Supplemental Information: Reply to Pachai *et al.*

William J Harrison

Peter J Bex

Supplemental Figures

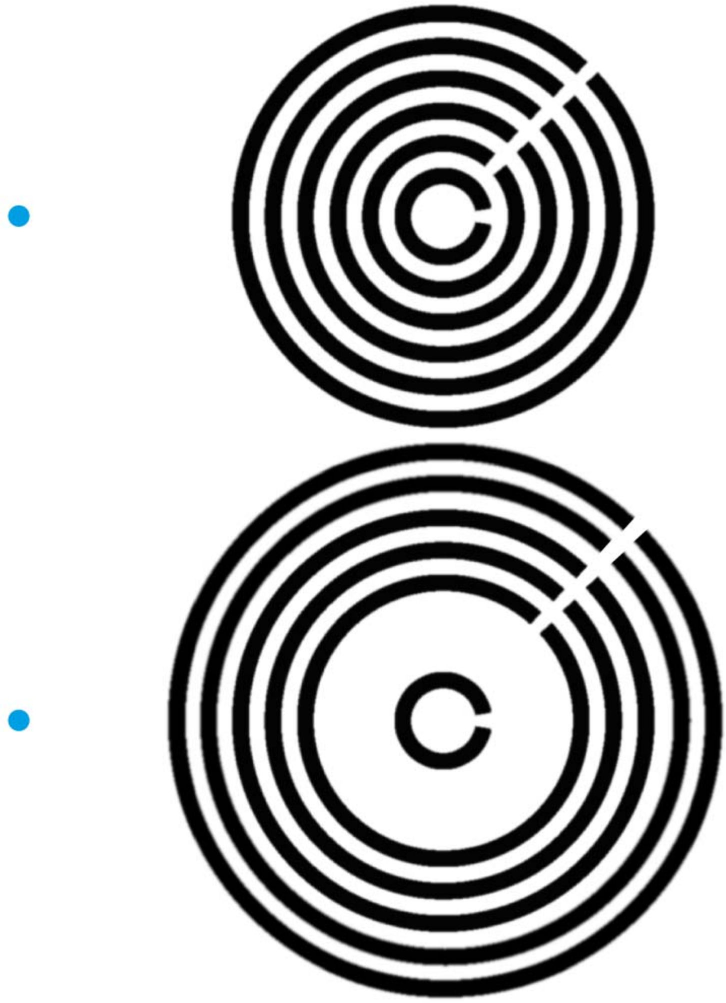

Figure S1. A demonstration that spacing matters. Pachai et al argue that the visibility of the central C in the top row is better explained by the degree to which the surrounding flankers are grouped than by the distance between target and flankers. In the bottom row, we increase the spacing between the target and nearest flanker; the apparent grouping cues within the flanks are unchanged. The reader can fixate each blue spot in turn and experience the improved target visibility when target-flanker spacing increases. We note that a release from crowding due to spatial distance is predicted even in a Gestalt grouping account of crowding [S1], and so this demonstration shows that space matters, but not necessarily that grouping doesn't. We suggest observers' perceptual errors in both cases are captured by a population code and texture segmentation.

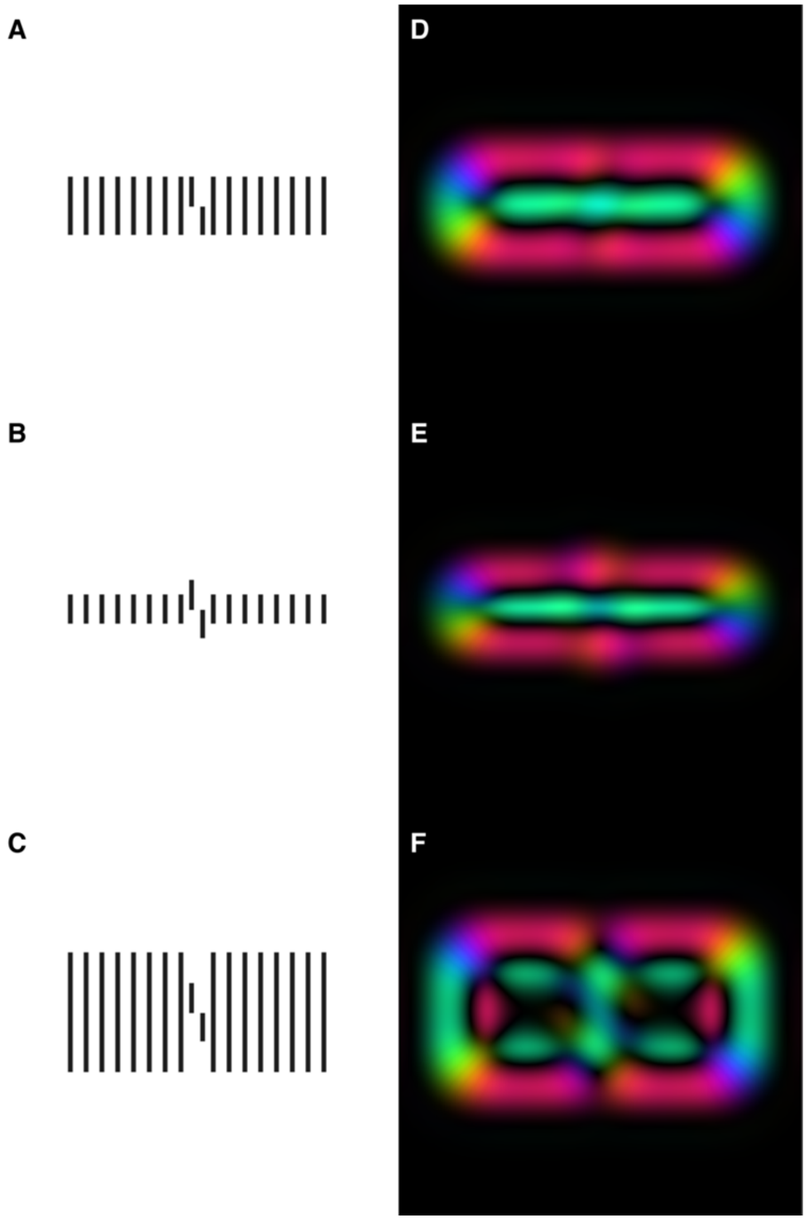

Figure S2. A texture analysis of crowded Vernier stimuli used by Herzog and colleagues [S2]. (A-C) Each stimulus is presented in peripheral vision, and an observer's task it to report the direction of offset of the centre bars. Observers' performance is worst in (A) and best in (C), with intermediate performance in (B). (D-F) Texture analysis of the stimuli. According to these analyses, the target elements form a common texture with the flanking distractors in (A) and to a lesser extent in (B), but not in (C).

Supplemental Experimental Procedures

Texture model

We used a standard filter-rectify-filter model of texture segmentation to analyse each stimulus (top panels, Fig. 1A-D). We first created stimulus images, each with a width that was slightly larger than the full width of the largest flanker in the five-flanker condition. The target was centred in the image, and had a radius of approximately 8% of the image width. For filtering, we used phase-invariant steerable filters (spatial frequency = 4 cycles/image) to detect local spectral energy at 16 orientations (0° to 168.75° in 11.25° steps). The full-wave rectified (absolute value) output for each orientation was again filtered with a filter of the same orientation but half the original spatial frequency (2 cycles/image). We then found the opponent energy [S3] by taking the energy of each orientation and subtracting the energy of the orthogonal orientation (i.e., the energy for the filter rotated by 90° from the current filter). Resulting negative values, which represent the orthogonal filter responses, were then set to zero. This process was repeated for all orientations, such that the opponent energies of two orthogonal filters were mutually exclusive. Finally, we normalised the opponent energies by dividing each orientation's energy by the sum of energy across all orientations. The data shown in the bottom panels of Figure 1A-D represents the mean of all orientations, normalised so the greatest model output has the most luminant colour.

Supplemental References

S1. Kubovy, M., Holcombe, A. O., and Wagemans, J. (1998). On the lawfulness of grouping by proximity. Cognitive Psychology *35*, 71–98.

S2. Herzog, M. H., and Manassi, M. (2015). Uncorking the bottleneck of crowding: a fresh look at object recognition. Current Opinion in Behavioral Sciences *1*, 86–93.

S3. Bergen, J. R., and Landy, M. S. (1991). Computational modeling of visual texture segregation. Computational models of visual processing.