

qualitative contrast to the experimental data. Clearly, human perception in this paradigm is not strictly linear and, thus, performance on the entire stimulus cannot be predicted by a simple summation of the performance levels of its parts.

We propose that the overall stimulus configuration and grouping play a crucial role in crowding [8,9]. Specifically, when the target *ungroups* from the flankers, performance is improved, and only when the target and flankers *group* is crowding strong. However, grouping does not explain why performance deteriorates in crowding. Instead, grouping specifies which elements are prone to crowd each other (see [8], p12, point 6). For this reason, we propose that a unified model of crowding needs both a grouping stage and a mechanism to account for the detrimental effects, such as the one proposed by Harrison and Bex [3]. Further research will explore how such a model may be constructed.

ACKNOWLEDGEMENTS

We thank Harrison and Bex for sharing their model code. This work was supported by the Swiss National Science Foundation (SNF) Sinergia project: "Learning from delayed and sparse feedback".

REFERENCES

1. Pelli, D.G., Palomares, M., and Majaj, N.J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *J. Vis.* 4, 1136–1169.
2. Strasburger, H., Harvey, Jr., L.O., and Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Percept. Psychophys.* 49, 495–508.
3. Harrison, W.J., and Bex, P.J. (2015). A unifying model of orientation crowding in peripheral vision. *Curr. Biol.* 25, 3213–3219.
4. Banks, W.P., Larson, D.W., and Prinzmetal, W. (1979). Asymmetry of visual interference. *Percept. Psychophys.* 25, 447–456.
5. Pöder, E. (2006). Crowding, feature integration, and two kinds of "attention". *J. Vis.* 6, 163–169.
6. Saarela, T.P., Sayim, B., Westheimer, G., and Herzog, M.H. (2009). Global stimulus configuration modulates crowding. *J. Vis.* 9, 1–11.
7. Levi, D.M., and Carney, T. (2009). Crowding in peripheral vision: why bigger is better. *Curr. Biol.* 19, 1988–1993.
8. Herzog, M.H., Sayim, B., Chicherov, V., and Manassi, M. (2015). Crowding, grouping, and object recognition: A matter of appearance. *J. Vis.* 15, 1–18.
9. Manassi, M., Sayim, B., and Herzog, M.H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *J. Vis.* 12, 1–14.

Laboratory of Psychophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.
E-mail: matt.pachai@epfl.ch

Correspondence

Reply to Pachai *et al.*

William J. Harrison^{1,2}
and Peter J. Bex³

Peripheral vision is fundamentally limited by the spacing between objects. When asked to report a target's identity, observers make erroneous reports that sometimes match the identity of a nearby distractor and sometimes match a combination of target and distractor features. The classification of these errors has previously been used to support competing 'substitution' [1] or 'averaging' [2] models of the phenomenon known as 'visual crowding'. We recently proposed a single model in which both classes of error occur because observers make their reports by sampling from a biologically-plausible population of weighted responses within a region of space around the target [3]. It is critical to note that there is no probabilistic substitution or averaging process in our model; instead, we argue that neither substitution nor averaging occur, but that these are misclassifications of the distribution of reports that emerge when a population response distribution is sampled. This is a fundamentally different way of thinking about crowding, and on this basis we claim to have provided a mechanism unifying categorically distinct perceptual errors. Our goal was not to model all crowding phenomena, such as the release from crowding when target and flankers differ in color or depth [4]. Pachai *et al.* [5] have suggested that our model is not unifying because it inaccurately predicts perceptual performance for a particular stimulus. Although we agree that our model does not predict their data, this specific demonstration overlooks the critical aspect of the model: perceptual reports are drawn from a weighted population code. We show that Pachai *et al.*'s [5] own data actually provide evidence for the population code we have described [3], and we suggest a biologically-plausible analysis of their stimuli that

provides a computational basis for their 'grouping' account of crowding.

In both Pachai *et al.*'s [5] work and our original study [3], following the presentation of a randomly oriented target Landolt C in peripheral vision, observers adjusted a foveal Landolt C so that it matched the target orientation. The target was presented alone (Figure 1A, top panel), or surrounded by a larger and independently oriented Landolt C (Figure 1B, top panel). In close agreement with our data [3], the mean error in their observers' reports is greater with the flanker than without. They further show that our model provides a good fit to these behavioural data. Pachai *et al.* [5] included a second flanker condition, in which the target was flanked by five concentric distractors, all with the same orientation (Figure 1C, top panel). For this condition, our model generally predicts that observers' performance should be worse than in the one-flanker condition. In contrast to this prediction, performance improved in this new condition. Pachai *et al.* [5] conclude that our model is fundamentally limited because it predicts crowding according to the distance between target and flankers. They advance a 'grouping' explanation, in which crowding is released because the flankers are somehow grouped independently of the target. Notice, however, they still found crowding even in their grouped-flanker condition. In Figure S1 in the Supplemental Information, we provide a demonstration that shows, consistent with our approach, crowding in this condition also relies on the distance between target and inner flanker. Nevertheless, we believe that the conclusions of Pachai *et al.* [5] overlook the critical aspect of our model.

In our model [3], populations of neurons coding a target's orientation also code information about flanking distractors and this contamination leads to perceptual errors (details given in [3]). Shown in the top panel of Figure 1E are the trial-by-trial errors made by observers in the one-flanker condition of Pachai *et al.* [5]. These data are very well captured by our population code model [3]. We think that a simple extension of

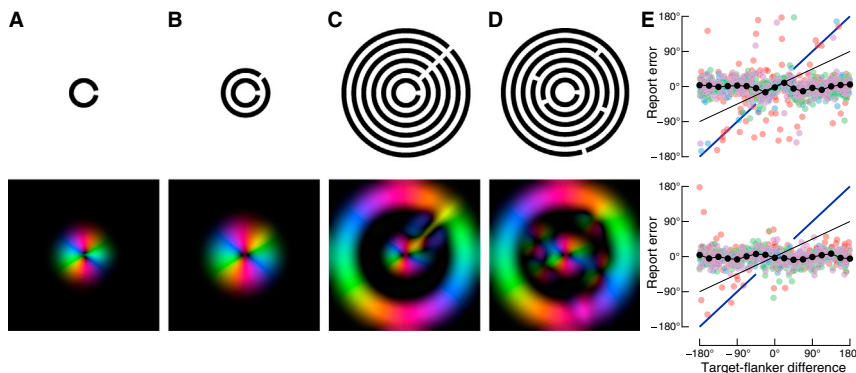


Figure 1. Our texture segmentation analysis and raw data of Pachai *et al.* [5].

(A–C) Top panels show the stimuli used by Pachai *et al.* [5] and bottom panels show the outputs of standard texture analyses. Colours in the bottom panel reveal the peak orientation of the filters used to detect a texture boundary (see text for more details). (D) Same as (C), but flankers have random orientations. (E) Raw data of Pachai *et al.* [5]. Top and bottom panels show one-flanker and five-flanker data, respectively. Different colours represent different observers, black points show the median error of all observers in 22.5° bins, with standard errors smaller than point size. The fine black line has a slope of 0.5, which corresponds to reports that might be (mis)classified as averaging; the heavy blue line has a slope of 1, which corresponds to reports that might be (mis)classified as substitutions. Our population code model captures all errors in the one-flanker condition. The same model can account for the five-flanker condition if we assume that texture segregation minimises the contribution of the flankers to the population code.

the population code model will also account for the novel five-flanker condition presented by Pachai *et al.* [5], although it will require modelling beyond the aims of our original paper and the scope of this correspondence.

The bottom panels of Figure 1A–C show the output of a standard filter-rectify-filter model of early visual texture processing [6]. Based on very simple and biologically-plausible filter-responses, these analyses reveal regions of common texture and texture change. Colours show texture boundaries (locations where one texture is segmented from another), whereas areas of black show regions of the image with a constant texture. For the uncrowded condition, the primary texture boundary contains the target; for the one-flanker condition, this texture boundary covers both target and flanker. In the five-flanker condition, separate texture boundaries occur for the target, the flanking gaps and the outer-most flanker. In our model, the population response is weighted by spatial separation between target and flanks. In an analogous way, texture boundaries could be used to weight the responses in a population code. To test this prediction, we generated a novel variant of the stimulus used by Pachai *et al.* [5] in which we randomized the

orientation of each flanker (Figure 1D, top panel). As shown in Figure 1D (bottom panel), there is no clear texture boundary generated by the flanker gaps, and we thus predict a relative increase in crowding in such a condition. Informal inspection of the figure suggests this is the case. We speculate that our approach is also compatible with models of peripheral vision in which images are encoded by higher-level image statistics within an area that increases with eccentricity [7,8]. Thus, observers' reports are statistical samples from the population response within regions of common texture. We think this hypothesis will go far to account for other recent data from Herzog and colleagues (see Figure S2).

Pachai *et al.* [5] argue that an unspecified 'grouping' process provides a better explanation of crowding than do models in which the strength of interactions depends on parameters like the distance between target and flankers. Our texture analysis suggests that the flankers may not be grouped in the manner described by Pachai *et al.*, but that the aligned gaps in fact define a texture boundary that segments this region from the rest of the image. Finally, we would like to comment that "grouping" explanations of crowding are circular:

it is assumed that when crowding is reduced, flankers must therefore be grouped; thus, grouping is measured by the degree of crowding and vice-versa. We argue that, in contrast to this approach, a weighted population code provides a testable mechanistic account for these data. An important advantage of our model is that we can now predict perceptual variability in crowding, rather than merely predicting when crowding will or will not occur.

SUPPLEMENTAL INFORMATION

Supplemental Information includes two Supplemental figures and a description of our texture analysis and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2016.03.024>.

ACKNOWLEDGMENTS

We thank Pachai *et al.* for sharing their raw data. This work was supported by NIH grant R01EY021553 (P.J.B.) and a National Health and Medical Research Council of Australia CJ Martin Fellowship (APP1091257; W.J.H.).

REFERENCES

1. Strasburger, H., and Malania, M. (2013). Source confusion is a major cause of crowding. *J. Vis.* 13, 24.
2. Parkes, L., Lund, J., Angelucci, A., Solomon, J., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat. Neurosci.* 4, 739–744.
3. Harrison, W.J., and Bex, P.J. (2015). A unifying model of orientation crowding in peripheral vision. *Curr. Biol.* 25, 3213–3219.
4. Kooi, F., Toet, A., Tripathy, S., and Levi, D.M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spat. Vis.* 8, 255–279.
5. Pachai, M.V., Doerig, A.C., and Herzog, M.H. (2016). How best to unify crowding? *Curr. Biol.* 26, R352–R353.
6. Bergen, J.R., and Landy, M.S. (1991). Computational modeling of visual texture segregation. In *Computational Models of Visual processing*, M.S. Landy and J.A. Movshon, eds. (Cambridge, MA: MIT Press), pp. 253–271.
7. Balas, B., Nakano, L., and Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *J. Vis.* 9, 13.
8. Freeman, J., and Simoncelli, E.P. (2011). Metamers of the ventral stream. *Nat. Neurosci.* 14, 1195–1201.

¹Department of Psychology, University of Cambridge, Cambridge CB2 3EB, UK.

²Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia.

³Department of Psychology, Northeastern University, Boston, MA 02115, USA.

E-mail: willjharri@gmail.com (W.J.H.),

p.bex@neu.edu (P.J.B.)