Supposing that crowding is compulsory grouping suggests a remarkably simple model for object recognition

— Denis Pelli & Sarah Rosen — Psychology & Neural Science, NYU

Optional abstract

Grouping and crowding have each received much study, but, because they can be avoided, they have seemed marginal to the mystery of how people recognize objects. We begin by showing that the same image parameters that promote grouping also promote crowding. Joining grouping and crowding theories suggests a unified account in which grouping and crowding are two aspects of the classification underlying all object recognition. In this unified account, recognition of a simple object, like a letter, is mediated by a single combining field driving a simple classifier that can only recognize one thing at a time. It follows that the human observer has a vast array of combining fields, one for each size and retinal location, and is usually free to use whichever best matches the target. For simplicity, we suppose they are all similar, differing only in position and size. As eccentricity increases, the smallest available combining field size increases, presumably because fewer neurons are available. Crowding arises when the smallest available combining field is too large to isolate the target object from adjacent objects, as a simple classifier cannot recognize two things at once. To test the notion that the same classifiers underlie both crowding and grouping, we show that the efficiency of identification by a single isolated combining field in the periphery is practically the same as the efficiency of unrestricted central vision of the same target. We conclude that the same unit for object recognition underlies both grouping and crowding, optional and compulsory. Critical spacing marks the boundary between optional and compulsory.

Optional table

The same factors promote grouping and crowding

GROUPING LAW	GROUPING Task	Grouped	CROWDING Task	Crowded	Fixation	Uncrowded	CROWDING PAPERS
Proximity Close elements tend to group.	*Do you see the dots grouped into columns or rows?*	● ● / ● ●	*While fixating the plus, is the middle letter an M or a W?*	XWX	+	X W X	Bouma, 1970 Townsend et al., 1971 Andriessen & Bouma, 1976 Bouma, 1978, 1980, 1991 Banks et al., 1979 Whitaker et al., 1989 Strasburger et al., 1991 Toet & Levi, 1992
Similarity Similar elements tend to group.	*Columns or rows?*		*Which way is the middle tilted?*		+		Andriessen & Bouma, 1976 Westheimer et al., 1976 Kooi et al., 1994, 1997 Levi et al., 2002 Hariharan et al., 2005
	Columns or rows?		*Which way is the middle tilted?*		+		Pizlo & Salach-Golyska, 1993, 1994 Kooi et al., 1994 Hess et al., 2000 Chung et al., 2001 Levi et al., 2002
	Columns or rows?	● ● / ● ●	*Is the middle letter M or W?*	XWX	+	XWX	Estes, 1982 Pizlo & Salach-Golyska, 1993, 1994 Kooi et al., 1994 Hess et al., 2000 Chung et al., 2001
	Columns or rows?	● ● / ● ●	*Is the middle letter M or W?*	XWX	+	XWX	Bouma, 1969 Kooi et al., 1994 Tripathy & Cavanagh, 2002 Poder, 2006, 2007 van den Berg et al., 2007
	Columns or rows?	■ ■ / ■ ■	*Is the middle letter M or W?*	XWX	+	OWO	Estes, 1982 Nazir, 1992 Kooi et al., 1994 Leat et al., 1999 Poder, 2006
	Columns or rows?	● ● / ● ●	*Is the middle letter M or W?*	XWX	+	XWX	Buffer & Westheimer, 1978 Levi et al., 1985 Felisberti & Solomon, 2005
	Columns or rows?	M O / B C	*Can you describe the middle?*	天大天	+	體大體	Leat et al., 1999 Felisberti et al., 2007 Zhang et al., 2009 Bernard & Chung, 2011
Continuity Continuous elements tend to group.	*Two lines or four?*		*Which way is the middle tilted?*	— — —	+	| — |	Field et al., 1993 Kapadia et al., 1995, 1999 Livne & Sagi, 2007, 2010 Levi & Carney, 2009 Chakravarthi & Pelli, 2011
Regularity Regular elements tend to group.	*Four dots or one dotted line?*		*Which way is the middle tilted?*	| | / | |	+	|| / ||	Malania et al., 2007 Sakaida et al., 2009, 2010 Yeotikar et al., 2011 Manassi et al., 2012 Rosen & Pelli, submitted
Common fate Elements that move together tend to group.	*Columns or rows?* * *Red arrows indicate grating motion*	← → / ← →	*Which way is the middle tilted?**		+	← → ←	Bex et al., 2003 Rajimehr et al., 2004 Bex & Dakin, 2005 Dakin et al., 2010 Maus et al., 2011
Connectedness Connected elements tend to group.	*Four lines or a rectangle?*		*Which way is the middle tilted?*	| | / | |	+	□ / □	Sayim et al., 2010 Manassi et a., 2012 Rosen et al., (submitted)
	One object or two?		*Which way is the middle tilted?*		+	□ / □	Rosen et al., (submitted)
Common region Elements in the same region tend to group.	*Four shapes or one face?*	☺	*Which way is the middle tilted?*	| | / | |	+	| | / | |	Presented here for the first time.
Closure Broken objects tend to be perceived as whole.	*Four lines or a rectangle?*		*Which way is the middle tilted?*	| | / | |	+	⊏⊐ / ⊏⊐	Livne & Sagi, 2007, 2010
	Five lines or one object?		*Which way is the middle tilted?*	⊏⊐⊏⊐⊏⊐	+	⊏⊐ / ⊏⊐	Presented here for the first time.
Amodal completion Incomplete objects tend to be perceived as complete.	*Two objects or three?*		*Can you describe the middle shape?*		+	◗ ⧗ ◖	Presented here for the first time.
Figure/ ground One object is seen as the figure while the other is seen as the background. It's bistable.	*Two faces or one vase?*		*A single vase is describable, but a pair of faces is a crowded jumble.*		+		Presented here for the first time.

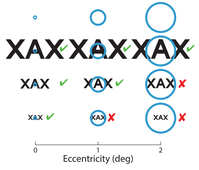
Facts

About crowding

Crowding is a bottleneck.

○ indicates the combining field, i.e. within the critical spacing of crowding.
✓ indicates recognition of the target **A**.
✗ indicates failure to recognize.

The extent of crowding is proportional to eccentricity and independent of target size.

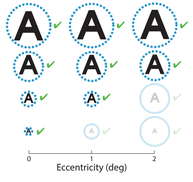


Facts

About grouping

Grouping allows us to recognize multiple elements as a simple object. It is promoted by several factors — Proximity, Similarity, etc. — which are known as the *Laws of Grouping*.

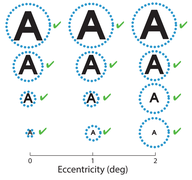
The extent of grouping is proportional to object size. But crowding sets a minimum (shown faintly).



Conjecture

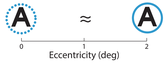
Universal recognition unit

It seems that the underlying operation of crowding and grouping may be the same: We call it "grouping" when recognition succeeds, and "crowding" when it fails. Suppose that recognition of any simple object, like a letter, is always mediated by a universal recognition unit. They're all the same, differing only in size and position. At each eccentricity, crowding is the smallest grouping.



Deductions

Every grouping is like a more peripheral crowding.



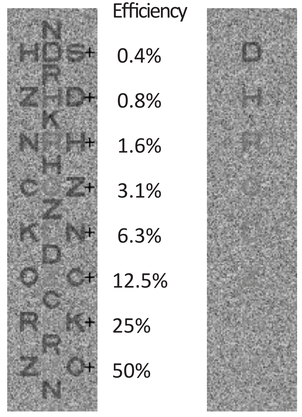
The bottleneck is ubiquitous.

Prediction

Unrestricted vs. isolated

Efficiency of letter identification in noise measures how well recognition works. Since a letter is always recognized by a universal recognition unit, the efficiency of an isolated unit should be practically the same as for unrestricted vision.

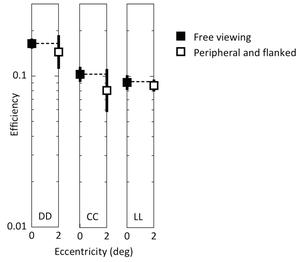
Test yourself



Isolating a recognition unit. Your efficiency of letter recognition is 6%, whether you are constrained by crowding to use a single peripheral recognition unit (left — fixate the plus and identify the middle letter) or free to use central vision (right). This is evidence that visual letter recognition is always mediated by recognition units like the one isolated here.

Results

Identify a letter in noise.



Conclusion

Parsimony and the experimental result support the idea that object recognition is meditated by a vast array of similar recognition units, of various sizes and locations, each capable of recognizing things no more complex than a single Roman letter.

Citation
Pelli, D. G., & Rosen, S. (2105) Supposing that crowding is compulsory grouping suggests a remarkably simple model for object recognition. Vision Sciences Society Annual Meeting, St. Pete Beach, Florida, May 15-20, 2015.
http://www.visionsciences.org/abstract_detail.php?id=572
http://psych.nyu.edu/pelli/posters.html

Rosen, S., and Pelli, D. G. (favorably reviewed, under revision) A review of crowding and grouping suggests a unit for object recognition. Submitted to *Journal of Vision* special issue on "Crowding".