Crowding: Recent advances and perspectives

Michael H. Herzog
Laboratory of Psychophysics, Brain Mind Institute, Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland

Bilge Sayim
Sciences Cognitives et Sciences Affectives (SCALab), CNRS, UMR 9193, University of Lille, Lille, France
Institute of Psychology, University of Bern, Bern, Switzerland

In crowding, the perception of a target strongly deteriorates in the presence of neighboring elements. Crowding is the standard situation in everyday life since elements are rarely seen in isolation (except in psychophysics laboratories). Crowding is not only crucial for normal object recognition, but also, for example, in reading, visual search, and perhaps numerosity estimation. Hence, any theory of vision needs to account for crowding.

What causes crowding has been debated heavily for more than a century. Not surprisingly, crowding is mainly discussed within the dominant framework of vision: In this classic framework, dating back to the work of Nobel Laureates Hubel and Wiesel, crowding is explained by lateral interactions, such as lateral inhibition, between neurons coding for similar features, for example, orientation and spatial frequency. Another explanation is the pooling of neural responses from lower to higher level neurons, with the latter having larger receptive fields and thus lower resolution, which causes crowding. However, such simplistic models cannot account for a large variety of data, which have shown that crowding depends on the complex spatial and temporal layout of elements across large parts of the visual field. For example, crowding is strong when a target, such as a Gabor, vernier, or letter, is flanked by other Gabors, verniers, or letters, respectively. A release of crowding occurs when more flankers are added that make up a group from which the target ungroups. Simple local approaches fail because the single flankers are contained in the multiflanker configurations (for a review, see Herzog, Sayim, Chicherov, & Manassi, 2015).

These results are more or less well-accepted. However, the mechanisms of crowding are as controversially debated as before, ranging from very basic to highly complex mechanisms (Levi, 2008; Whitney & Levi, 2011; Herzog et al., 2015), as this special issue shows. Because local pooling models have clear problems to explain why complex configurations determine crowding strength, complex pooling models, such as the texture tiling model (TTM), were proposed that pool information in different feature channels across large regions of the visual field (Balas, Nakano, & Rosenholtz, 2009; Rosenholtz, Yu, & Keshvari, 2019). Whereas the TTM is a one-stage model, other models propose that crowding occurs only within perceptual groups, requiring at least two processing stages to account for crowding: first perceptual groups need to be computed and then interference occurs through a different mechanism (Herzog et al., 2015). Attentional accounts of crowding propose that imprecise attention to the target location (Strasburger, 2005) or insufficient attentional resolution, in contrast with sufficient visual resolution, underlie crowding (He, Cavanagh, & Intriligator, 1996).

In this special issue, some of the major questions are where and how crowding occurs in the processing hierarchy (low level vs. high level), whether it occurs in different channels (magnocellular vs. parvocellular; fovea vs. periphery; different color systems), what is the role of other mechanisms, such as attention and redundancy masking, whether crowding needs only one or at least two processing stages (e.g., a grouping stage), and when and under what conditions crowding rules hold and are specific to crowding.

Even though complex interactions seem to prevail in crowding, there are approaches that attempt to explain crowding by low-level mechanisms. For example, Rodriguez and Granger (2021) propose that grouping effects can be explained by a generalized contrast—still a rather simple—mechanism, which does not involve an explicit grouping stage. Instead, basic center-surround mechanisms in the early visual stream underlie their proposed mechanism. Previously, it was argued that local pooling cannot explain crowding but complex pooling models, such as the TTM (Balas et al., 2009), can. As mentioned, the TTM does not require any explicit grouping stage—one stage is sufficient (Rosenholtz et al., 2019). However, Bornet et al. (2021) tested the TTM in six key experimental studies that highlighted high-level effects in crowding. The TTM performed similar to simple pooling mechanisms and was thus not able to account for the results.

https://doi.org/10.1167/jov.22.12.15 ISSN 1534-7362

Reuther, Chakravarthi, and Martinovic (2022) approached the question from a different angle, comparing different psychophysical paradigms. They tested crowding, masking, and grouping with the same stimuli but different task instructions to determine commonalities between tasks. The central idea of their study is that grouping needs large-scale integration to be successful, an operation deleterious to masking and crowding. They found that none of the tasks correlated with each other after normalizing for differences in basic contrast perception. Thus, the different tasks seem to be based on different mechanisms. Crowding seems to be even more complex: Choung, Bornet, Doerig, and Herzog (2021) show that the whole determines performance in crowding and that performance with complex crowding stimuli cannot be predicted by the performance of the parts making up the complex stimuli.

Strong target-flanker grouping usually goes hand in hand with low performance. By contrast, Rummens and Sayim (2022) demonstrate how strong target-flanker grouping (and high target-flanker similarity) decreased crowding: When target-flanker configurations were informative about target identity, performance was better than in uninformative configurations. They further show how emergent features and redundancy masking, the reduction of the number of perceived items in repeating patterns (Sayim & Taylor, 2019; Yildirim, Coates, & Sayim, 2022), could have driven this inversion of the usual effect of target-flanker similarity.

Poder (2020) argues that grouping is a rather vague term, which can be better conceptualized by saliency. He suggests that crowding is a phenomenon concomitant with position-invariant recognition, occurring at all levels of processing, but only within a window determined by attentional selection. Computer simulations of neural networks show evidence for this proposal. Verissimo, Holsken, and Olivers (2021) used a correlation approach and show that slow visual search correlates with larger crowding zones. Their results provide evidence for models of visual search that claim that RTs in serial search are determined by limits of peripheral vision, that is, crowding. The authors propose that search is serial when eye movements are needed to overcome crowding.

Whereas most studies locate the mechanism(s) of crowding somewhere in the brain hierarchy, little research has looked into the different visual pathways. Atilgan, Yu, and He (2020) tailored their stimuli to target the magno(M)- and the parvo(P)-cellular system separately by using different spatial frequencies and color (the M-system is insensitive to color). They found that the M-system is more vulnerable to crowding than the P-system. In addition, form processing is more affected than color and motion processing. One explanation could be the larger receptive fields of the M-system, well in the spirit of the pooling idea, in which larger receptive fields lead to stronger crowding. Lee, Reuther, Chakravarthi, and Martinovic (2021) asked a related question. Isolating S-cone and achromatic luminance mechanisms (and using their combination), they sought to investigate where crowding occurs in the visual processing hierarchy. They found that signatures of crowding only emerged reliably with orientation discrimination above threshold for both, S-cone and achromatic stimuli, suggesting that crowding occurs at a stage where feature information, such as orientation, is already extracted and S-cone and luminance mechanisms are combined.

An important question in crowding research is whether crowding in the fovea is based on the same mechanisms as crowding in the periphery. Although many studies using vernier targets showed similar effects in the fovea and in the periphery (e.g., Sayim, Westheimer, & Herzog, 2010; Manassi, Sayim, & Herzog, 2012), results for letter targets are less clear. Coates, Jiang, Levi, and Sabesan (2022) investigated crowding with tumbling Es flanked by bars at several eccentricities in the parafovea. Using adaptive optics, they found that the different results at different eccentricities could be well-summarized by a single function, unifying the varying results. They suggest that common mechanisms underlie crowding across the visual field, including near the fovea.

Possibly the most central characteristics of crowding is its dependence on the distance between the target and the flankers: Flankers closer to the target usually yield worse performance than flankers farther away. However, whether this relation holds, strongly depends on the entire target-flanker configuration (Herzog, Sayim, Chicherov, & Manassi, 2015). A recent study showed that the relationship was even inverted when emergent features of target-flanker configurations were informative about target identity (Melnik, Coates, & Sayim, 2018). The general effect of target-flanker distance is often summarized by Bouma's law, which states that flankers start to interfere with target perception at a distance of approximately 0.5 times the eccentricity of the target. Coates, Ludowici, and Chung (2021) review and reanalyze a large number of data sets to shed light on the varied findings in different studies. They report how Bouma's fraction varies with contrast and duration, and introduce a model that captures differences of the results of several crowding paradigms. Based on their results, they give detailed recommendations on what factors should be considered when conducting crowding experiments.

Chakravarthi, Rubruck, Kipling, and Clarke (2021) investigated another characteristic of crowding—its inward–outward asymmetry where a flanker placed on the target's side further from fixation (outside) interferes more strongly with target perception than a flanker placed closer to fixation (inside). They found a clear inward–outward asymmetry at all four tested cardinal locations in three different cueing conditions (blocked, randomized, and precued). Their results confirm that the inward–outward asymmetry is a stable characteristic of crowding.

Another characteristic of crowding is its visual field dependence. Yildirim et al. (2022) show how redundancy masking, the decrease in the number of perceived items in repeating patterns, has different visual field asymmetries than crowding, adding to the increasing evidence that—although related to crowding—redundancy masking is a unique phenomenon. In particular, they found that, in contrast with crowding, redundancy masking was stronger on the horizontal than on the vertical meridian and did not differ between the upper and lower visual fields. They suggest that these visual field asymmetries are linked to visual field–dependent capacities to extract stimulus regularity and variations of compression of visual space.

Another avenue to investigate crowding is to study to what extent we can modify brain processing by transcranial direct current stimulation to decrease crowding. Chen, Zhu, He, and Fang (2021) used transcranial direct current stimulation, and showed that contralateral 20-minute stimulation can alleviate crowding of an orientation discrimination and letter identification task. Sham and ipsilateral stimulation showed no effects.

Another crucial question is to what extent crowding can be overcome by learning, a question particularly important for rehabilitation. Plank et al. (2021) show, in accordance with previous studies, that perceptual learning can improve performance in crowding tasks. Transfer to tumbling Es occurred when healthy observers trained with Landolt C, but not the other way around. Hence, for rehabilitation the right choice of stimuli might be crucial. Interestingly, learning can be specific for the target and the flankers. Eberhardt, Pittino, and Huckauf (2021) investigated how the presentation depth of target and flankers modulated the emotional conditioning effect where performance deteriorates when flanker features are negatively conditioned. They presented target and flankers in different (real) depth planes, and found that negatively conditioned flankers (but not targets) yielded worse performance than neutrally conditioned flankers when the flankers were presented closer than or at fixation depth, suggesting that emotional conditioning of stimulus parts does not only depend on whether they are targets or flankers but also at what relative position they are presented.

Learning to "uncrowd" is particularly important in clinical contexts. Tailor, Theodorou, Dahlmann-Noor, Dekker, and Greenwood (2021) investigated whether increased foveal crowding in idiopathic infantile nystagmus syndrome resembles foveal crowding in amblyopia. In particular, they asked if nystagmic crowding is due to variations of stimulus locations because of involuntary eye movements or due to neural changes similar to amblyopia. They found increased foveal crowding compared to healthy controls, and—in contrast with amblyopic crowding—stronger increases with horizontal compared with vertical flankers as predicted by an account based on eye movements. Hence, they suggest that increased foveal crowding in idiopathic infantile nystagmus syndrome is due to the involuntary eye movements.

Eye movements have also been proposed to be linked to typical peripheral crowding. In this special issue, Raveendran, Krishnan, and Thompson (2020) investigated the question of whether fixation stability is linked to poor performance in peripheral crowding. They found that fixation stability was decreased when performing a peripheral task compared to foveally presented stimuli; however, there was no link between crowding strength and fixation stability, suggesting that fixation stability is not a contributor to poor performance in peripheral crowding.

Taken together, the collection of articles in this special issue is another step to better understanding crowding, the interaction of multiple elements in the visual system, and visual perception in general. A plethora of different approaches, varying perspectives and highly interesting findings promise to contribute to—and shape—the future of crowding research, helping to advance our understanding of how we see.

References

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

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision, 9*(12):13, 1–18, http://journalofvision.org/9/12/13/, http://doi.org/10.1167/9.12.13.

Bornet, A., Choung, O.-H., Doerig, A., Whitney, D., Herzog, M. H., & Manassi, M. (2021). Global and high-level effects in crowding cannot be predicted by either high-dimensional pooling or target cueing. *Journal of Vision, 21*, 10, https://doi.org/10.1167/jov.21.12.10.

Chakravarthi, R., Rubruck, J., Kipling, N., & Clarke, A. D. F. (2021). Characterizing the in-out asymmetry in visual crowding. *Journal of Vision, 21*, 10, https://doi.org/10.1167/jov.21.11.10.

Chen, G., Zhu, Z., He, Q., & Fang, F. (2021). Offline transcranial direct current stimulation improves the ability to perceive crowded targets. *Journal of Vision, 21*, 1, https://doi.org/10.1167/jov.21.2.1.

Choung, O. H., Bornet, A., Doerig, A., & Herzog, M. H. (2021). Dissecting (un)crowding. *Journal of Vision, 21*, 10, https://doi.org/10.1167/jov.21.10.10.

Coates, D. R., Jiang, X., Levi, D. M., & Sabesan, R. (2022). Cortical distance unifies the extent of parafoveal contour interactions. *Journal of Vision, 22*, 15, https://doi.org/10.1167/jov.22.2.15.

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

Eberhardt, L. V., Pittino, F., & Huckauf, A. (2021). Close - but not distant - conditioned flanker emotion affects crowding. *Journal of Vision, 21*, 22, https://doi.org/10.1167/jov.21.8.22.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334–337.

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

Lee, R. J., Reuther, J., Chakravarthi, R., & Martinovic, J. (2021). Emergence of crowding: The role of contrast and orientation salience. *Journal of Vision, 21*, 20, https://doi.org/10.1167/jov.21.11.20.

Levi, D. M. (2008). Crowding-an essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654.

Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision, 12*(10), 13, http://www.journalofvision.org/content/12/10/13, http://doi.org/10.1167/12.10.13.

Melnik, N., Coates, D. R., & Sayim, B. (2018). Emergent features in the crowding zone: When target-flanker grouping surmounts crowding. *Journal of Vision, 18*(9), 19, 1–15, https://doi.org/10.1167/18.9.19.

Plank, T., Lerner, L., Tuschewski, J., Pawellek, M., Malania, M., & Greenlee, M. W. (2021). Perceptual learning of a crowding task: Effects of anisotropy and optotype. *Journal of Vision, 21*, 13, https://doi.org/10.1167/jov.21.11.13.

Poder, E. (2020). Crowding and attention in a framework of neural network model. *Journal of Vision, 20*, 19, https://doi.org/10.1167/jov.20.13.19.

Raveendran, R. N., Krishnan, A. K., & Thompson, B. (2020). Reduced fixation stability induced by peripheral viewing does not contribute to crowding. *Journal of Vision, 20*, 3, https://doi.org/10.1167/jov.20.10.3.

Reuther, J., Chakravarthi, R., & Martinovic, J. (2022). Masking, crowding, and grouping: Connecting low and mid-level vision. *Journal of Vision, 22*, 7, https://doi.org/10.1167/jov.22.2.7.

Rodriguez, A., & Granger, R. (2021). On the contrast dependence of crowding. *Journal of Vision, 21*, 4, https://doi.org/10.1167/jov.21.1.4.

Rosenholtz, R., Yu, D., & Keshvari, S. (2019). Challenges to pooling models of crowding: Implications for visual mechanisms. *Journal of Vision, 19*(7), 15.

Rummens, K., & Sayim, B. (2022). Multidimensional feature interactions in visual crowding: When configural cues eliminate the polarity advantage. *Journal of Vision, 22*, 2, https://doi.org/10.1167/jov.22.6.2.

Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science, 21*(5), 641–644, http://www. ncbi.nlm.nih.gov/pubmed/20483840

Sayim, B., & Taylor, H. (2019). Letters lost: Capturing appearance in crowded peripheral vision reveals a new kind of masking. *Psychological Science, 30*(7), 1082–1086, https://doi.org/10.1177/0956797619847166.

Strasburger, H. (2005) Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*, 8–8.

Tailor, V. K., Theodorou, M., Dahlmann-Noor, A. H., Dekker, T. M., & Greenwood, J. A. (2021). Eye movements elevate crowding in idiopathic infantile nystagmus syndrome. *Journal of Vision, 22*, 9, https://doi.org/10.1167/jov.21.13.9.

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

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Science, 15*(4), 160–168.

Yildirim, F. Z., Coates, D. R., & Sayim, B. (2022). Atypical visual field asymmetries in redundancy masking. *Journal of Vision, 22*, 4, https://doi.org/10.1167/jov.22.5.4.