Object Recognition: Visual Crowding from a Distance

Immediately before a large eye movement, a target object is crowded by clutter near the target's future location. This new finding, from a recent study, shows that the brain's remapping for the anticipated eye movement unavoidably combines features from the target's current and future retinal locations into one perceptual object.

Denis G. Pelli[1] and Patrick Cavanagh[2]

Object recognition proceeds through the selection and combination of features, governed by rules of grouping and crowding [1–5]. This is variously called binding, grouping, region growing, or crowding, depending partly on whether the observer recognizes the object. In 'crowding', target identification fails because the feature combination extends unavoidably over an inappropriately large area, jumbling elements of the adjacent clutter with those of the target, which wrecks identification [6,7]. The size of this minimum area over which features are combined is roughly the same for all simple objects, like letters, and grows proportionally with eccentricity (distance from fixation) [5,8,9]. Transformed by the cortical magnification factor, the minimum size of the combination region corresponds to a fixed area in mm^2 on the surface of the cortex [10,11]. This compact region of compulsory combination is roughly centered on the target. In their recent *Current Biology* paper, Harrison *et al.* [12] reported their finding that, immediately before a large eye movement (saccade), features from *two* widely separated regions, at the *present* and *future* locations of the target, are combined to produce *one* perceptual object.

In their study, Harrison *et al.* [12] presented target and clutter briefly, immediately before a large saccade (Figure 1). When clutter was placed next to the *future* retinal location of the target — far away, in the other hemifield — the clutter impaired recognition nearly as strongly as when it was placed next to the target's *current* location. As in normal crowding, the strength of the effect was found to depend on similarity between the target and the clutter [13]: an L crowds an H, but an X does not crowd an O. The sensitivity to similarity shows that the impairment is due to crowding, not to the indiscriminate impairment of saccadic suppression [14,15].

The story of remapping began with physiological studies of cells that direct saccades to their targets, showing a correction in their target locations for an upcoming eye movement [16]. It was then extended to the allocation of attention, where a similar correction for upcoming saccades was found [17,18]. In their new work, Harrison *et al.* [12] have used crowding to show that remapping displaces not just location information, but also content, effectively combining the content of two distant regions.

The discovery of remapping by Duhamel, Colby, and Goldberg in 1992 launched intense scientific activity evaluating the mechanisms, brain areas, and pathways involved. They

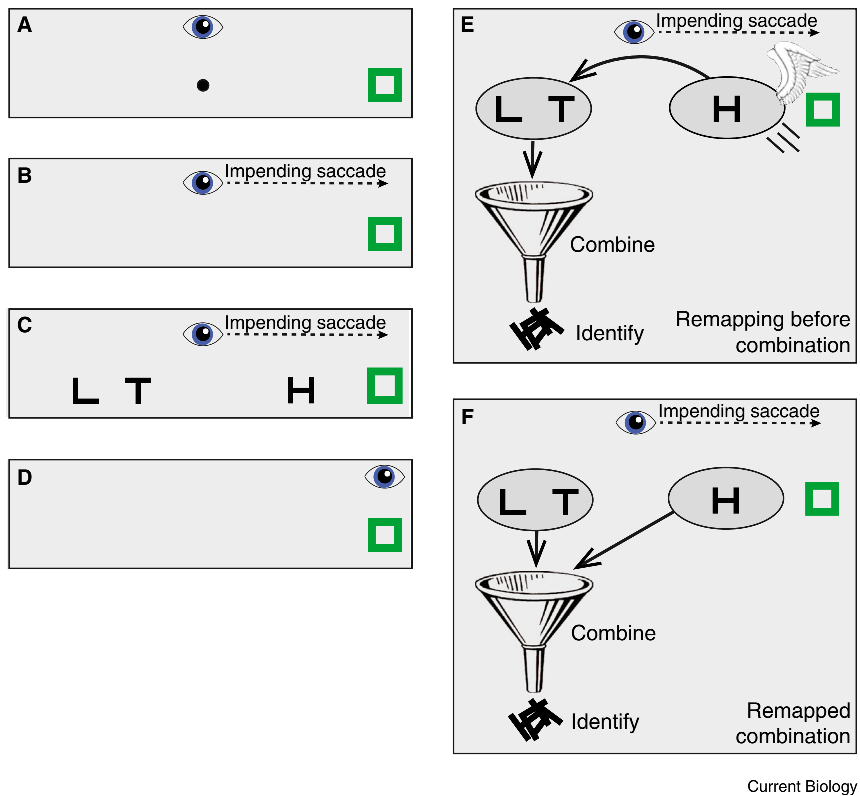

Figure 1. Visual crowding at a distance before a saccade.

(A) Initially, the observer fixates the central dot. (B) Disappearance of that central dot cues the observer to move his or her eyes rightward to the green square. (C) The random target (shown as H) is briefly presented to the right of fixation. At the same time, clutter is briefly presented to the left of fixation. The duration is 17 ms. (D) The brief target and clutter are gone by the time the eye moves and finally arrives at the green square. (Our simplified diagram omits the other placeholders and the changing content of the green square, which are not discussed here.) (E) Remapping before Combination: in this, the target's features are remapped from its original retinotopic location to its anticipated post-saccadic location; that places the target amid the clutter. Identification unavoidably combines information over a region that includes the (displaced) target and clutter. Each combination region, the critical zone for crowding, is represented by a gray oval. Combining the target with clutter wrecks identification, despite the large offset between the target and clutter in the display. (F) Remapped Combination: in this, the remapping causes identification to be based on a combination of features from two locations: the present and future locations of the target [19]. Again, combining the features of the target and clutter wrecks identification. Based on Harrison *et al*. [12]. (Funnel image by permission; from Merriam-Webster's Learner's Dictionary ©2013 www.learnersdictionary.com by Merriam-Webster, Inc. www.Merriam-Webster.com) found that almost all the neurons in lateral intraparietal cortex (LIP) begin to fire in response to a stimulus that will be brought into their respective fields by an eye movement, even if the stimulus is extinguished before the eyes arrive. This remapped response is only found for attended targets, either flashed or task-relevant (for example, [19]), and always moves in the direction opposite to that of the saccade, predicting where the attended target will be after the saccade. The predictive response of these neurons can begin as early as 100 ms before the saccade and tends to peak at the onset of the saccade, much earlier than the cell would be able to respond if the stimulus simply appeared in the cell's receptive field following the eye movement. This pre-saccadic stimulus activates the cells for the remapped location (which depends on the saccade vector) and, at the same time, it also activates cells that normally respond to the target's location. As a result, activity can be seen just before the saccade in two widely separated sets of cells, at the target's actual retinal location and its remapped location, both in response to the same brief stimulus.

These physiological studies were followed by behavioral studies which showed a similar pre-saccadic shift of attention by placing probes [17] at the target's remapped location just before a saccade. Moreover, there is compelling evidence from time stamping and masking, and now crowding, showing that remapping does not just displace location information but generates a perceived target object that combines target information from the target's pre- and (expected) post-saccadic retinal locations. In time stamping, observers are asked to saccade to a clock with rapidly spinning hands and report the time that their eyes arrive, but they actually report times 40–60 ms earlier [20]. This shows that what the observers think they see at fixation actually incorporates imagery from the peripheral retinal location of the clock, as if observers were already looking at it. Masking and crowding are very different [7]. Masking requires overlap; crowding requires proximity. Pre-saccadic masking is effective at the target's current and future retinal locations. Without a saccade, a mask must overlap the target to impair its visibility, but immediately before a saccade a mask is effective at the target's post-saccadic location [18].

These 'bi-local' results are now buttressed by the Harrison *et al.* [12] finding of remapped crowding. As already noted, without a saccade, clutter must be near the target to crowd it, but, immediately before a saccade, clutter is also effective at the target's post-saccadic retinal location [12]. Because the range of masking and crowding is fairly small, and the size of saccades is quite variable, it would be interesting to measure the spatial profile of remapping to discover how accurately the pre-saccadic remapping predicts the saccade. Thus, Harrison *et al.* [12] have shown that the brain's remapping for the anticipated eye movement unavoidably combines features from the current and future retinal locations of the target into one perceptual object. (See Figure 1 for two proposals on how this might happen.)

References

1. Pelli, D.G., Burns, C.W., Farell, B., and Moore-Page, D.C. (2006). Feature detection and letter identification. Vision Res. *46*, 4646–4674.
2. Suchow, J.W., and Pelli, D.G. (2013). Learning to detect and combine the features of an object. Proc. Natl. Acad. Sci. USA *110*, 785–790.
3. Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II [Laws of organization in perceptual forms]. Psychologische Forschung *4*, 301–350.
4. Levi, D.M. (2008). Crowding–an essential bottleneck for object recognition: a mini-review. Vision Res. *48*, 635–654.
5. Pelli, D.G., and Tillman, K.A. (2008). The uncrowded window of object recognition. Nat. Neurosci. *11*, 1129–1135.
6. Stuart, J.A., and Burian, H.M. (1962). A study of separation difficulty: Its relationship to visual acuity in normal and amblyopic eyes. Am. J. Ophthalmol. *53*, 471–477.
7. 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.
8. Bouma, H. (1970). Interaction effects in parafoveal letter recognition. Nature *226*, 177–178.
9. Toet, A., and Levi, D.M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. Vision Res. *32*, 1349–1357.
10. Pelli, D.G. (2008). Crowding: A cortical constraint on object recognition. Curr. Opin. Neurobiol. *18*, 445.
11. Freeman, J., and Simoncelli, E.P. (2011). Metamers of the ventral stream. Nat. Neurosci. *14*, 1195–1201.
12. Harrison, W.J., Retell, J.D., Remington, R.W., and Mattingley, J.B. (2013). Visual crowding at a distance during predictive remapping. Curr. Biol. *23*, 793–798.
13. Kooi, F.L., Toet, A., Tripathy, S.P., and Levi, D.M. (1994). The effect of similarity and duration on spatial inter-action in peripheral vision. Spatial Vision *8*, 255–279.
14. Burr, D.C., Morrone, M.C., and Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. Nature *371*, 511–513.
15. Castet, E., and Masson, G.S. (2000). Motion perception during saccadic eye movements. Nat. Neurosci. *3*, 177–183.
16. Duhamel, J.R., Colby, C.L., and Goldberg, M.R. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. Science *255*, 90.
17. Rolfs, M., Jonikaitis, D., Deubel, H., and Cavanagh, P. (2010). Predictive remapping of attention across eye movements. Nat. Neurosci. *14*, 252–256.
18. Hunt, A.R., and Cavanagh, P. (2011). Remapped visual masking. J. Vis. *11*, 13.
19. Kusunoki, M., and Goldberg, M.E. (2003). The time course of perisaccadic receptive field shifts in the lateral intraparietal area of the monkey. J. Neurophysiol. *89*, 1519–1527.
20. Hunt, A.R., and Cavanagh, P. (2009). Looking ahead: The perceived direction of gaze shifts before the eyes move. J. Vis. *9*, 1, 1–7.

[1]Psychology Department and Center for Neural Science, New York University, New York, New York, USA. [2]Laboratoire Psychologie de la Perception, Université Paris Descartes, CNRS UMR 8158, Paris, France.
E-mail: denis.pelli@nyu.edu

http://dx.doi.org/10.1016/j.cub.2013.04.022

Plant Cytoskeleton: DELLA Connects Gibberellins to Microtubules

A new study reveals that DELLA proteins directly interact with the prefoldin complex, thus regulating tubulin subunit availability in a gibberellin-dependent manner. This finding provides a mechanistic link between the growth-promoting plant hormone gibberellin and cortical microtubule organization.

Ram Dixit

Plant cells are surrounded by a rigid wall that precludes their movement. Therefore, plant growth and development relies in large part on regulation of the extent and direction of cell expansion. The cortical microtubule cytoskeleton is a key part of the cellular machinery that defines