Vision Research 153 (2018) 13–23

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

# Critical resolution: A superior measure of crowding

Leili Soo*, Ramakrishna Chakravarthi, Søren K. Andersen

*School of Psychology, University of Aberdeen, United Kingdom*

ARTICLE INFO

*Keywords:*
Object recognition
Visual crowding
Psychophysics
Critical spacing
Critical resolution
Visual perception
Flanker interference

ABSTRACT

Visual object recognition is essential for adaptive interactions with the environment. It is fundamentally limited by crowding, a breakdown of object recognition in clutter. The spatial extent over which crowding occurs is proportional to the eccentricity of the target object, but nevertheless varies substantially depending on various stimulus factors (e.g. viewing time, contrast). However, a lack of studies jointly manipulating such factors precludes predictions of crowding in more heterogeneous scenes, such as the majority of real life situations.

To establish how such co-occurring variations affect crowding, we manipulated combinations of 1) flanker contrast and backward masking, 2) flanker contrast and presentation duration, and 3) flanker preview and pop-out while measuring participants' ability to correctly report the orientation of a target stimulus. In all three experiments, combining two manipulations consistently modulated the spatial extent of crowding in a way that could not be predicted from an additive combination. However, a simple transformation of the measurement scale completely abolished these interactions and all effects became additive. Precise quantitative predictions of the magnitude of crowding when combining multiple manipulations are thus possible when it is expressed in terms of what we label the 'critical resolution'. Critical resolution is proportional to the inverse of the smallest flanker free area surrounding the target object necessary for its unimpaired identification. It offers a more parsimonious description of crowding than the traditionally used critical spacing and may thus constitute a measure of fundamental importance for understanding object recognition.

## 1. Introduction

Object recognition is essential for visually guided adaptive behaviour. For example, while driving on a rainy evening, timely recognition of a pedestrian about to cross the street may be essential to avoiding an accident. Our ability to recognise an object in the periphery as a pedestrian would be impaired if she were standing next to an object of similar size and shape, such as for example, a road sign. This reduction in the ability to identify objects in clutter is called visual crowding (Bouma, 1970; Levi, 2008; Pelli & Tillman, 2008; Pelli, Palomares, & Majaj, 2004; Whitney & Levi, 2011). Crowding fundamentally limits our ability to process visual scenes as diverse as driving, reading or searching for a particular object. In most situations crowding, rather than visual acuity, is the limiting factor on visual perception. In recent years, substantial efforts have been undertaken to uncover the limits of object recognition, using crowding as a tool (Chung, Levi, & Legge, 2001; Harrison & Bex, 2015; He, Cavanagh, & Intriligator, 1996; Herzog & Manassi, 2015; Herzog, Sayim, Manassi, & Chicherov, 2016; Pelli et al., 2004).

The Bouma Law (coined by Pelli & Tillman, 2008) describes one of the most fundamental properties of crowding. It states that the distance between a target and its flankers below which the flankers start to interfere with the identification of the target is proportional to the target's eccentricity, i.e. its distance from fixation (Bouma, 1970). This distance between target and flankers is known as the 'critical spacing' and is considered to be the measure that best characterises the interference between nearby objects. It was initially reported to be approximately half the target's eccentricity (Bouma, 1970). There is evidence that the Bouma Law holds true for a large variety of objects and features, such as orientation, hue, lightness, size (van den Berg, Roerdink, & Cornelissen, 2007), spatial frequency (Chung et al., 2001), letters (Bouma, 1970; Kooi, Toet, Tripathy, & Levi, 1994; Pelli et al., 2004; Wolford & Chambers, 1984), faces (Farzin, Rivera, & Whitney, 2009), real-world objects (Wallace & Tjan, 2011) and natural scenes (Wallis & Bex, 2012). This consistency has led some researchers to propose the Bouma Law as a general principle of object recognition (Pelli & Tillman, 2008) that has implications for the neural mechanisms of feature integration. According to this idea, neurons (in say V1) responding to object features will pool their responses if they are within a certain distance (6 mm in the radial direction) of each other in the cortex (Pelli, 2008), leading to crowding.

However, this notion seems inconsistent with studies that have

* Corresponding author at: School of Psychology, William Guild Building, Room T32, Kings College, Old Aberdeen, AB24 3FX, United Kingdom.
*E-mail address:* leilisoo@abdn.ac.uk (L. Soo).

https://doi.org/10.1016/j.visres.2018.08.005
Received 24 April 2018; Received in revised form 30 July 2018; Accepted 31 August 2018

revealed large variations in the proportionality constant that links critical spacing and eccentricity. For example, critical spacing is reduced (less target-flanker interference) if target and flankers differ in some property such as colour (Andriessen & Bouma, 1976; Chung et al., 2001; Kennedy & Whitaker, 2010; Kooi et al., 1994; Nazir, 1992; Põder, 2007; Scolari, Kohnen, Barton, & Awh, 2007) or if the flankers are previewed (Scolari et al., 2007; Watson & Humphreys, 1997). On the other hand, critical spacing is increased, and indeed can be much larger than half the eccentricity, if the flankers' luminance contrast is higher than that of the target (Rashal & Yeshurun, 2014), if the target is mildly masked, (Vickery, Shim, Chakravarthi, Jiang, & Luedeman, 2009), or if display duration is reduced (Kooi et al., 1994; Tripathy, Cavanagh, & Bedell, 2014), whereas masking the flankers reduces critical spacing (Chakravarthi & Cavanagh, 2009; Wallis & Bex, 2011).

These findings suggest substantial variability in the distance over which features are integrated, depending on stimulus properties. Thus, the amount of crowding may differ vastly between dissimilar scenes or even objects within the same scene. To understand how crowding limits visual perception, it is, therefore, necessary to know how various stimulus manipulations affect crowding and what the combined effect of such manipulations is. The latter is especially important for two reasons. First, real-world scenes combine multiple object properties in a variety of ways. For example, a flanker might differ from the target in contrast, spatial frequency, and orientation, simultaneously. In addition, effective viewing durations might vary a lot due to movements of eyes, observers, or objects. Masking can occur when an object or its flankers are occluded by other (perhaps moving) objects. In order to move towards an understanding of the limitations of object recognition in the real world, it is therefore necessary to understand the effects of combinations of stimulus properties. Second, the magnitude of the effects of different stimulus properties on crowding can only be compared across studies if they are either independent of each other or if the way in which these effects are combined is exactly understood. For example, doubling the contrast of flankers (while keeping target contrast constant) approximately doubled the critical spacing in a previous study (Rashal & Yeshurun, 2014). Would such a surprisingly large effect also have been observed if stimuli had not been presented very briefly and with a backward mask? It could even be the case that the effect of one manipulation is contingent upon a certain combination of other factors. If this were the case, manipulating flanker contrast might only have a (detectable) effect when measured under these specific conditions. Perhaps surprisingly, previous studies have typically tested the effects of manipulating stimulus properties on crowding in isolation (e.g., Kooi et al., 1994; Rashal & Yeshurun, 2014; Scolari et al., 2007). It is therefore unknown what the combined effect of such manipulations is and whether it follows a regular pattern across different manipulations.

The present study examined how the effects of stimulus properties that affect object recognition in a cluttered scene are combined. We manipulated flanker contrast together with backward masking (Experiment 1) and display duration (Experiment 2). Additionally, we manipulated flanker preview and target-flanker similarity in a third experiment (Experiment 3). We employed full-factorial designs in order to assess both main effects and interactions of these manipulations on critical spacing. This allows us to determine whether the effects of combining two properties can be predicted from the extent of visual crowding observed when manipulating these properties separately.

## 2. Methods

### 2.1. Participants

All participants were students at the University of Aberdeen. Experiment 1 had fifteen participants (11 female; 13 right-handed; mean age = 22.2 years; age range: 18–25 years), Experiment 2 had ten participants (6 female; all right-handed; mean age = 22.6 years; age range: 19–27 years) and Experiment 3 had twelve participants (8 female; 11 right-handed; mean age = 24.1 years; age range: 20–26). In all experiments, participants had normal or corrected to normal visual acuity. Participants gave written informed consent prior to participation. They received either £5 or course credits as compensation for their participation. All experiments were approved by the University of Aberdeen Psychology Ethics Committee (Project number: PEC/3146/2014/10) and the work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

### 2.2. Experiment 1

#### 2.2.1. Materials

Stimuli were generated and presented using Matlab (The MathWorks, Natick, MA) with the Cogent Graphics toolbox (developed by John Romaya, Laboratory of Neurobiology, Wellcome Department of Imaging Neuroscience) on a 19 in. CRT monitor set to a resolution of 1024 × 768 pixels and a refresh rate of 60 Hz, viewed from a distance of 60 cm. The target was the letter 'T' (1.3° of visual angle) and was presented 9° from fixation in either the left or right visual field along the horizontal meridian in one of four cardinal orientations: upright, inverted, rotated 90° right or 90° left. The target letter appeared either in isolation or was surrounded by three flanking stimuli (above, below and on the outer side of the target). No flanker was presented on the inner side of the target as such a flanker would have approached or intersected fixation at large target-flanker distances. Flankers were letter 'H's (same size as the target stimulus), presented either upright or rotated 90°. Flankers, when present, could be at one of seven possible distances from the target measured centre to centre: 1.5°, 2°, 2.5°, 3°, 4°, 5° and 7° of visual angle. The experiment manipulated the presence of backward masking and flanker contrast. The backward mask was a rectangle of size 8.2° × 26.7°, created by tiling patches of size 0.2° × 0.2°. Each individual patch of the mask had a random grey scale luminance value sampled from a uniform distribution between 0.02 and 57.44 cd/m$^2$.

The Weber Contrast of stimuli was calculated as follows:

$$contrast = \frac{I - I_b}{I_b} \tag{1}$$

where $I$ is the luminance of the stimulus and $I_b$ is the luminance of the background. Targets had a luminance of 19.6 cd/m$^2$ corresponding to a contrast of 0.25 against the grey background (15.7 cd/m$^2$). The flankers either had the same contrast as the target or had a luminance of 39.5 cd/m$^2$ corresponding to a contrast of 1.5 from the background.

#### 2.2.2. Procedure

The sequence of events during Experiment 1 are depicted in Fig. 1A. Each trial started with a fixation cross presented for 1000 ms. Then, the target and three flankers were presented for 100 ms. In half the trials, a noise mask was presented for 300 ms after the offset of the target display (target-mask SOA of 100 ms). Flanker contrast was the same as the target's in half the trials and higher in the other half. Target and flanker orientations were randomly chosen for each trial. Participants were instructed to report the target orientation by pressing the corresponding arrow key (left, right, up or down) on a keyboard. Auditory feedback was provided on each trial; percentage correct averaged over all the trials within a block was displayed at the end of that block.

Participants underwent training for 1–3 blocks of 32 trials each, at the beginning of the experiment. The main experiment consisted of a total of 1024 trials. There were 256 different types of trials: 2 sides (L/R) × 2 flanker contrasts (equal/higher) × 2 masking conditions (yes/no) × 8 flanker distances (1.5°, 2°, 2.5°, 3°, 4°, 5°, 7° and no flankers) × 4 target orientations. Each type of trial was repeated 4 times and all trials were presented in random order. After every block of 128 trials, participants were given a self-paced break during which they received written feedback on their average accuracy in the preceding block. For purposes of analysis, data was averaged over sides and target