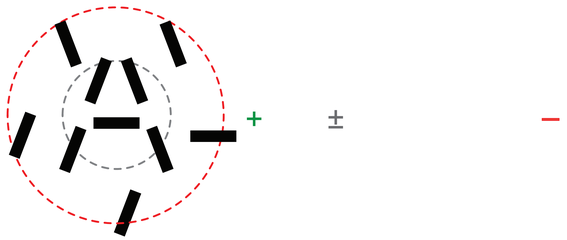
Crowding shows that faces have parts and bodies don't

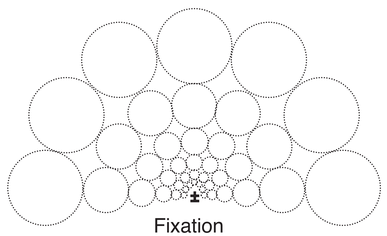
Kat Tillman, Momo Araki, & Denis Pelli
Psychology & Neural Science, NYU

Crowding



To recognize an object, the visual system must first detect and then combine its features. If you fixate on the red minus above, you fail to recognize the A. The features (black bars) are being combined over too large an area (red dashed circle) to isolate the features of the A from the nearby junk. This is crowding. The critical spacing to avoid crowding is the diameter of the integration area over which features are combined.

Critical spacing depends only on where the object is (Pelli & Tillman 2008). Critical spacing is proportional to the distance from fixation (Bouma 1970).



Parts & wholes

Some objects, like letters, are recognized through a single integration of features over the whole object. These objects are recognized as ***wholes***, like the letter "n" in the demo to the right.

+ Holistic

Other objects, like words, require separate integration of features over distinct regions of the object. These regions are ***parts***. In the demo, the parts are sufficiently spaced for separate integration, so you do recognize the word "one."

+

Parts within an object crowd one another unless they are more than critically spaced (Martelli et al. 2005). In the demo, the letters are not sufficiently spaced, so you cannot recognize them, and thus cannot recognize the word.

–

Using crowding to count parts

Thanks to crowding, we can count the number of parts in an object simply by measuring the object's threshold size in the periphery. If the object has only one part, its threshold size is tiny, limited by acuity, not crowding. If the object has multiple parts, its threshold size is determined by crowding. At threshold size the parts are critically spaced. Assuming the parts are evenly spaced, there is one part per squared critical spacing. Thus the number of parts can be estimated by dividing the area of the object by the area per part:

$$\text{Complexity} = \frac{\text{Area of object}}{\text{Area per part}}$$

$$= \frac{\text{Area of object}}{(\text{Critical spacing})^2}$$

Complexity ≤ 1
Holistic

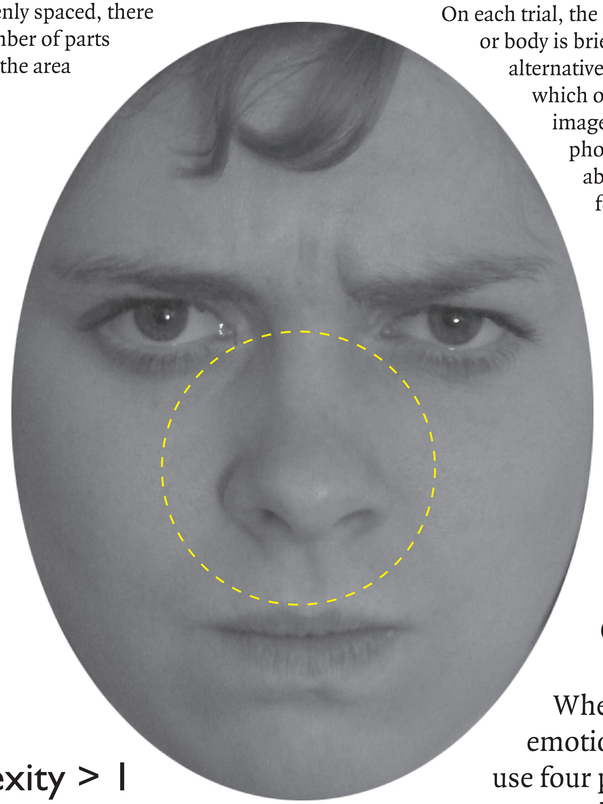
■

Complexity > 1
By-parts

If the object can be recognized even when the whole thing falls inside one critical spacing, then it is recognized as a whole. If the object must be larger than the critical spacing to be recognized, then it contains parts (which crowd one another unless integrated separately).

Method

We measure the observer's threshold size for recognition of the emotion portrayed by a face photograph or body silhouette in peripheral vision and then calculate complexity.



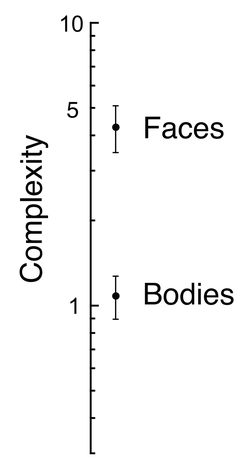
On each trial, the observer fixates a small square while an emotional face or body is briefly flashed at 8 deg in the left visual field. The 5-alternative forced choice task requires the observer to decide which of five basic emotions the image portrays. The set of images includes 56 faces and 59 bodies. All are based on photographs of the same actress. Sample stimuli appear above. A control experiment (not shown) confirms that, for emotion recognition, the body silhouettes have similar size thresholds as body photographs. The QUEST staircase procedure adjusts the size of the image on the next trial to home in on the threshold size for 82% correct recognition of emotion. Faces and bodies were run in separate blocks. 11 observers participated.

Results & conclusion

When recognizing the emotion in a face, observers use four parts (4.3 ± 0.8, *mean* ± *se* across 11 observers).

When recognizing the emotion in a body, observers use one part (1.1 ± 0.2).

Thus, in this task, faces are recognized by parts, and bodies are recognized as wholes.



Tillman K A, Araki M, Pelli D G, 2008, "Crowding shows that faces have parts and bodies do not" *Perception* 37 ECVP Abstract Supplement, page 33