al[71] proposed a model in which complex cells and simple cells interact by mutual inhibition. In this model, isolated visual contours are processed by simple cells, which suppress weak complex cell responses. However, in the presence of nearby similarly oriented flanking contours in a small area, complex cells respond vigorously because of spatial pooling, and they then suppress simple cell activity within their receptive field area. This texture model nicely predicts several aspects of their data; however, the pooling parameter was based on the simulations that best fit the data, rather than on physiology or some other principled approach.

There are three more recent approaches to modeling crowding. Van den Berg and colleagues[72] propose a quantitative model for spatial integration of orientation signals, based on the principles of population coding. Their model nicely predicts several properties of crowding, including critical spacing, “compulsory averaging”, and the inner/outer asymmetry. However, in its current form it fails to predict the effect of target flank similarity [41,42], the configuration effects [24,28,52], and object-centered or holistic crowding [13,35,57]. Dayan & Solomon[73] take a very different approach, in which spatial selection of a target among flankers emerges through a process of Bayesian inference, in a computational form. “Interference” (also known as crowding) in this model results from the

spatial uncertainty inherent in large receptive fields, and receptive field size is assumed to increase with eccentricity according to the cortical magnification factor. The model was developed to explain the Eriksen flanker task. It remains to be seen how well it accounts for many of the key features of crowding reviewed here.

None of the models naturally accounts for the radial/tangential anisotropy in a principled way. Van den Berg et al [72] simply use different parameters to define the radial and tangential integration fields. In contrast, Nandy and Tjan (*Vision Sciences Society Annual Meeting*, 2010) begin with a model of cortical area V1 and its geometry and lateral connections, quite similar to the model of Neri & Levi [74], combined with the important role of natural image statistics[75]. However, the novel insight and advance in their model is the idea that image statistics are acquired primarily at attended spatial locations via a gating mechanism, and that spatial attention and any subsequent eye movement that it elicits overlap in time. Nandy & Tjan further argue that learning image statistics during development leads to the formation of lateral connections that distort the true image statistics in the peripheral field, leading in turn to the radial-tangential anisotropy of crowding. Whether this model will be able to account for grouping, similarity, feature/object specific crowding, remains to be seen.

Any successful computational model of crowding needs to account for each of the characteristics above, including the diagnostic criteria and the factors that modulate crowding. It is not sufficient for a model of crowding to simply mimic or reproduce the phenomenological “jumble” that is representative of crowding [75]; the model must hold to the diagnostic criteria for crowding and it must account for what escapes crowding, what breaks crowding, and what modulates crowding.

Multiple levels of crowding

Although there is great heterogeneity in the results on crowding, there is sufficient evidence to cast doubt on the idea that crowding is a unitary effect due to a single stage of processing, though this is implicitly assumed in most studies of crowding (reviewed in Refs [1,2]); rather, the collective work suggests that crowding happens independently at several stages of visual processing. In support of this view are the observations that crowding is specific to the similarity between, and the configuration of, target and flanks (discussed above and in Box 1), and the fact that there is “compound” crowding: in a given scene, crowding occurs selectively between features [1], object parts [59], and whole objects [13,57]. These make an ‘all-convergent’ crowding stage unlikely. If crowding occurs at multiple levels of visual analysis, or if different ‘channels’ (chromatic, spatial frequency, object, etc.) each possess their own unique crowding bottleneck, then one would expect that the gradient of crowding as a function of eccentricity might be channel or stimulus specific. There is intriguing evidence building in favor of this view (Box 1), but much work remains (Box 3).

In natural scenes, crowding may be ubiquitous, but it may occur in layers, with location, content, and attention-dependence. One of the major challenges in future work on crowding will be to develop a parsimonious model that can account for the diversity of findings. Individual models that ignore content, category boundaries, similarity, and attention are not likely to succeed.

The crowded future

Our discussion of crowding has been confined exclusively to visual perception, but the implications of crowding are far and wide and will likely expand rapidly in the near future (see Box 3). For example, crowding may exert a fundamental limit on visually guided actions in naturally cluttered scenes. Although the impact of crowding, per se, on visually

guided action has rarely been investigated [14,76,77], several studies suggest that clutter impairs action (e.g., Refs[78,79]). More intriguingly, there is evidence that crowded visual information is differentially used by perceptual and visuomotor systems [76]. Such dissociations between the perceptual and motor responses to crowded scenes may help address the paradoxical but unanswered question of why eye (and hand) movements are not more random than they are in natural scenes—if peripheral object identities are crowded and unrecognizable, how do we make accurate, non-random eye and hand movements to those objects? The practical consequences of visuomotor crowding, divorced from perceptual crowding, would be extensive, ranging from clinical settings to ergonomics and human factors; therefore, the search should be on for dissociations between perception and action in crowded scenes (Box 3).

Because crowding limits visually guided action and it defines the resolution of conscious visual perception, another important question is how crowding develops from infancy to old age. Although little work has been done specifically on crowding in older populations (c.f., Ref[80]), studies with infants as young as 6 months[81], toddlers[82], and adolescents from 8 to 11 years old[83] have reported that children are much more impaired by crowding than adults, even when their acuity is fully developed[83]. Given this protracted development of crowding, an important question is what sorts of knock-on effects might occur later in life if crowding developed abnormally in infancy. For example, crowding may be uniquely and differentially impacted in neurodevelopmental, visual, and cognitive disorders including autism[84], dyslexia[85], amblyopia[86], and macular degeneration[1], among others. Therefore, the clinical and practical implications of crowding, as well as its training and possible rehabilitation, are widespread.

Box 1: Is Bouma's law a law?

Many studies have confirmed that the critical spacing for crowding depends on target eccentricity, not target size [3-5,87]. Recently, it was suggested that Bouma's rule of thumb should be elevated to the status of a "law" [2,67]. If it were a "law," it would have important implications: because of the log conformal mapping of the visual world onto the retinotopic visual cortex [88], it suggests that regardless of the eccentricity, the critical distance represents a fixed distance on the cortex. Thus, objects can only be recognized when they are sufficiently separated on retinotopically organized cortex [67,89]. Pelli[67] argues that the critical spacing on the cortex is ~6 mm in area V1 (which is also the range of dichoptic interactions in the region of the cortex corresponding to the blind spot[66]).

The notion that crowding depends simply on cortical distance is simple and appealing. But is Bouma's law truly a law? Is it physiologically plausible that the critical spacing is independent of stimulus, task, attention and other factors? Figure I suggests not. Specifically, Fig. B1 (top panel) shows that Bouma's "constant" (b) is substantially larger when the polarity, color or shape of targets and flankers are the same than when they differ [41,46,47]. Figure B1 shows that b is also larger when the complexity of targets and flankers are the same compared to when they differ[90]. Recent work also suggests that b is larger when the target and flankers are letter-like symbols rather than actual letters (Fig. B1 middle panel[91]) and that b is smaller when the target location is cued versus not cued (Fig. B1 lower panel[21]). Yeshurun & Rashal's [21] study differs from a number of previous studies, which have shown either no effect or a small effect of cueing on crowding [42,71,92], because they took measures to avoid having the cue mask the target. Because even very weak target masking can greatly enhance the measured crowding effect [93], it is crucial to control for interactions between crowding and masking. In addition to modulations by attention [21,39,60,61], b can be modulated by perceptual learning [94,95]. The implications of these and other related findings

[13,57,96,97] are discussed in the main text (see also Fig. 4), but this small sample of studies shows that Bouma's constant strongly depends on a number of stimulus, task, and attentional factors, as well as on the orientation of the target and flankers and their location in the visual field [7,9]. Differences in the methods used to calculate b cannot explain the asymmetries in Fig B1.

Collectively, the evidence shows that Bouma's rule is not hard-and-fast. Critical spacing depends on stimulus characteristics, task requirements, and attentional factors. To be sure, critical spacing is a hallmark of crowding, but Bouma's rule should not be used in black-and-white terms, or as a single value (e.g., 0.5). Rather, it should be understood as a continuum, and whether a particular stimulus is crowded should be evaluated based on where it falls on this continuum relative to other similar stimuli under similar conditions.

Box 2: Neurophysiological Studies of crowding

Very few neurophysiological studies have been conducted measuring crowding per se. Motter (*Vision Sciences Society Annual Meeting*, 2002) has found crowding-like effects in V4 neurons, using letter-like stimuli (though Merigan [70] found no effect of V4 lesions on crowding). Many other physiological studies have investigated responses to pairs or groups of stimuli, natural scenes, and visual search arrays (e.g., for reviews, see [98-100]). All of these would likely involve crowding to some extent. However, the stimuli in the majority of the neurophysiological studies were not presented to ensure or tightly control crowding, and so are related in either an uncertain or indirect way. As discussed in the main text, there are very specific diagnostic criteria for what constitutes crowding, and future neurophysiological and imaging studies will need to specifically control and manipulate these factors to isolate the mechanism(s) of crowding.

Box 3: Questions for future research

- Does crowding happen in different modalities (e.g., in audition, touch)? Does it happen crossmodally (visual flankers with an auditory target)? Does optimal cue combination (e.g., between vision and audition) reduce or counteract crowding?
- Is there a benefit of crowding? Ensemble statistics can be perceived in crowds, but are they always? Do crowding and ensemble perception share a common mechanism?
- How and when does crowding limit action? Does the motor system have a compensatory mechanism? Is there a dissociation between the crowding of perception and action?
- What is the developmental timecourse of crowding, and how does it impact the development of visually guided action? How does crowding change over the lifespan?
- Crowding reveals the spatial resolution of conscious vision; many neurodevelopmental disorders are anecdotally reported as having "coarse" grained attention/perception. Is there a significant difference in crowding thresholds in clinical populations with Autism, Fragile X disorder, 22q deletion syndrome, Williams syndrome, or other neurodevelopmental disorders?
- How can crowding be mitigated to improve reading, particularly in those with macular degeneration or amblyopia?

- Does the gradient of crowding as a function of eccentricity differ systematically for different “channels” (e.g., chromatic, 1st- vs 2nd-order, spatial frequency, shapes, letters, object parts, object wholes, etc.)?

Acknowledgments

This work was supported by grants from the National Eye Institute R01EY01728 (DL) and R01EY018216 and NSF CAREER 0748689 (DW). We are grateful to Jason Fischer for helpful comments on an earlier draft of this manuscript.

Glossary

Aftereffect	A delayed or prolonged physiological or psychological response following exposure to a stimulus
Amblyopia	(from the Greek, <i>amblyos</i> —blunt; <i>opia</i> —vision) is a developmental abnormality that results from physiological alterations in the visual cortex and impairs form vision
Anisotropy	A difference in the extent of crowding when measured along different axes
Eccentricity	Refers to where in the visual field, relative to the locus of fixation, a stimulus appears
Flanker	An object or feature that is close to the target, and is irrelevant to the observer’s task. Sometimes referred to as a distractor
Isotropy	Uniformity in all directions
Lateral interaction	The capacity of a feature, object, or neuron to influence the perception (or excitation) of a neighboring feature, object or neuron
Macular degeneration	A condition affecting mostly older adults that results in a loss of central vision due to degenerative changes in the macular (the part of the retina with the highest cone density that is responsible for high visual acuity and reading). It is the leading cause of blindness in Americans over 65
Metacontrast Masking	A type of backward visual masking (see below) in which the visibility of a brief target stimulus is reduced when a second, non-overlapping but adjacent stimulus is presented in quick succession
Visual Masking	The reduction in visibility of a visual stimulus (target) caused by the presentation of a second stimulus, either at the same time as the target (simultaneous masking), before the target (forward masking) or following the target (backward masking)
Surround suppression	The reduction in visibility of a visual stimulus or of neuronal firing rate caused by the presentation of a surrounding stimulus

References

1. Levi DM. Crowding--an essential bottleneck for object recognition: a mini-review. *Vision Res.* 2008; 48:635–654. [PubMed: 18226828]
2. Pelli DG, Tillman KA. The uncrowded window of object recognition. *Nat Neurosci.* 2008; 11:1129–1135. [PubMed: 18828191]

3. Levi DM, et al. Suppressive and facilitatory spatial interactions in peripheral vision: peripheral crowding is neither size invariant nor simple contrast masking. *J Vis.* 2002; 2:167–177. [PubMed: 12678590]
4. Levi DM, et al. Suppressive and facilitatory spatial interactions in amblyopic vision. *Vision Res.* 2002; 42:1379–1394. [PubMed: 12044744]
5. Pelli DG, et al. Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis.* 2004; 4:1136–1169. [PubMed: 15669917]
6. Bouma H. Interaction effects in parafoveal letter recognition. *Nature.* 1970; 226:177–178. [PubMed: 5437004]
7. Toet A, Levi DM. The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Res.* 1992; 32:1349–1357. [PubMed: 1455707]
8. Feng C, et al. Horizontal and vertical asymmetry in visual spatial crowding effects. *J Vis.* 2007; 7:13. 11–10. [PubMed: 18217828]
9. He S, et al. Attentional resolution and the locus of visual awareness. *Nature.* 1996; 383:334–337. [PubMed: 8848045]
10. Liu T, et al. Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. *Curr Biol.* 2009; 19:127–132. [PubMed: 19135367]
11. Banks WP, et al. Asymmetry of Lateral Interference in Visual Letter Identification. *Percept Psychophys.* 1977; 22:232–240.
12. Petrov Y, et al. Crowding and surround suppression: not to be confused. *J Vis.* 2007; 7:12, 11–19. [PubMed: 18217827]
13. Farzin F, et al. Holistic crowding of Mooney faces. *J Vis.* 2009; 9:18, 11–15. [PubMed: 19761309]
14. 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]
15. Bex PJ, et al. The shape and size of crowding for moving targets. *Vision Res.* 2003; 43:2895–2904. [PubMed: 14568377]
16. Chakravarthi R, Cavanagh P. Temporal properties of the polarity advantage effect in crowding. *J Vis.* 2007; 7:11, 11–13. [PubMed: 18217826]
17. Huckauf A, Heller D. On the relations between crowding and visual masking. *Percept Psychophys.* 2004; 66:584–595. [PubMed: 15311658]
18. Ng J, Westheimer G. Time course of masking in spatial resolution tasks. *Optom Vis Sci.* 2002; 79:98–102. [PubMed: 11868853]
19. Enns JT. Object substitution and its relation to other forms of visual masking. *Vision Res.* 2004; 44:1321–1331. [PubMed: 15066393]
20. Bonneh YS, et al. Spatial and temporal crowding in amblyopia. *Vision Res.* 2007; 47:1950–1962. [PubMed: 17502115]
21. Yeshurun Y, Rashal E. Precueing attention to the target location diminishes crowding and reduces the critical distance. *J Vis.* 2010; 10:1–12.
22. Tyler CW, Likova LT. Crowding: a neuroanalytic approach. *J Vis.* 2007; 7:16, 11–19. [PubMed: 18217831]
23. Greenwood JA, et al. Crowding changes appearance. *Curr Biol.* 2010; 20:496–501. [PubMed: 20206527]
24. Levi DM, Carney T. Crowding in peripheral vision: why bigger is better. *Curr Biol.* 2009; 19:1988–1993. [PubMed: 19853450]
25. Parkes L, et al. Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci.* 2001; 4:739–744. [PubMed: 11426231]
26. Baldassi S, et al. Visual clutter causes high-magnitude errors. *PLoS Biol.* 2006; 4:e56. [PubMed: 16494527]
27. Bulakowski PF, et al. Reexamining the possible benefits of visual crowding: dissociating crowding from ensemble percepts. *Atten Percept Psychophys.* 2011 DOI 10.3758/s13414-13010-10086-13412.
28. Livne T, Sagi D. Configuration influence on crowding. *J Vis.* 2007; 7:4, 1–12. [PubMed: 18217819]

29. Felisbert FM, et al. The role of target salience in crowding. *Perception*. 2005; 34:823–833. [PubMed: 16124268]
30. Mareschal I, et al. Cortical distance determines whether flankers cause crowding or the tilt illusion. *J Vis*. 2010; 10:13. [PubMed: 20884588]
31. Song S, Levi DM. Spatiotemporal mechanisms for simple image feature perception in normal and amblyopic vision. *J Vis*. 2010; 10(13):21, 1–22. [PubMed: 21106686]
32. Aghdaee SM. Adaptation to spiral motion in crowding condition. *Perception*. 2005; 34:155–162. [PubMed: 15832566]
33. Whitney D. Motion distorts perceived position without awareness of motion. *Curr Biol*. 2005; 15:R324–326. [PubMed: 15886084]
34. Blake R, et al. Strength of early visual adaptation depends on visual awareness. *Proc Natl Acad Sci U S A*. 2006; 103:4783–4788. [PubMed: 16537384]
35. Dakin SC, et al. Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *J Vis*. 2010; 10:1–16.
36. Greenwood JA, et al. Positional averaging explains crowding with letter-like stimuli. *Proc Natl Acad Sci U S A*. 2009; 106:13130–13135. [PubMed: 19617570]
37. Haberman J, Whitney D. Rapid extraction of mean emotion and gender from sets of faces. *Curr Biol*. 2007; 17:R751–753. [PubMed: 17803921]
38. Popple AV, Levi DM. The perception of spatial order at a glance. *Vision Res*. 2005; 45:1085–1090. [PubMed: 15707915]
39. Strasburger H. Unfocused spatial attention underlies the crowding effect in indirect form vision. *J Vis*. 2005; 5:1024–1037. [PubMed: 16441200]
40. Huckauf A, et al. Semantic processing of crowded stimuli? *Psychol Res*. 2008; 72:648–656. [PubMed: 18841386]
41. Kooi FL, et al. The effect of similarity and duration on spatial interaction in peripheral vision. *Spat Vis*. 1994; 8:255–279. [PubMed: 7993878]
42. Nazir TA. Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Res*. 1992; 32:771–777. [PubMed: 1413560]
43. Andriessen JJ, Bouma H. Eccentric Vision - Adverse Interactions between Line Segments. *Vision Res*. 1976; 16:71–78. [PubMed: 1258390]
44. Hariharan S, et al. “Crowding” in normal and amblyopic vision assessed with Gaussian and Gabor C’s. *Vision Res*. 2005; 45:617–633. [PubMed: 15621179]
45. Chung ST, et al. Spatial-frequency and contrast properties of crowding. *Vision Res*. 2001; 41:1833–1850. [PubMed: 11369047]
46. Gheri C, et al. The relationship between search efficiency and crowding. *Perception*. 2007; 36:1779–1787. [PubMed: 18283928]
47. Kennedy GJ, Whitaker D. The chromatic selectivity of visual crowding. *J Vis*. 2010; 10:15. [PubMed: 20884564]
48. Poder E, Wagemans J. Crowding with conjunctions of simple features. *J Vis*. 2007; 7:23. 21-12. [PubMed: 18217838]
49. Hubbard EM, et al. Individual differences among grapheme-color synesthetes: brain-behavior correlations. *Neuron*. 2005; 45:975–985. [PubMed: 15797557]
50. Banton T, Levi DM. Spatial localization of motion-defined and luminance-defined contours. *Vision Res*. 1993; 33:2225–2237. [PubMed: 8273289]
51. Chung ST, et al. Crowding between first- and second-order letter stimuli in normal foveal and peripheral vision. *J Vis*. 2007; 7:10, 11–13. [PubMed: 18217825]
52. Livne T, Sagi D. How do flankers’ relations affect crowding? *J Vis*. 2010; 10:1, 1–14. [PubMed: 20377278]
53. Malania M, et al. Grouping of contextual elements that affect vernier thresholds. *J Vis*. 2007; 7:1, 1–7. [PubMed: 18217816]
54. Poder E. Crowding, feature integration, and two kinds of “attention”. *J Vis*. 2006; 6:163–169. [PubMed: 16522143]

55. Saarela TP, et al. Global stimulus configuration modulates crowding. *J Vis.* 2009; 9:5, 1–11. [PubMed: 19271915]
56. Saarela TP, et al. The effect of spacing regularity on visual crowding. *J Vis.* 2010; 10:1–7.
57. Louie EG, et al. Holistic crowding: selective interference between configural representations of faces in crowded scenes. *J Vis.* 2007; 7:24. 21-11. [PubMed: 18217839]
58. Cavanagh, P. What's up in top-down processing?. In: Gorea, A., editor. *Representation of vision: Trends and tacit assumptions in vision research.* Cambridge Press; 1991. p. 295-304.
59. Martelli M, et al. Are faces processed like words? A diagnostic test for recognition by parts. *J Vis.* 2005; 5:58–70. [PubMed: 15831067]
60. Chakravarthi R, Cavanagh P. Bilateral field advantage in visual crowding. *Vision Res.* 2009; 49:1638–1646. [PubMed: 19362572]
61. Freeman J, Pelli DG. An escape from crowding. *J Vis.* 2007; 7:22. 21-14. [PubMed: 18217837]
62. Dakin SC, et al. Dissociable effects of attention and crowding on orientation averaging. *J Vis.* 2009; 9:28. 21-16. [PubMed: 20053091]
63. Chakravarthi R, Cavanagh P. Recovery of a crowded object by masking the flankers: determining the locus of feature integration. *J Vis.* 2009; 9:1–9.
64. Wallis TS, Bex PJ. Visual crowding is correlated with awareness. *Curr Biol.* 2011; 21:254–258. [PubMed: 21277208]
65. Flom MC, et al. Contour Interaction and Visual Resolution: Contralateral Effects. *Science.* 1963; 142:979–980. [PubMed: 14069233]
66. Tripathy SP, Levi DM. Long-range dichoptic interactions in the human visual cortex in the region corresponding to the blind spot. *Vision Res.* 1994; 34:1127–1138. [PubMed: 8184557]
67. Pelli DG. Crowding: a cortical constraint on object recognition. *Curr Opin Neurobiol.* 2008; 18:445–451. [PubMed: 18835355]
68. Bi T, et al. The effect of crowding on orientation-selective adaptation in human early visual cortex. *J Vis.* 2009; 9:13, 11–10. [PubMed: 20053076]
69. Motter BC. Modulation of transient and sustained response components of V4 neurons by temporal crowding in flashed stimulus sequences. *J Neurosci.* 2006; 26:9683–9694. [PubMed: 16988039]
70. Merigan WH. Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. *Vis Neurosci.* 2000; 17:949–958. [PubMed: 11193111]
71. Wilkinson F, et al. Lateral interactions in peripherally viewed texture arrays. *J Opt Soc Am A Opt Image Sci Vis.* 1997; 14:2057–2068. [PubMed: 9291601]
72. van den Berg R, et al. A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Comput Biol.* 2010; 6:e1000646. [PubMed: 20098499]
73. Dayan P, Solomon JA. Selective Bayes: attentional load and crowding. *Vision Res.* 2010; 50:2248–2260. [PubMed: 20435055]
74. Neri P, Levi DM. Spatial resolution for feature binding is impaired in peripheral and amblyopic vision. *J Neurophysiol.* 2006; 96:142–153. [PubMed: 16421195]
75. Balas B, et al. A summary-statistic representation in peripheral vision explains visual crowding. *J Vis.* 2009; 9:13, 11–18. [PubMed: 20053104]
76. Bulakowski PF, et al. Visuomotor crowding: the resolution of grasping in cluttered scenes. *Front Behav Neurosci.* 2009; 3:49. [PubMed: 19949462]
77. Vlaskamp BN, Hooge IT. Crowding degrades saccadic search performance. *Vision Res.* 2006; 46:417–425. [PubMed: 15893785]
78. Cohen EH, et al. The relationship between spatial pooling and attention in saccadic and perceptual tasks. *Vision Res.* 2007; 47:1907–1923. [PubMed: 17499833]
79. Meegan DV, Tipper SP. Reaching into cluttered visual environments: spatial and temporal influences of distracting objects. *Q J Exp Psychol A.* 1998; 51:225–249. [PubMed: 9621840]
80. Habak C, et al. Preservation of shape discrimination in aging. *J Vis.* 2009; 9:18, 11–18. [PubMed: 20053109]
81. Farzin F, et al. Spatial resolution of conscious visual perception in infants. *Psychol Sci.* 2010; 21:1502–1509. [PubMed: 20817914]

82. Atkinson J, et al. Visual acuity testing of young children with the Cambridge Crowding Cards at 3 and 6 m. *Acta Ophthalmol (Copenh)*. 1988; 66:505–508. [PubMed: 3218472]
83. Jeon ST, et al. Developmental changes during childhood in single-letter acuity and its crowding by surrounding contours. *J Exp Child Psychol*. 2010; 107:423–437. [PubMed: 20633893]
84. Keita L, et al. Far visual acuity is unremarkable in autism: Do we need to focus on crowding? *Autism Res*. 2010:333–341. [PubMed: 20928845]
85. Martelli M, et al. Crowding, reading, and developmental dyslexia. *J Vis*. 2009; 9:14, 11–18. [PubMed: 19757923]
86. Levi DM, Klein SA. Vernier acuity, crowding and amblyopia. *Vision Res*. 1985; 25:979–991. [PubMed: 4049747]
87. Tripathy SP, Cavanagh P. The extent of crowding in peripheral vision does not scale with target size. *Vision Res*. 2002; 42:2357–2369. [PubMed: 12350424]
88. Schwartz EL. Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Res*. 1980; 20:645–669. [PubMed: 7445436]
89. Levi DM, et al. Vernier acuity, crowding and cortical magnification. *Vision Res*. 1985; 25:963–977. [PubMed: 4049746]
90. Zhang JY, et al. Legibility of Chinese characters in peripheral vision and the top-down influences on crowding. *Vision Res*. 2009; 49:44–53. [PubMed: 18929592]
91. Grainger J, et al. Crowding affects letters and symbols differently. *J Exp Psychol Hum Percept Perform*. 2010; 36:673–688. [PubMed: 20515197]
92. Scolaro M, et al. Spatial attention, preview, and popout: which factors influence critical spacing in crowded displays? *J Vis*. 2007; 7:7, 1–23. [PubMed: 18217822]
93. Vickery TJ, et al. Supercrowding: weakly masking a target expands the range of crowding. *J Vis*. 2009; 9:12, 11–15. [PubMed: 19271922]
94. Chung ST. Learning to identify crowded letters: does it improve reading speed? *Vision Res*. 2007; 47:3150–3159. [PubMed: 17928026]
95. Sun GJ, et al. Ideal observer analysis of crowding and the reduction of crowding through learning. *J Vis*. 2010; 10:1–14.
96. Latham K, Whitaker D. Relative roles of resolution and spatial interference in foveal and peripheral vision. *Ophthalmic Physiol Opt*. 1996; 16:49–57. [PubMed: 8729566]
97. Huckauf A. Task set determines the amount of crowding. *Psychol Res*. 2007; 71:646–652. [PubMed: 16721622]
98. Bichot NP, Desimone R. Finding a face in the crowd: parallel and serial neural mechanisms of visual selection. *Prog Brain Res*. 2006; 155:147–156. [PubMed: 17027386]
99. Gallant, J. Neural mechanisms of natural scene perception. In: L, M.; C, W, JS., editors. *The visual neurosciences*. MIT; 2003. p. 1590-1602.
100. Chelazzi L. Serial attention mechanisms in visual search: a critical look at the evidence. *Psychol Res*. 1999; 62:195–219. [PubMed: 10472200]



Figure 1.

Visual crowding—the deleterious effect of clutter on peripheral object recognition—is ubiquitous in natural scenes. **A.** It seriously impacts virtually all everyday tasks including reading, driving, and interacting with the environment. For example, fixating the bull’s-eye, near the construction zone, note that it is difficult or impossible to recognize the child on the left side of the road, simply because of the presence of the nearby signs. The child on the right, on the other hand, is relatively easier to recognize. **B.** While fixating the crosses, identifying the middle shape, letter, or line orientation—or even the number of tilted lines—is difficult or impossible on the bottom half of the panel. Crowding impairs the ability to recognize and scrutinize objects, but it does not make them disappear; one can see that some *thing* is present in panel (**A**), but it is difficult to identify the thing as a child as opposed to another sign. Crowding defines the spatial resolution of conscious object recognition throughout most of the visual field.

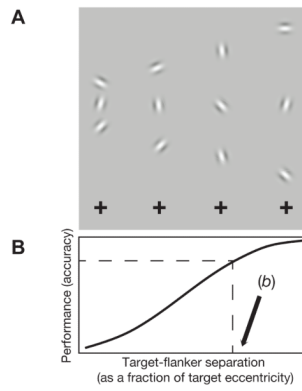


Figure 2.

The critical spacing of crowding and Bouma's proportionality constant (b). **A.** Fixating the crosses along the bottom, notice that the target orientation (central Gabor patch in each column) is easier to recognize on the right. **B.** Performance accuracy increases as the target-flanker separation increases. Bouma's constant, b , may be defined as the target-flank separation (as a ratio of target eccentricity) that results in criterion performance (shown by the dashed line). Although the analytic methods and criteria used to compute b vary from study to study, it generally corresponds to the point at which performance begins to drop as flankers are advanced toward the target.

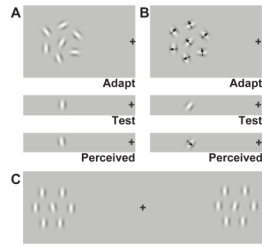


Figure 3.

What escapes crowding? Although crowding limits conscious access to object identities in the periphery, there is much information that gets through or escapes crowding. **A-B.** Adaptation to low-level features. Adapting to the orientation (**A**) or motion (**B**) of a crowded pattern that is unidentifiable nevertheless causes strong local orientation and motion aftereffects, respectively. **C.** While fixating the central cross, notice that the array of Gabor patches on the right appears tilted more clockwise, while the array on the left appears relatively more tilted counterclockwise. In fact, the only difference between the two arrays is the single central Gabor patch; all the flankers are identical. Crowded objects can be unidentifiable, but nevertheless contribute their features to the ensemble or texture.

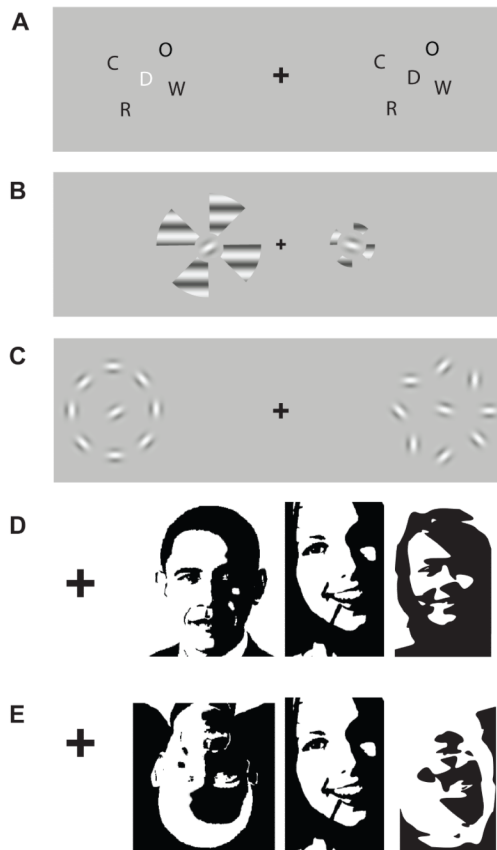


Figure 4.

Crowding can be modulated or released depending on the nature of the flankers, and the target-flanker relationship. **A.** Reversing the contrast polarity of the target (left panel) reduces crowding compared to the right panel. **B.** Flankers of the same size as the target crowd more effectively (more crowding in right panel). **C.** Perceptual grouping of the flankers on the left reduces crowding (more crowding in the right panel). **D-E.** Object-centered, holistic crowding. Faces crowd each other, and upright faces are more effective flankers than inverted ones. Crowding therefore occurs not just between features (Gabor patches, letters), but also between holistic representations of faces).

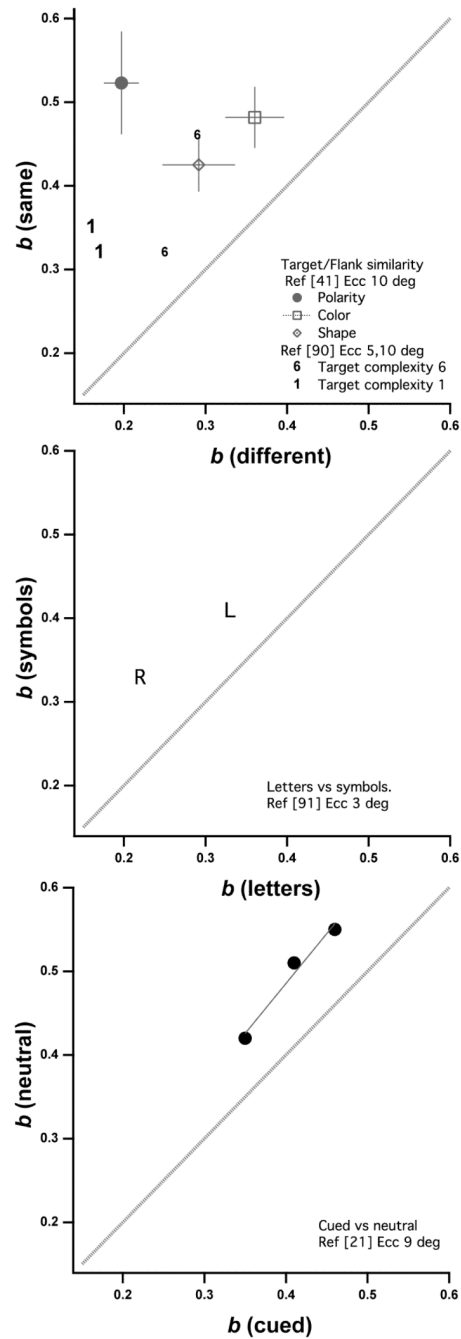


Figure I (Box 1).

Bouma's constant (reflecting the critical spacing between target and distractors) varies depending on the similarity of the targets and flankers, and their complexity (top panel), the type of stimuli used (middle panel), and the attentional requirements of the task (bottom panel). That critical spacing can vary systematically suggests that Bouma's rule is a rule-of-thumb, not a hard-and-fast law. Differences in methods and analytic approaches to calculating (b) cannot explain these graphs. The graphs show a *within*-study modulation of crowding (b), depending on similarity, complexity, stimulus type, and attention.