Macaque monkeys experience visual crowding

Erin A. Crowder
Department of Neuroscience and Center for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, PA, USA

Carl R. Olson
Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA, USA

In peripheral vision, objects that are easily discriminated on their own become less discriminable in the presence of surrounding clutter. This phenomenon is known as crowding. The neural mechanisms underlying crowding are not well understood. Better insight might come from single-neuron recording in nonhuman primates, provided they exhibit crowding; however, previous demonstrations of crowding have been confined to humans. In the present study, we set out to determine whether crowding occurs in rhesus macaque monkeys. We found that animals trained to identify a target letter among flankers displayed three hallmarks of crowding as established in humans. First, at a given eccentricity, increasing the spacing between the target and the flankers improved recognition accuracy. Second, the critical spacing, defined as the minimal spacing at which target discrimination was reliable, was proportional to eccentricity. Third, the critical spacing was largely unaffected by object size. We conclude that monkeys, like humans, experience crowding. These findings open the door to studies of crowding at the neuronal level in the monkey visual system.

Introduction

In comparison to foveal vision, our view of the periphery is impoverished. This is due in part to the fact that there are fewer cones and fewer retinal ganglion cells dedicated to peripheral than to foveal locations (Wässle, Grünert, Röhenbeck, & Boycott, 1989). Foveal overrepresentation persists as visual information flows through the thalamus to primary visual cortex, where the amount of tissue devoted to a given eccentricity is directly proportional to acuity. Thus acuity is believed to decrease in the periphery as a direct result of coarser sampling by cones (Cowey & Rolls, 1974).

Peripheral vision suffers not only from reduced acuity but also from information loss due to crowding. The essence of crowding is that a peripheral item recognizable on its own becomes illegible when surrounded by other nearby items. Crowding is usually quantified in terms of critical spacing, the maximum distance at which surrounding clutter interferes with object recognition. Critical spacing, like acuity, scales with eccentricity (Bouma, 1970). However, unlike acuity, critical spacing possesses no well-established neuronal explanation. Mechanisms that appear to have been ruled out include surround suppression (Petrov, Popple, & McKee, 2007) and impaired feature detection (Levi, Harihara, & Klein, 2002a; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli, Palomares, & Majaj, 2004). Pooling of feature information within neuronal receptive fields, however, remains a plausible explanation (Flom, Weymouth, & Kahneman, 1963b). The rate at which critical spacing scales with eccentricity in humans is explicable by pooling within windows roughly the size of neuronal receptive fields in area V2 of the monkey (Freeman & Simoncelli, 2011). Yet pooling within V2 receptive fields cannot be the full explanation, because these receptive fields lack the anisotropic structure required to account for radial-tangential differences in critical spacing (Toet & Levi, 1992). Elliptical zones of integration might conceivably arise from a top-down selection process (He, Cavanagh, & Intriligator, 1996) tied to the saccadic system, in which an anisotropy for precision mirrors the anisotropy for critical spacing (Harrison, Mattingley, & Remington, 2013; Nandy & Tjan, 2012).

To draw firm conclusions concerning the neuronal processes that underlie crowding will require studying the phenomenon by means of invasive techniques such as are typically employed in nonhuman primates. However, nonhuman primates will be appropriate for study only if they exhibit crowding. The aim of the
present study was to determine whether they do. The universal hallmark of crowding is Bouma’s law, which in its most general form states that the critical spacing at which an object becomes unidentifiable among similar flankers depends solely on eccentricity, regardless of the nature of the object (Pelli & Tillman, 2008). It follows that critical spacing is independent of object size (Levi, 2008; Pelli et al., 2004). To determine whether monkeys exhibit crowding, we trained two macaques to perform a visual discrimination task in which we could vary the spacing, eccentricity, and size of the target and the flanking distractors. We found that psychometric functions relating accuracy to target–flanker distance resembled those of humans, that critical spacing was proportional to eccentricity, and that critical spacing was largely unaffected by object size. We conclude that monkeys experience crowding. This observation paves the way for investigations into the neural mechanisms underlying crowding in awake, behaving monkeys.

Methods

Animals and equipment

Two adult male rhesus macaque monkeys (Macaca mulatta) were used in these experiments (Monkey 1 and Monkey 2). Experimental procedures were approved by the Carnegie Mellon University Institutional Animal Care and Use Committee and were in compliance with the United States Public Health Service Guide for the Care and Use of Laboratory Animals.

For behavioral testing, each monkey was seated in a primate chair with the head stabilized by a surgically implanted post. Events during each trial were controlled by Cortex software (National Institute of Mental Health, Bethesda, MD). Visual stimuli were presented on a 17-in. LCD screen with a resolution of 1024 × 768 pixels positioned 18 in. from the animal’s eyes. Eye position was tracked by an infrared system (ISCAN, Woburn, MA). The system was calibrated by requiring the monkey, at the beginning of each block of trials, to fixate a small target presented successively at four locations corresponding to the corners of a 14° × 14° square centered on the screen. Off-line, the readings on each trial were converted to degrees of visual angle by performing a linear transformation based on the stored calibration voltages.

Task design

On each trial, the monkey responded to presentation of a target in the right visual field by making a saccade directly above or below fixation (Figure 1A). The targets were Sloan letters A, F, H, U, and Z (courtesy of Denis Pelli) and counterparts obtained by rotating them 90°. Each letter had an aspect ratio of 1. A letter and its rotated counterpart were associated with saccades in opposite directions according to rules counterbalanced across animals. The target–saccade mapping shown in Figure 1B was used for Monkey 1, as indicated here, was reversed for Monkey 2. Flankers, when present, consisted of Sloan letters K, P, T, and Y. Their arrangement varied from trial to trial (Figure 1C). They were always the same size as the target. We chose letters as stimuli because their use is common in human studies of crowding. The monkeys’ prior experience with letters was fundamentally different from the experience of human subjects. To allay concern that this might affect crowding, we collected data from two human subjects using identical displays, as described in a later section.

Within each block of trials, the size and eccentricity of the target were fixed. Size and eccentricity were manipulated across three experiments: Experiment 1 (size 1° at eccentricity 6°), Experiment 2 (size 0.5° at eccentricity 3°), and Experiment 3 (size 0.5° at eccentricity 6°). Characterizing the impact of size and eccentricity on choice accuracy required cross-block comparison. To minimize the influence of random fluctuations from block to block, each monkey completed 12 blocks of trials for each experiment. Each block required performing a discrimination under 192 different conditions, as described next.

Within each block, the variable of interest was the center-to-center spacing between the target and the
flankers. On a given trial, this could assume any of six values with equal likelihood. In a block involving the presentation of targets at an eccentricity of 6°, the possible spacings were 1.1°, 1.45°, 1.8°, 2.15°, 2.5°, and infinity (target alone). In a block involving the presentation of targets at an eccentricity of 3°, the possible spacings were 0.6°, 0.8°, 1.0°, 1.2°, 1.4°, and infinity (target alone). Other incidental factors varying within a block were fully counterbalanced against spacing. These factors included target identity, saccade direction, placement of the target in the upper or lower visual field, and flanker configuration.

In each block, we employed as targets two letters and their rotated counterparts. The four targets appeared with equal frequency. Saccades in upward and downward directions were demanded with equal frequency because targets associated with the two directions were equally common. In each block, the target appeared equally often above and below the 0° horizontal meridian. In blocks involving the presentation of targets at 3° and 6° eccentricity, the vertical displacement from the horizontal meridian was 0.5° and 1°, respectively. The flankers could appear in any of four configurations (Figure 1C).

Full counterbalancing required assessing behavior under 192 conditions corresponding to all possible combinations of six spacings, four targets, two vertical locations, and four flanker configurations. The conditions were imposed in random order with the sole exception that each combination of target, spacing, and flanker occurred once in the first half of the block (when the display was centered at one vertical location) and again in the second half of the block (when the display was centered at the other vertical location). The sequence of vertical locations was upper-then-lower for half of the target–spacing–flanker combinations and lower-then-upper for the other half.

During a single block, the monkey had to complete a trial successfully under each of the 192 conditions. A trial was considered successful if the monkey made a saccade in the correct direction. This culminated in a juice reward followed by an immediate advance to the next trial. A trial was aborted if the monkey’s gaze deviated by more than 2° horizontally or 3° vertically from the central fixation point. In practice, gaze rarely deviated more than 1° horizontally or 2° vertically (Supplementary Figure S1). Breaking fixation or making an erroneous response resulted in withholding of the reward and a time-out of several seconds. The condition was returned to the pool from which future trials would be drawn. We based behavioral analysis exclusively on those 192 trials in which the monkey made the first saccadic decision under a given condition, without regard to whether the decision was correct or incorrect.

**Schedule of training and testing**

After training on basic skills such as maintaining gaze on a central fixation point and making a saccade to a suddenly appearing peripheral circle, the monkeys were introduced to a visual discrimination task in which a 1° target appearing at fixation instructed an upward or downward saccade. This phase took 1 month in Monkey 1 and 3 months in Monkey 2. Next, the monkeys were eased into performing the same discrimination on 1° targets presented at an eccentricity of 6°. This phase took 1 month in Monkey 1 and 3 months in Monkey 2. Next, they were habituated to performing in the presence of flanking distractors at various spacings through presentation of the distractors at very low contrast initially with a gradual contrast increase. This phase took 1 month in Monkey 1 and 4 months in Monkey 2. We continued to train the monkeys with flankers fully visible until their performance stabilized. This took 4 months in Monkey 1 and 2 weeks in Monkey 2. We then introduced them to task variants with 0.5° targets centered at an eccentricity of 3° or 6°. To achieve stable behavior under multiple interleaved conditions took 2 weeks in Monkey 1 and 6 weeks in Monkey 2. Finally, we collected behavioral data over the course of 1 month in each animal, interleaving blocks of trials with 0.5° objects at an eccentricity of 3°, 0.5° objects at an eccentricity of 6°, and 1° objects at an eccentricity of 6°.

**Data analysis**

The universal measure of crowding is critical spacing, the maximal spacing at which flankers seriously interfere with target discrimination (Bouma, 1970; Pelli & Tillman, 2008). Human studies typically adopt a definition of critical spacing based on a fixed threshold halfway between chance and perfect accuracy, taking the critical spacing to be that spacing at which a psychometric function fitted to the data intersects the threshold (Chung, 2007; Kooi, Toet, Tripathy, & Levi, 1994; Toet & Levi, 1992). The use of a predefined threshold would be problematic in monkeys, because their performance is more erratic than the performance of humans. Even under undemanding conditions, overall accuracy rarely approaches 100%. Furthermore, overall accuracy can be affected by minor changes in a task, including, in the present instance, alterations of target size and eccentricity and the addition of flankers. It is impossible, in such cases, to distinguish between a bottom-up cause (such as poor acuity) and a top-down cause (such as poor motivation or confusion in the face of difficulty). To circumvent this difficulty, we adopted the following approach.
We defined threshold as the inflection point of a sigmoidal function fitted to points representing accuracy as a function of flanker spacing:

\[
P(s) = \beta_1 + \frac{\beta_0 - \beta_1}{1 + \left( \frac{s}{s_C} \right)^{b_2}}
\]

where \(P(s)\) is the probability of a correct response at a given spacing \(s\), \(\beta_0\) represents the lower asymptote, \(\beta_1\) is the upper asymptote, \(b_2\) determines the slope at the inflection point, and \(s_C\) is the inflection point. Model parameters were fitted using nonlinear least squares (provided in the MATLAB Curve-Fitting Toolbox). We operationally defined critical spacing as the model coefficient \(s_C\). This approach conforms in spirit to the practice in human studies of selecting a threshold midway between chance and perfect performance.

Human studies often include data from trials in which flankers were absent in the set of data to which the psychometric curve is fitted (Levi et al., 2002a; Pelli et al., 2004). The inclusion of singleton data would be problematic in monkeys due to reasons noted. Accordingly, we based our estimate of critical spacing exclusively on trials in which flankers were present.

To be sure that the results obtained from monkeys were not an artifact of these choices with regard to how critical spacing was measured, we repeated all analyses using two alternative models: a model in which \(\beta_1\) was fixed at the performance level when no flankers were present and a model in which \(b_2\) was fixed at the average slope across experiments. The essential findings were the same (Supplementary Figure S2). We also applied to human data the measurement procedure customized for use in monkeys. The essential findings were the same (Supplementary Figure S3).

**Results**

Both animals were able to discriminate the target at a rate well above chance when the flankers were sufficiently far away. Both experienced a falloff in accuracy as the flankers moved closer to the target. To determine whether the pattern of falloff was consistent with expectations based on crowding, we assessed performance as a function of the eccentricity of the display and the size of the letters within it.

**Experiment 1**

We first assessed performance with displays consisting of 1° letters with the target at 6° eccentricity (Figure 2A). These parameters are within the range commonly used to demonstrate crowding in humans (Chung, 2007; Levi et al., 2002a; Pelli et al., 2004; Tripathy & Cavanagh, 2002). As in humans (Kooi et al., 1994; Tripathy & Cavanagh, 2002; Yeshurun & Rashal, 2010), accuracy increased as a function of target–flanker spacing, in a pattern well fitted by a sigmoidal function (Figure 2C, D; goodness of fit: \(R^2 = 1.0\) in Monkey 1 and 0.99 in Monkey 2). That the fit was good is not surprising, inasmuch as there were five data points and the model had four free parameters. The purpose of curve fitting was to allow us to establish the inflection point of the best-fit curve, which served as an operational measure of critical spacing (see Methods). This measure possesses the virtue of being insensitive to asymptotic accuracy, which typically varies from monkey to monkey. Monkey 1 (Figure 2C) was superior to Monkey 2 (Figure 2D) in asymptotic accuracy. Nevertheless, the measured critical spacing was virtually identical in the two animals: 1.45° in Monkey 1 and 1.47° in Monkey 2.

**Experiment 2**

If the critical spacing, as measured in Experiment 1, genuinely arose from crowding, then according to Bouma’s law (Bouma, 1970; Pelli & Tillman, 2008), it should decline with a reduction in eccentricity. To test
this prediction, we scaled the display down by a factor of 0.5, reducing target eccentricity to 3° and letter size to 0.5° and contracting the range of target–flanker spacings proportionately. As in the first experiment, the data were well fitted by a sigmoidal function (Figure 3C, D; goodness of fit: $R^2 = 0.99$ in Monkey 1 and 0.90 in Monkey 2). The measured critical spacing was 0.82° in Monkey 1 (diminished from Experiment 1 by a factor of 0.57) and 0.90° in Monkey 2 (diminished from Experiment 1 by a factor of 0.61). These values were close to the value of 0.5 predicted from Bouma’s law.

**Experiment 3**

The reduction in critical spacing from Experiment 1 to Experiment 2 might in principle have arisen either from scaling down the eccentricity of the display or from scaling down the size of the letters. However, classic accounts of crowding predict that critical spacing should be largely independent of letter size (Levi et al., 2002a; Pelli et al., 2004; Tripathy & Cavanagh, 2002), unlike lateral masking, which should scale with size (Levi, Klein, & Hariharan, 2002b; Pelli et al., 2004). To test this prediction, we presented the display at the original eccentricity of 6° while employing small letters (0.5°). The data were well fitted by a sigmoidal function (Figure 4C, D; goodness of fit: $R^2 = 0.93$ in Monkey 1 and 0.90 in Monkey 2). The critical spacing derived from the fitted curves was 1.42° in Monkey 1 and 1.43° in Monkey 2. These values were very close to those measured in Experiment 1 with letters twice as large. This outcome is compatible with observations in humans viewing crowded displays. It is incompatible with an explanation based solely on lateral masking.

The results of Experiments 1–3 are summarized in Figure 5. From this figure, it is clear that the critical spacing depended primarily on eccentricity (6° in Experiments 1 and 3 as compared to 3° in Experiment 2) and not on letter size (as indicated by the dashed white lines superimposed on the bars). As a basis for statistical comparison among the outcomes of the three experiments, we computed critical spacing for each of the 12 blocks of trials completed by each monkey in each experiment.

To determine whether eccentricity influenced the critical spacing with size held constant, we carried out an ANOVA with monkey (1 or 2) and eccentricity (3° in Experiment 2 or 6° in Experiment 3) as factors. In accordance with Bouma’s law, there was a highly significant main effect of eccentricity, $p = 3.4 \times 10^{-6}$. The interaction between monkey and eccentricity was not significant, $p = 0.12$. 

---

Figure 3. Experiment 2. (A) The target was placed at an eccentricity of 3° in the right visual field. Each target and flanker subtended 0.5°. (B) Flankers were spaced at five center-to-center distances from the target. In a sixth condition, flankers were absent. (C) Accuracy as a function of spacing in Monkey 1. Each data point reflects the mean over all blocks. Error bars indicate the standard error of the mean across blocks. The blue curve is fitted to five points representing performance when flankers were present. The dashed blue line indicates the critical spacing, defined as the inflection point of the fitted curve. The red curve is carried over from Experiment 1 for comparison. (D) Equivalent psychometric data for Monkey 2.

Figure 4. Experiment 3. (A) The target was placed at an eccentricity of 6° in the right visual field. Each target and flanker subtended 0.5°. (B) Flankers were spaced at five center-to-center distances from the target. In a sixth condition, flankers were absent. (C) Accuracy as a function of spacing in Monkey 1. Each data point reflects the mean over all blocks. Error bars indicate the standard error of the mean across blocks. The green curve is fitted to five points representing performance when flankers were present. The dashed green line indicates the critical spacing, defined as the inflection point of the fitted curve. The red and blue curves are carried over from Experiments 1 and 2 for comparison. (D) Equivalent psychometric data for Monkey 2.
To determine whether letter size influenced critical spacing with eccentricity held constant, we carried out an ANOVA with monkey (1 or 2) and size (1° in Experiment 1 or 0.5° in Experiment 3) as factors. In accordance with Bouma’s law, size had no significant main effect on critical spacing, \( p = 0.21 \). However, the interaction between monkey and size did approach significance, \( p = 0.065 \). Post hoc analysis revealed that this effect arose from a tendency in Monkey 1 for the critical spacing to increase in conjunction with letter size (two-tailed \( t \) test), \( p = 0.01 \). In Monkey 1, a 100% increase in letter size produced a 25% increase in critical spacing. In Monkey 2, it produced a 4% decrease. Even the effect observed in Monkey 1 was far too small to support an explanation based solely on lateral masking.

Each of the aforementioned ANOVAs revealed a marginally significant main effect of monkey (\( p = 0.03 \) when eccentricity was a factor and \( p = 0.07 \) when size was a factor). This arose from a tendency for the critical spacing to be smaller in Monkey 1 than in Monkey 2. It is not surprising that there should have been a difference between the monkeys. Humans also show interindividual differences in critical spacing (Toet & Levi, 1992).

In each block of trials, the monkey was required to discriminate between two pairs of targets out of the five that were available for testing (Figure 1B). To be sure that the results generalized across targets, we sorted the data from all of the blocks by target pair. We found that overall accuracy varied with target identity—for an ANOVA with target pair as a factor, \( p = 6.9 \times 10^{-7} \) in Monkey 1 and \( 4.2 \times 10^{-5} \) in Monkey 2—with the pattern of dependence differing between monkeys as if each had learned some target pairs better than others. To test whether the dependence of critical spacing on eccentricity in Experiments 2 and 3 was a function of target identity, we carried out an ANOVA on data from each monkey with target pair and eccentricity as factors. This revealed no significant main effect of target pair, \( p = 0.25 \) and 0.19 in Monkey 1 and Monkey 2, respectively; a significant main effect of eccentricity, \( p = 1.3 \times 10^{-7} \) and \( 1.3 \times 10^{-6} \) in Monkey 1 and Monkey 2; and, critically, no significant interaction, \( p = 0.35 \) and 0.94 in Monkey 1 and Monkey 2. To test whether the lack of dependence of critical spacing on size in Experiments 1 and 3 was a function of target identity, we carried out an ANOVA on data from each monkey with target pair and size as factors. This revealed no significant main effect of target pair, \( p = 0.27 \) and 0.61 in Monkey 1 and Monkey 2; no significant main effect of size, \( p = 0.08 \) and 0.81 in Monkey 1 and Monkey 2; and, critically, no significant interaction, \( p = 0.72 \) and 0.85 in Monkey 1 and Monkey 2. We conclude that the key results of the study did not depend on the identity of the targets.

Finally, we asked whether an identical approach would yield comparable results in humans. This step was motivated by two considerations. First, prior exposure to letters was different in the monkeys than it typically is in human subjects. The monkeys were intensively trained on arbitrary letter–saccade associations, whereas human subjects are typically literate. Second, the method by which we computed the critical distance in monkeys was adapted to their particular pattern of performance. In particular, we fitted a curve to data derived exclusively from flanker-present conditions. Using stimuli and methods of analysis identical to those employed in the monkey study, we assessed crowding in two human participants (see supplementary materials: Human experiments paralleling the monkey experiments). The results were closely similar to those obtained in monkeys (Supplementary Figure S3). An ANOVA on critical-spacing values with species (monkey or human) and experiment (1, 2, or 3) as factors revealed a marginally significant effect of species, \( p = 0.035 \), a highly significant effect of experiment, \( p = 3.5 \times 10^{-14} \), and no interaction effect, \( p = 0.21 \). The absence of a significant interaction effect indicates that the pattern of variation across experimental conditions (the signature of crowding) did not differ between species. In one additional human experiment, we showed that expanding the set of test conditions to include smaller target–flanker spacings exerted no systematic effect on measured critical spacing (see Supplementary Materials: Human experiment with narrow spacing and Supplementary Figure S4). We conclude that the observations obtained in monkeys are not an artifact either of their specialized experience with letters or of the methods of analysis necessary for characterizing their behavior.
Discussion

We carried out tests in macaque monkeys to determine whether they exhibit visual crowding. The key findings are the following. First, the ability of the monkeys to identify a peripheral target declined with decreasing distance between the target and the surrounding flanks. Second, the critical spacing scaled with eccentricity. Third, the critical spacing did not scale with object size. Together these results meet the standard diagnostic criteria for crowding (Levi, 2008; Pelli et al., 2004; Pelli & Tillman, 2008). This is the first demonstration that nonhuman primates exhibit crowding. The finding that crowding occurs in monkeys means that it will be possible in the future to investigate the neural underpinnings with invasive methods not generally applicable in human studies.

Comparison with human crowding literature

The average critical spacing for monkeys in this study was 0.26φ, where φ denotes target eccentricity. For humans tested under identical conditions, the average critical spacing was 0.25φ. These values are lower than the value of 0.5φ reported in the classic study by Bouma (1970). Many other reports on crowding also describe values less than 0.5φ. Pelli et al. note, with regard to results from a particular series of experiments, “Bouma was right to say ‘roughly’ 0.5. For some of our data, this value drops as low as 0.3” (2004, p. 1144). Similarly, Chung, Levi, and Legge (2001) list, in their table 1, prior studies yielding critical spacings as low as 0.1φ and as high as 0.5φ. Variability in measurements of critical spacing can arise from many sources. These include the arrangement of the elements in the display (Toet & Levi, 1992), the degree of similarity between the targets and distractors (Kooi et al., 1994), the duration of the display (Tripathy & Cavanagh, 2002), the predictability of the display’s location (Yeshurun & Rashal, 2010), and the amount of prior training of the observers (Chung, 2007). The outcome is also dependent on the method for computing critical spacing. At present, no single method can be taken as representing a gold standard. The clipped line fit (Pelli et al., 2004) gives comparatively large readings because it yields a critical distance that lies close to the shoulder of the performance-versus-distance function. The approach of fitting a continuous curve to the data and noting the point at which it intersects a criterial performance level (Tripathy & Cavanagh, 2002) yields comparatively small readings because, with commonly used criteria, the intersection occurs on the slope rather than at the shoulder of the performance-versus-distance function. Our approach falls into the latter category.

To determine how our measurements of critical spacing compare to results obtained in previous studies, we carried out detailed analyses of data from two studies employing methods similar to ours and reporting results independently for multiple individuals. Details are provided in the supplementary materials. Toet and Levi (1992) graphically depict, in their figure 6, the critical spacings of six individuals. Most relevant for comparison are horizontal critical spacings for displays centered on the horizontal meridian at eccentricities of 2.5°, 5°, and 10°. We found, by taking measurements directly from the figure, that the critical spacing as measured across all individuals and eccentricities had a mean of 0.27φ. Chung’s (2007) table 1 contains values from which the critical spacings of eight observers may be derived. We found that the critical spacings of individuals studied before intensive training, and thus comparable to our human observers, had a mean of 0.20φ. These studies examined crowding induced by two rather than four flanks. There is little difference between critical spacings measured under two-flanker and four-flanker conditions (Pelli et al., 2004). Nevertheless, we thought it worthwhile to compare our results to those of a study that also employed four flanks. The “same colour” data points in figure 2 of Pöder (2007) represent the average performance of seven observers required to identify a 0.5° letter flanked by four other letters at an eccentricity of 3.3°—a close match to the geometry of displays employed in Experiment 2 of our study. Applying our estimation procedure to these data points yielded a critical spacing of 0.23φ (see Supplementary Materials: Analysis of data from prior studies and Supplementary Figure S5). We conclude that our measurements of critical spacing are compatible with results reported previously.

In our study, both monkey and human observers exhibited some degree of flanker cost: Performance even at the largest spacing tested was worse than when targets were presented in isolation as singletons. Among the monkeys, the mean percentage correct for the largest spacing was 80%, as compared to 86% under the singleton condition, giving a flanker cost of 6%. Among the humans, the mean percentage correct for the largest spacing was 93%, as compared to 98% under the singleton condition, giving a flanker cost of 5%. Flanker cost could arise from any of several sources. It might reflect genuine albeit weak crowding arising when flankers encroach on the penumbra of the crowding field. It might arise from a reduction of attention to the target induced by the presence of flanks. It might arise from basing report on a flanker rather than on the target. Although we cannot identify the origin of the flanker cost in our study, we can ask...
whether it is comparable to the cost observed previously under parametrically equivalent conditions. In the monkeys of our study, the largest spacing was, on average across all experiments, 1.71 times the critical spacing. In our human observers, the ratio was 1.74. On the assumption that the flanker cost arises from genuine but weak crowding occurring when the flankers lie at the edge of the crowding field, we selected for comparison a prior study in which the ratio was approximately the same (Chung, 2007). Measurements taken from figure 4 and parameters taken from table 1 of that article indicate that the largest tested spacing was, on average across all eight observers, 1.76 times the critical spacing. Among observers in that study, the mean percentage correct score for the largest spacing was 82%, as compared to 99% under the singleton condition, giving a flanker cost of 16% with rounding error. Details are provided in the supplementary materials. We conclude that the flanker costs observed among monkeys and humans in our study did not exceed flanker costs observed in previous studies of crowding.

Crowding versus masking and attention

Crowding differs from ordinary masking in several ways, preeminent among which is the pattern of dependence on target eccentricity and size (Levi et al., 2002a; Pelli et al., 2004; Tripathy & Cavanagh, 2002). In crowding, the critical spacing scales with eccentricity independently of size, whereas in masking, the reverse is true. To a very close approximation, the results that we obtained in both monkeys and humans conformed to the pattern expected from crowding. In monkeys, there was a slight deviation from the ideal insofar as, when letters were presented at an eccentricity of 6°, the critical spacing was slightly greater if the target subtended 1° than if it subtended 0.5° (Figure 5). This effect was not significant in the combined data but did achieve significance in a post hoc test on data from Monkey 1. We do not believe that this argues against interpreting our results as due to crowding, for the following reasons. First, the results for both monkeys conformed much more closely to the pattern expected from crowding than to that expected from masking. Second, the dependence on target size was weak and inconsistent. Third, it has been observed in some studies of crowding in humans that the critical spacing increases slightly as target size increases (Levi et al., 2002a; Tripathy & Cavanagh, 2002).

Distinguishing the effects of crowding from the effects of attention is a difficult challenge because the two processes may be closely related. It has been hypothesized that the critical spacing arises from a limit on the spatial resolution of visual attention (Cavanagh, He, & Intriligator, 1999; He et al., 1996; Intriligator & Cavanagh, 2001). We know at present only that crowding and attention interact. For example, precuing the hemifield in which the display will appear (Yeshurun & Rashal, 2010) and planning an eye movement to the target (Harrison et al., 2013) shrink the critical spacing. The effects of attention and crowding are dissociable under special circumstances, for example when subjects judge average orientation across a group of Gabor patches (Dakin, Bex, Cass, & Watt, 2009). However, the design of our study did not allow a firm dissociation.

Limitations of the present study

The tasks presented in this article were designed to be performed by nonhuman primates in conjunction with neurophysiological recording. The design therefore differs in some regards from methods established in psychophysical studies on humans. We tested only two subjects of each species, we presented a limited library of target and flanker letters, and we required subjects to discriminate only between two orientations of each letter. Possibly as a result of these experimental choices, our results exhibited some patterns not typically present in psychophysical studies of crowding in humans.

First, flanker cost was highly variable across tasks and across subjects. Most strikingly, Monkey 2 exhibited a flanker cost of 19% during Experiment 2 (Figure 3D) but only 6% for Experiment 1 (Figure 2D) and 1% for Experiment 3 (Figure 4D). In contrast, Monkey 1 exhibited a sizable flanker cost for Experiment 1 (8%, Figure 2C), compared to smaller flanker costs for Experiment 2 (2%, Figure 3C) and Experiment 3 (3%, Figure 4C). Flanker cost varied across subjects and experiments in the human studies as well, but to a lesser degree (Supplementary Figure S3). It is possible that with more practice, greater motivation, or better sampling of target–distractor distances, the mean and variance of flanker cost would have decreased. Minimizing and stabilizing flanker cost would be a desirable goal for future experiments in nonhuman primates.

We also observed considerable variability in the lower asymptote of the psychometric function and in its slope. To accommodate this variability, we adopted a curve-fitting procedure that allowed these parameters to vary independently. Variability across individuals might have been less had we tailored letter size to the acuity limit of each subject (Chung, 2007). Additionally, we might have used a larger set of targets so as to prevent performance based on idiosyncratically selected diagnostic features. Steps such as these, aimed at equating to the greatest extent possible the performance...
of different individuals, would be a desirable feature of future experiments in nonhuman primates.

Finally, Monkey 1 exhibited a significant reduction in critical spacing when letter size was decreased with all other factors held constant. This effect suggests that crowding was confounded with masking when the larger letters closely abutted. It would be desirable, in future experiments, to avoid this problem through the use of relatively small targets.

Questions for future physiological inquiry

We know from human studies that crowding is a cortical phenomenon (Flom, Heath, & Takahashi, 1963a) increasing in magnitude along the ventral stream (Anderson, Dakin, Schwarzkoopf, Rees, & Greenwood, 2012; Bi, Cai, Zhou, & Fang, 2009). However, much remains to be learned about the underlying neuronal mechanisms. Having established that monkeys experience crowding, we are now in a position to attack these questions in experiments based on neuronal recording. There are at least two fundamental outstanding questions about the neuronal correlates of crowding that could be addressed in such experiments.

Does crowding arise from a reduction in the strength of the neuronal representation of the target? We take inferotemporal cortex as an example in terms of which to consider this question. This region is a logical candidate for the study of crowding because it is necessary for the efficient discrimination of letter-like images (Cowey & Gross, 1970) and because it contains neurons selective for letter-like stimuli (Sripati & Olson, 2010). We speculate that crowding is manifest among inferotemporal neurons in the form of a reduction in the strength of the signal encoding the identity of the target. The neuronal representation of an image is reduced by the simultaneous presence of other images even when they are separated by distances greater than those at which crowding occurs (Zhang et al., 2011; Zoccolan, Cox, & Dicarlo, 2005). The reduction takes a form well described in terms of divisive normalization (Carandini & Heeger, 2011). When images are far apart, neuronal activity representing the identity of a given item can be restored to almost normal strength by allocating top-down attention to it (Chelazzi, Duncan, Miller, & Desimone, 1998; Moran & Desimone, 1985; Zhang et al., 2011). When images are close together, this mechanism might fail either because bottom-up pooling of information about the target and flankers has rendered the target unavailable for independent selection (Parkes et al., 2001) or because the spatial resolution of top-down attention is limited (He et al., 1996). In either event, one would expect the neuronal signal representing the identity of the target to become progressively weaker, and attentional selection to become progressively less effective, as the elements of the display move closer to each other over the range of distances at which crowding operates. Testing this prediction would require no more than recording target-discriminating neuronal activity during performance of the crowding task.

Does crowding arise from a qualitative change in the representation of the target? In the simplest form of the bottom-up pooling model (Parkes et al., 2001) and in any model based on the limited spatial resolution of top-down attention (He et al., 1996), one would expect inferotemporal cortex neurons to fire at a rate representing the average of rates elicited by the separate elements of the display considered individually. However, there are other models in which the weakening of the representation of the target results from a more profoundly nonlinear interaction with the flankers. It has been proposed that the degradation of the target representation arises from the computation of textural statistics (Balas, Nakano, & Rosenholtz, 2009), from the illusory conjunction of features (Greenwood, Bex, & Dakin, 2010; Pelli et al., 2004; Pöder & Wagemans, 2007), from substitution of a flanker for a target (Freeman, Chakravarthi, & Pelli, 2012), and from a combination of these processes (Hanus & Vul, 2013). Each of these scenarios gives rise to specific predictions, potentially testable in a neuronal recording experiment, concerning the impact of the presence of the flankers on neuronal selectivity for the target.

Conclusions

1. The ability of monkeys to identify a visual target is progressively reduced as flankers move closer to it.
2. The critical spacing at which flankers begin to interfere with recognition varies in proportion to eccentricity.
3. At a given eccentricity, the critical spacing does not scale with object size.
4. Given that these phenomena are hallmarks of crowding in humans, we conclude that monkeys exhibit crowding.
5. These findings establish the feasibility of studying crowding at the level of single neurons in the primate visual system.

Keywords: crowding, monkey, peripheral vision, psychophysics

Acknowledgments

This work was supported by National Institutes of Health grants EY018620 and P50 MH084053. Addi-
tionally, EAC was supported by the National Science Foundation Graduate Research Fellowship, grant number 1247842. EAC was also supported by the Department of Defense National Defense Science and Engineering Graduate Fellowship, under FA9550-05-C-0059.

Commercial relationships: none.
Corresponding author: Carl R. Olson.
Email: colson@cnbc.cmu.edu.
Address: Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA, USA.

References


