Comparing crowding in human and ideal observers

Ronald van den Berg — Department of Neuroscience, Baylor College of Medicine, Houston, USA; Laboratory for Experimental Ophthalmology, University of Groningen, University Medical Center Groningen, Netherlands

Addie Johnson — Department of Psychology, University of Groningen, Groningen, Netherlands

Angela Martinez Anton — Department of Psychology, University of Groningen, Groningen, Netherlands; Laboratory for Experimental Ophthalmology, University of Groningen, University Medical Center Groningen, Netherlands

Anne L. Schepers — Department of Psychology, University of Groningen, Groningen, Netherlands; Laboratory for Experimental Ophthalmology, University of Groningen, University Medical Center Groningen, Netherlands

Frans W. Cornelissen — Department of Psychology, University of Groningen, Groningen, Netherlands; Laboratory for Experimental Ophthalmology, University of Groningen, University Medical Center Groningen, Netherlands

A visual target is more difficult to recognize when it is surrounded by other, similar objects. This breakdown in object recognition is known as crowding. Despite a long history of experimental work, computational models of crowding are still sparse. Specifically, few studies have examined crowding using an ideal-observer approach. Here, we compare crowding in ideal observers with crowding in humans. We derived an ideal-observer model for target identification under conditions of position and identity uncertainty. Simulations showed that this model reproduces the hallmark of crowding, namely a critical spacing that scales with viewing eccentricity. To examine how well the model fits quantitatively to human data, we performed three experiments. In Experiments 1 and 2, we measured observers' perceptual uncertainty about stimulus positions and identities, respectively, for a target in isolation. In Experiment 3, observers identified a target that was flanked by two distractors. We found that about half of the errors in Experiment 3 could be accounted for by the perceptual uncertainty measured in Experiments 1 and 2. The remainder of the errors could be accounted for by assuming that uncertainty (i.e., the width of internal noise distribution) about stimulus positions and identities depends on flanker proximity. Our results provide a mathematical restatement of the crowding problem and support the hypothesis that crowding behavior is a sign of optimality rather than a perceptual defect.

Keywords: crowding, substitution, feature integration, modeling, uncertainty, ideal observer

Citation: van den Berg, R., Johnson, A., Martinez Anton, A., Schepers, A. L., & Cornelissen, F. W. (2012). Comparing crowding in human and ideal observers. *Journal of Vision, 12*(6):13, 1–14, http://www.journalofvision.org/content/12/6/13, doi:10.1167/12.6.13.

Introduction

Numerous studies have demonstrated that a visual target is more difficult to identify when other, similar objects surround it (Figure 1). This effect is called visual crowding and fundamentally limits conscious visual perception. Crowding is typically quantified in terms of contrast threshold elevation, that is, the increase in target contrast required to equalize recognition performance for a flanked target with that of the same target presented in isolation (Levi, 2008). The largest target-flanker spacing at which such threshold elevation occurs is commonly called the critical spacing.

doi: 10.1167/12.6.13 Received September 4, 2011; published June 12, 2012

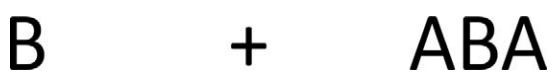

Figure 1. Example of crowding. The two Bs are at equal distance from the "+." When fixating the "+," the isolated B on the left is easy to recognize, whereas the B on the right appears to be jumbled with the flanking As. This is called crowding.

It has been found that critical spacing is proportional to the retinal eccentricity of the target (Bouma, 1970; Pelli, Palomares, & Majaj, 2004) and independent of feature (van den Berg, Roerdink, & Cornelissen, 2007). This proportionality is a hallmark of crowding and distinguishes it from visual masking (Levi, 2008; Pelli et al., 2004).

A long-standing scientific interest in the crowding phenomenon (Bouma, 1970; Korte, 1923; Stuart & Burian, 1962) has produced a wealth of experimental data and theories (Levi, 2008). A prevalent theory argues that crowding is a result of integrating features over inappropriately large areas (Greenwood, Bex, & Dakin, 2010; Levi & Carney, 2009; Pelli et al., 2004; Pelli & Tillman, 2008; van den Berg, Roerdink, & Cornelissen, 2010), but it has also been suggested that crowding is primarily a result of substituting object positions due to spatial uncertainty (Chung & Legge, 2009; Greenwood, Bex, & Dakin, 2009; Huckauf & Heller, 2002; Nandy & Tjan, 2007; Strasburger, 2005) or of attentional limitations (Chakravarthi & Cavanagh, 2007; He, Cavanagh, & Intriligator, 1996; Petrov & Popple, 2007). Despite the abundance of data and theories, quantitative models for crowding are sparse (Whitney & Levi, 2011) and have been put forward only recently (Dakin, Cass, Greenwood, & Bex, 2010; Greenwood et al., 2009; Nandy & Tjan, 2007; Solomon & Dayan, 2011; van den Berg et al., 2010). Whereas several researchers have proposed that crowding may be explained using ideal-observer theory, as yet it is unclear to what extent the behavioral hallmark of crowding, namely critical spacing that scales with target eccentricity, would be compatible with an ideal-observer account of crowding.

To address this question, we derived an ideal-observer model for object identification in the presence of flankers and perceptual noise. Using simulations, we examined the qualitative behavior of this model and found that it produces behavior qualitatively similar to the behavior typically found with human observers. To assess how well the model quantitatively accounts for human data, we performed three experiments. In the first two experiments, we measured identification and localization performance for objects presented in isolation at a range of stimulus eccentricities and with a range of contrasts. In the third experiment, we measured target-identification performance in the presence of flankers. To preview our results, we found that about half of the errors in the third experiment could be explained by the perceptual uncertainty measured in Experiments 1 and 2. The remainder of the errors could not be explained by an increase in either position or identity uncertainty alone, but a model in which both types of uncertainty depend on flanker proximity did accurately account for the data.

Model

Derivation

Our model is based on the notion that humans are noisy observers (Faisal, Selen, & Wolpert, 2008) in the sense that stimulus positions and identities are never perfectly encoded but always involve some noise. The ideal observer is the observer who performs as well as is theoretically possible in the presence of such noise. To qualitatively assess whether ideal observers exhibit crowding, we derived a model for a simple target-identification task with one to-be-ignored flanker. The two stimuli are positioned at adjacent locations, one of which is defined as the target location (Figures 2a and b). The task is to report the identity of the target. Because both the positions and identities of stimuli are encoded in a noisy manner, the ideal observer will occasionally confuse the object positions or misinterpret the identity of the target. We refer to these errors as "position confusions" (PC) and "identity confusions" (IC), respectively.

The probability of a PC can be computed as follows. Let the target be defined as the stimulus closest to fixation. Then the probability of a PC is equal to the probability of perceiving the flanker closer to fixation than the target. If we assume Gaussian noise in the encoded positions, then this probability can be expressed as follows:

$$\begin{aligned} p(\mathrm{PC}|\boldsymbol{\varphi},\boldsymbol{\sigma}) &= p(x_\mathrm{D} < x_\mathrm{T}|\boldsymbol{\varphi},\boldsymbol{\sigma}) \\ &= p(x_\mathrm{D} - x_\mathrm{T} < 0|\boldsymbol{\varphi},\boldsymbol{\sigma}) \\ &= \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\varphi_\mathrm{T} - \varphi_\mathrm{D}}{\sqrt{2(\sigma_\mathrm{D}^2 + \sigma_\mathrm{T}^2)}}\right), \end{aligned}$$

where $\boldsymbol{\varphi} = (\varphi_\mathrm{T}, \varphi_\mathrm{D})$ represents the (one-dimensional) positions of the target and distractor, respectively, and $\boldsymbol{\sigma} = (\sigma_\mathrm{T}, \sigma_\mathrm{D})$ the spatial uncertainty at these locations, and $\mathbf{x} = (x_\mathrm{T}, x_\mathrm{D})$ the internal representations of the stimulus positions. The probability of an IC equals the probability of reporting a "B" when the target (or, in case of a PC, the flanker) is an "A," or vice versa. This is equal to 1 minus the probability of reporting the correct stimulus. We assume that the probability of reporting the correct stimulus follows a scaled and

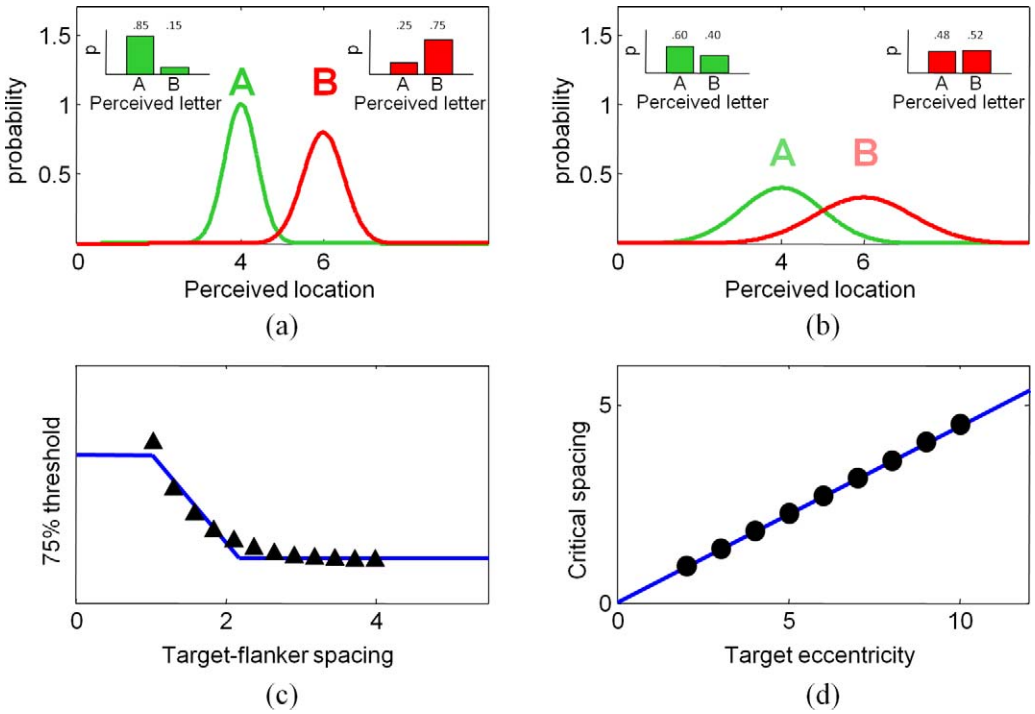

Figure 2. An ideal-observer model for target identification under perceptual uncertainty. In this example, we define the leftmost stimulus as the target and the other as the flanker. (a) We assume that perceived stimulus positions follow a Gaussian distribution around the actual positions (spatial uncertainty; represented by the green and red probability distribution functions) and that an "A" is sometimes perceived as a "B" and vice versa (identity uncertainty; represented by the discrete probability distributions in the insets). When stimulus contrast is high, both kinds of uncertainty are low, and few errors will be made when asked to report the identity of the target. (b) When stimulus contrast is low, uncertainty about the identities and positions of the stimuli is high. As a result, observers will more often confuse stimulus positions and identities with each other, leading to an increased number of errors. (c) Simulation data from an identification task with a target presented at 6°. Performance was estimated for a range of target-flanker spacings, and 75% performance thresholds were determined. In this example, fitting a clipped ramp function (blue curve) gives an estimated critical spacing of 2.2°. (d) Assuming that positional uncertainty (reflected in the widths of the curves in panels a and b) is proportional to eccentricity, a linear relation between critical spacing and target eccentricity follows, which is reminiscent of crowding behavior in human observers.

translated cumulative normal distribution as a function of the logarithm of stimulus contrast:

$$F(c;\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2)=\frac{1}{2}+\frac{1}{2}\int_{-\infty}^{\ln c}\frac{1}{\sqrt{2\pi\sigma_{\mathrm{IC}}^2}}e^{-\frac{(x-\mu_{\mathrm{IC}})^2}{2\sigma_{\mathrm{IC}}^2}}dx,$$

where μ_{IC} and σ_{IC} are the mean and width of the cumulative normal function and c the contrast of the encoded stimulus (which, as will be shown later, is a good description of human performance; Figure 5). Hence, the probability of an IC can be written as

$$\begin{aligned}p(\mathrm{IC}|c,\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2)&=1-F(c;\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2)\\&=\frac{1}{2}-\frac{1}{2}\int_{-\infty}^{\ln c}\frac{1}{\sqrt{2\pi\sigma_{\mathrm{IC}}^2}}e^{-\frac{(x-\mu_{\mathrm{IC}})^2}{2\sigma_{\mathrm{IC}}^2}}dx.\end{aligned}$$

When stimulus contrast is close to 0, the probability of reporting the wrong stimulus is equal to chance (0.5; i.e., the observer must guess); when stimulus contrast is high, the probability of an IC approaches 0.

The probability of reporting the wrong stimulus on a trial in which no PC occurs is equal to the probability of incorrectly decoding the target signal (i.e., to the probability that an IC occurs):

$$p(\mathrm{error}|\text{no PC},c_{\mathrm{T}},\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2)=p(\mathrm{IC}|c_{\mathrm{T}},\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2),$$

where c_{T} is the contrast of the target. Responses on trials in which a PC has occurred are based on the signal from the flanker. Because there are only two possible stimulus identities and the target and distractor are always different from each other in the experiment that we are modeling, this leads to an error only if the flanker signal is decoded correctly, that is, if no IC occurs. Hence, we get

$$p(\mathrm{error}|\mathrm{PC},c_{\mathrm{F}},\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2)=1-p(\mathrm{IC}|c_{\mathrm{F}},\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2),$$

where c_{F} denotes the contrast of the flanker. The total

probability of an error being made is thus:

$$p(\text{error}|\boldsymbol{\varphi}, \mathbf{c}, \mu_{\text{IC}}, \sigma_{\text{IC}}^2) = p(\text{no PC}|\boldsymbol{\varphi}, \boldsymbol{\sigma}) p(\text{IC}|c_{\text{T}}, \mu_{\text{IC}}, \sigma_{\text{IC}}^2) + p(\text{PC}|\boldsymbol{\varphi}, \boldsymbol{\sigma}) \times \left(1 - p(\text{IC}|c_{\text{F}}, \mu_{\text{IC}}, \sigma_{\text{IC}}^2)\right). \quad (1)$$

Simulations

In the simulations conducted to study the qualitative behavior of this model, flanker contrast, c_{F}, was set to 20%. (All contrasts throughout the paper are specified as the Weber contrast.) The width of the positional noise distribution was set at 1° for both the target and the flanker, i.e., $\sigma_{\text{T}} = \sigma_{\text{F}} = 1°$. The two parameters that determine the probability of an identity confusion, μ_{IC} and σ_{IC}, were set to 10% and 1%, respectively. Using Equation 1, we computed performance as a function of target contrast, c_{T}, for a range of target-flanker spacings. From the resulting curves, we estimated the target-contrast thresholds that yield 75% correct performance. Plotting these thresholds as a function of stimulus spacing (Figure 2c) revealed a curve qualitatively similar to those typically found when human subjects perform target identification under conditions of crowding; the identification threshold declines as spacing is increased, up until a "critical" spacing beyond which the identification threshold is more or less stable. A trilinear function was fit to these thresholds (Figure 2c) to estimate the critical spacing, as has been done in previous studies (Pelli et al., 2004; van den Berg et al., 2007).

The previous results indicate that critical spacing is an inherent property of ideal observers, even when assuming that the standard deviation of the positional noise distribution does not depend on the retinal eccentricity of the stimulus. A recent paper, however, reported that positional uncertainty is described well by a Gaussian noise distribution with a standard deviation that depends linearly on eccentricity (Michel & Geisler, 2011). Using this assumption of linearity, we repeated our simulations for several target eccentricities, φ_{T}, yielding an estimate of critical spacing as a function of target eccentricity (Figure 2d). We found that the model produces a critical spacing that scales with eccentricity. Hence, the behavioral hallmark of crowding (Levi, 2008; Pelli et al., 2004) seems to be a qualitative property of any ideal observer who performs target identification in the face of visual uncertainty.

These results show that, in theory, one merely needs to assume that perception is noisy and that humans behave as ideal observers to explain critical spacing in crowding. Whether this is a good model for describing human-observer data depends on whether the levels of uncertainty in the model match the levels of uncertainty found in human subjects—something rarely measured in crowding studies. Therefore, prior to measuring subjects' behavior in a crowded condition (Experiment 3), we first measured their levels of perceptual uncertainty for targets presented in isolation (Experiments 1 and 2).

Experiment 1

In Experiment 1, we measured spatial uncertainty for isolated targets as a function of both viewing eccentricity and stimulus contrast.

Observers

Six human observers (4 male, 2 female; age range, 21 to 28 years) participated. One of the observers was an author (AM); the other five were paid students who were naive regarding the purpose of the experiment. All observers had normal or corrected-to-normal vision.

Apparatus

Stimuli were generated and data collected using Mathworks™ Matlab in combination with the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). Stimuli were displayed on a 22-inch LaCie RGB monitor with a frame rate of 60 Hz. Prior to the experiments, we used a Konica Minolta CS-100A colorimeter to measure the relationship between the monitor's RGB input triplets (grayscale, R = G = B) and its luminance output (in cd/m^2). The viewing distance was 60 cm. Luminance of the (gray) background screen was 35 cd/m^2. A chin rest was used to minimize head movements.

Stimuli and procedure

The stimuli in Experiment 1 consisted of the letters A through F presented in the Sloan font (Pelli & Robson, 1988). On each trial, a letter was selected randomly and presented to both the left and right of a central fixation mark at different eccentricities (Figure 3a). Stimuli were presented for 83 ms. The task of the observer was to indicate which of the two letters was closest to the fixation point. On each trial, the eccentricities at which the letters were presented were picked as follows: A base eccentricity, $\bar{\varphi}$, was chosen from the set {3.0, 4.25, 5.50, 6.75, 8.0}°, and a difference value, Δ_{φ}, from a set of 10 positive and 10 negative values (with $|\Delta_{\varphi}|$ logarithmically spaced). Subsequently, the eccentricities

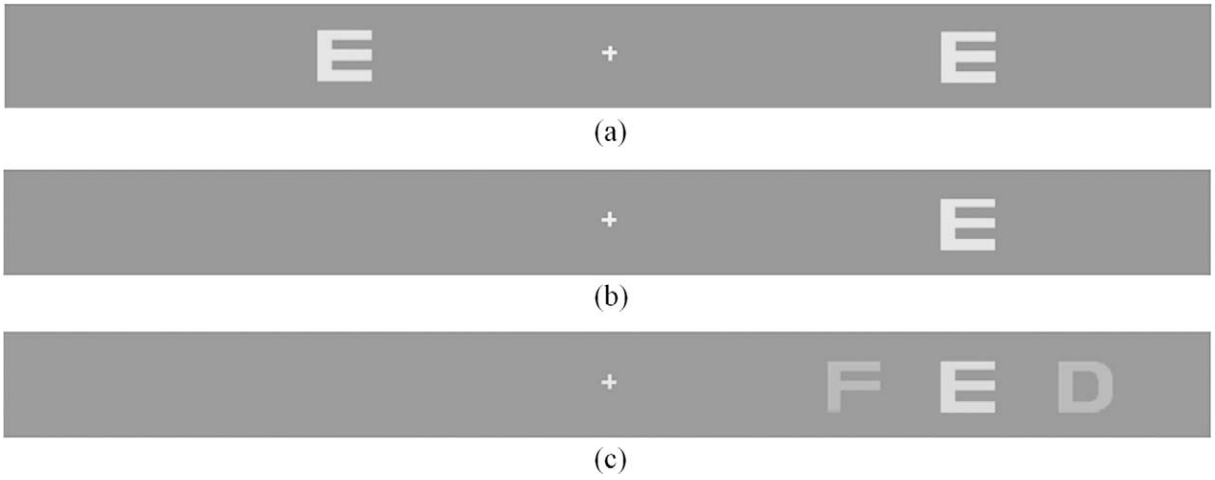

Figure 3. Examples of stimuli used in the three experiments. (a) In Experiment 1, aimed at estimating location uncertainty, observers indicated which of two peripheral, briefly presented stimuli was closest to fixation. (b) In Experiment 2, aimed at estimating identity uncertainty, observers reported the identity of the target. (c) In Experiment 3, aimed at measuring crowding effects, the task was to report the identity of a target (defined as "the central letter") surrounded by two flankers.

of the left and right stimulus were set to $\varphi_L = \bar{\varphi} - \frac{1}{2}\Delta_\varphi$ and $\varphi_R = \bar{\varphi} + \frac{1}{2}\Delta_\varphi$, respectively. (Hence, the left stimulus was closer to fixation on trials with $\Delta_\varphi > 0$ and the right stimulus on trials with $\Delta_\varphi < 0$.) To reduce the number of values of Δ_φ at which subjects perform at floor or ceiling performance, the boundaries of the range of $|\Delta_\varphi|$ were automatically updated after each 300 trials. If an observer's mean performance over the last 300 trials was below 70%, then the maximum value was increased by 10%. If the mean performance was above 80%, the minimum value was decreased by 10%. Following this change, a new set of 20 values for Δ_φ was computed.

Stimuli were presented at one of five possible contrast values (logarithmically spaced), randomly chosen on each trial. Both stimuli were presented at the same contrast. Maximum stimulus contrast was set to 50% for all observers. To minimize detection errors, the minimum contrast was set to 1.1 times an observer's 99%-correct detection contrast threshold. An observer's contrast threshold was determined prior to the main experiment using a detection task in which a random letter was presented at 6° eccentricity randomly to the left or right of fixation, at random contrast. The observer reported whether the stimulus appeared at the left or right side. A cumulative Gaussian was fit to performance as a function of (log) contrast. The thus found 99%-correct detection contrasts thresholds were 8.7%, 18.5%, 16.6%, 16.6%, 12.8%, and 13.3% for subjects AM, EK, GK, KN, SF, and XL, respectively.

After each trial, observers were given feedback through a change in the color of the otherwise gray fixation mark to green (correct) or red (error) during the intertrial interval. Each observer completed 15 blocks of 150 trials each.

Results

Estimates of observers' positional uncertainty were computed by fitting an ideal observer model to the data. Under the assumption that observed stimulus locations are normally distributed around the real stimulus location, the probability that the perceived location of the stimulus on the left is closer to fixation than the perceived location of the stimulus on the right equals

$$p(x_L < x_R | \varphi_L, \varphi_R, \sigma_L, \sigma_R) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\varphi_R - \varphi_L}{\sqrt{2(\sigma_L^2 + \sigma_R^2)}}\right),$$

where φ_L and φ_R denote the eccentricities of the stimuli on the left and the right, respectively; x_L and x_R denote their internal representations; and σ_L and σ_R denote the trial-to-trial variability over these internal representations.

Previous studies have found that the noise distribution for perceived positions is well-described by a Gaussian distribution with a standard deviation linearly related to stimulus eccentricity (Michel & Geisler, 2011) and weakly depends on stimulus contrast (Hess & Holliday, 1992). To verify whether this holds for the data of Experiment 1, we compared four different functions relating spatial uncertainty (σ) to retinal eccentricity (φ) and contrast (c) (Table 1). Bayesian model comparison (see Appendix for details) revealed that the model that assumes a linear relationship between σ and stimulus eccentricity, and a power-law relationship between σ and stimulus contrast, provides the most likely description of the data.

Function number	Relationship between σ and φ	Relationship between σ and c	Function
1	Linear	Linear	$\sigma(\varphi, c) = a\varphi c$
2	Linear	Power law	$\sigma(\varphi, c) = a\varphi c^t$
3	Power law	Linear	$\sigma(\varphi, c) = a\varphi^s c$
4	Power law	Power law	$\sigma(\varphi, c) = a\varphi^s c^t$

Table 1. Overview of the functions that were examined to relate spatial uncertainty to stimulus eccentricity and contrast.

Specifically, this function (Function 2 in Table 1) outperforms the first, third, and fourth functions shown in Table 1, with 116 ± 28, 118 ± 28, and 4.3 ± 0.3 log-likelihood points, respectively. Note that a difference of 4.3 means that one model is exp(4.3) ≈ 74 times more likely than the other, given the dataset.

For each subject, we computed maximum-likelihood values for parameters a and t under Function 2 in Table 1 (see Appendix for details). Data and model fits are shown in Figure 4. The maximum likelihood values of the parameters were $a = 0.30 \pm 0.07$ and $t = -0.33 \pm 0.06$ (mean ±1 *SEM*).

Experiment 2

Methods

In Experiment 2, we measured observers' uncertainty about stimulus identity as a function of stimulus contrast. The observers, apparatus, and general procedure in Experiment 2 were identical to those in Experiment 1, except for the following differences: Only one stimulus was presented on each trial, and the task was to report the identity of this stimulus (Figure 3b). To discourage observers from making anticipatory eye movements, the stimulus was presented to the left or right side of fixation with equal probability. The stimulus was presented at 6° eccentricity at one of 12 contrasts in the range 1.5% to 24% (logarithmically spaced). Each observer completed 12 blocks of 150 trials.

Results

The probability of an error decreased with increasing contrast. We fitted these data with a cumulative Gaussian with a baseline of 1/6 (chance performance):

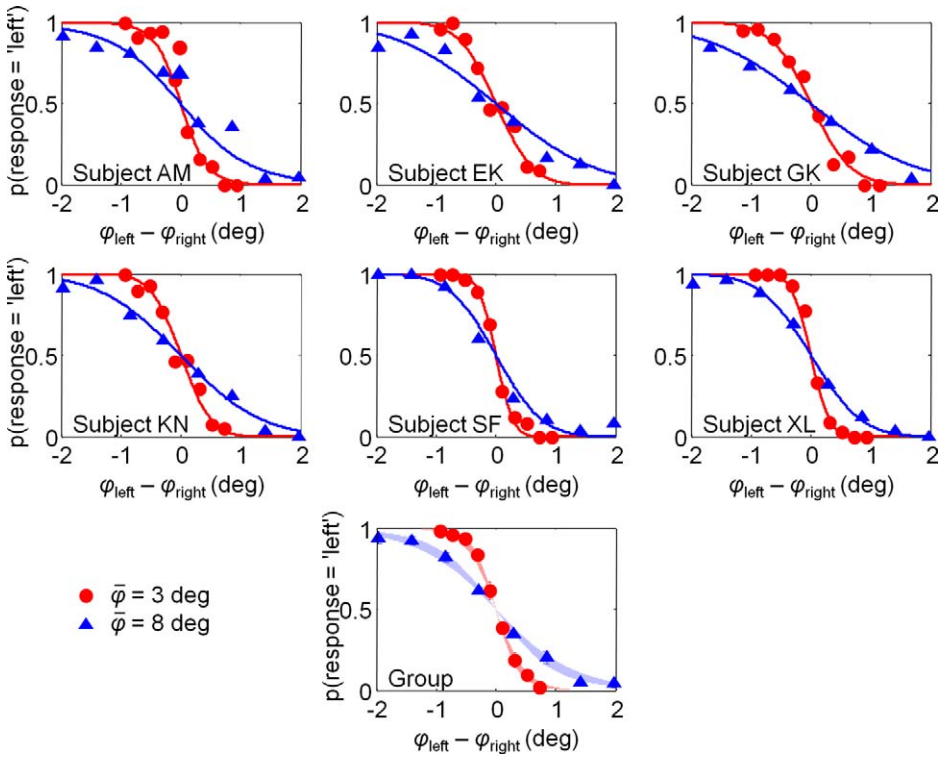

Figure 4. Spatial uncertainty estimates from Experiment 1. The probability of responding that the left stimulus was closer to fixation decreases as a function of the difference in stimulus eccentricity. The data are well fit by an ideal-observer model with a Gaussian noise distribution. Data are from individual subjects (top and middle rows) and the mean over all subjects (bottom row). To avoid clutter, only the results for the smallest and largest value of $\bar{\varphi}$ are shown. The error bars and shaded areas in the panel with the group results indicate 1 *SEM*. Note that most error bars are invisible, due to being smaller than the marker.

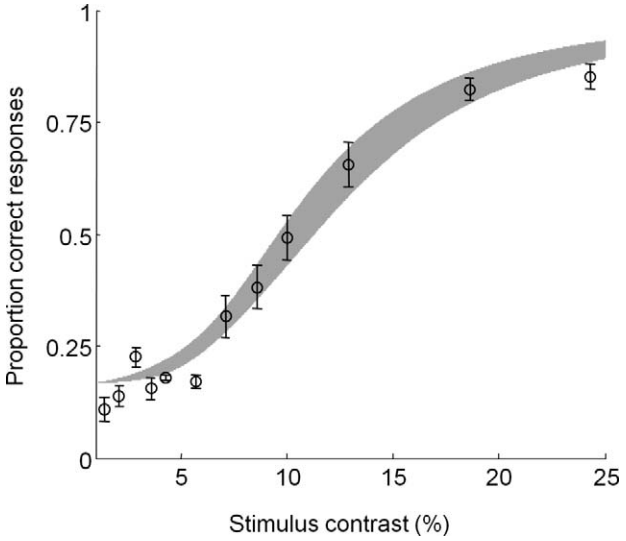

Figure 5. Identification performance as a function of stimulus contrast (open circles) and the fit (shaded area) of a cumulative Gaussian with a baseline of chance performance (1/6), pooled across subjects and target letters. Error bars and shaded area represent 1 *SEM*.

$$p(\text{correct identification}|c, \mu_{IC}, \sigma_{IC}^2) = \frac{1}{6} + \frac{5}{6}\int_{-\infty}^{\ln c} \frac{1}{\sqrt{2\pi\sigma_{IC}^2}} e^{-\frac{(x-\mu_{IC})^2}{2\sigma_{IC}^2}} dx. \quad (2)$$

This function provided a good fit to the data of each subject. Pooled results are shown in Figure 5.

Experiment 3

Methods

In Experiment 3, we measured target identification performance in the presence of flankers. The experiment was identical to Experiment 2, except for the following differences: The target was flanked on either side by a letter randomly chosen from the range A through F. The flankers were always different from the target and from each other. The center-to-center spacing between the target and flankers was randomly chosen on each trial from a set of 12 values in the range 0.9 to 4.5°. The contrast of both flankers was set to 1.5 times the minimum contrast used in Experiment 1. The contrast of the target was randomly chosen on each trial from a set of 12 linearly spaced contrast values with minimum and maximum values as in Experiment 1. Each observer completed 15 blocks of 150 trials each.

Results

For each observer and target-flanker spacing in Experiment 3, we computed the proportion of trials on which the observer reported a nontarget letter and the proportion of trials on which the observer reported a flanker letter. We call the former "errors" and the latter "flanker correspondences" (Strasburger, 2005). Note that the flanker correspondences are a subset of the errors. Analyzing flanker correspondences may help to disentangle the relative contributions of position and identity uncertainty from crowding, because these two types of uncertainty have different effects on the flanker correspondences. For example, if identity and position confusion occur independently of each other, then every position confusion will cause a flanker correspondence, whereas (in the present task) only two out of every five identity confusions will result in a flanker correspondence.

Figure 6a shows the proportions of error and flanker-correspondence trials as a function of target-flanker spacing. Both are inversely related to spacing. Figure 6b shows the flanker correspondences as a proportion of the errors. In line with a previous study (Strasburger, 2005), the actual number of flanker correspondences was substantially higher than chance level, especially at small spacings. Figure 6c shows target-contrast thresholds as a function of spacing. We fitted a trilinear function to these data and found a critical spacing of approximately 3.5°, which is 0.58 times the target eccentricity.

Model derivation for Experiment 3

In this section, we derive a model for the task in Experiment 3. We first explain how we modeled the probabilities of position and identity confusions. Thereafter, we derive expressions for the error and flanker-correspondence probabilities in Experiment 3.

Modeling the probabilities of position and identity confusions

Because the target in Experiment 3 was defined as "the central stimulus," a PC occurred when one of the flankers was perceived at the central location. To express this mathematically, let $\boldsymbol{\varphi} = (\varphi_{IN}, \varphi_T, \varphi_{OUT})$ denote the locations of the foveal (inward) flanker, target, and peripheral (outward) flanker, respectively; x_{IN}, x_T, and x_{OUT}, their normally distributed internal representations; and $\mathbf{c} = (c_{IN}, c_T, c_{OUT})$ their contrasts.

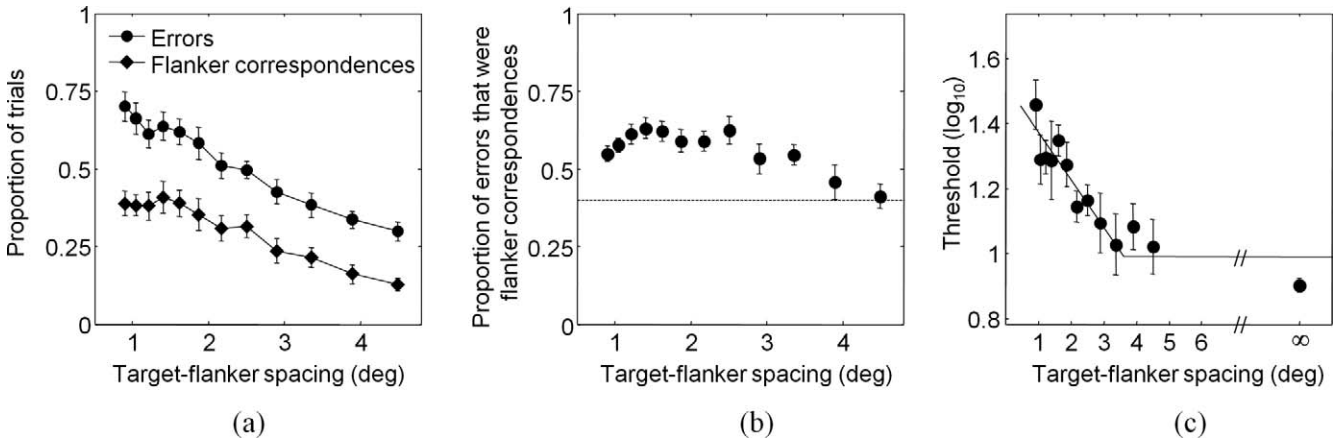

Figure 6. Subject data from Experiment 3. (a) The proportions of errors and flanker correspondences decrease as a function of target-flanker spacing. Note that no clear critical spacing can be observed. (b) For small target-flanker spacings, the proportion of errors that were flanker correspondences is well above chance level (dashed line). (c) Fitting a clipped-line curve to the target-identification thresholds gives an estimated critical spacing of approximately 3.5°. Error bars represent 1 *SEM*. Infinity represents the condition without flankers.

Then the probability of a PC in Experiment 3 equals

$$p(\mathrm{PC}|\boldsymbol{\varphi},\mathbf{c}) = 1 - p(x_{\mathrm{IN}} < x_{\mathrm{T}} < x_{\mathrm{OUT}}|\boldsymbol{\varphi},\boldsymbol{\sigma}) - p(x_{\mathrm{OUT}} < x_{\mathrm{T}} < x_{\mathrm{IN}}|\boldsymbol{\varphi},\boldsymbol{\sigma}), \quad (3)$$

where $\boldsymbol{\sigma} = (\sigma_{\mathrm{IN}}, \sigma_{\mathrm{T}}, \sigma_{\mathrm{OUT}})$ is the trial-to-trial variability in the internal representations of the stimuli as estimated from Experiment 1. Evaluating Equation 3 involves taking the integral of the error function, for which no closed-form solution exists. Therefore, we evaluated this equation numerically using Monte Carlo simulation.

Because we do not know what form the noise distribution over letter identities takes, we cannot derive the probability of an IC analytically. Therefore, we approximated the probability of an IC as the probability of misperceiving the identity of a target as measured in Experiment 2. Hence, the probability of an IC can be specified using Equation 2:

$$p(\mathrm{IC}|c,\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2) = 1 - p(\text{correct identification}|c,\mu_{\mathrm{IC}},\sigma_{\mathrm{IC}}^2). \quad (4)$$

Modeling error rates in Experiment 3

Following the same logic as in the derivation in the Model section, we find the following expressions for the error probabilities on trials with and without a target-position confusion:

$$p(\mathrm{error}|\boldsymbol{\varphi},\mathbf{c},\boldsymbol{\theta},\text{no PC}) = p(\mathrm{IC}|c_{\mathrm{T}},\boldsymbol{\theta})$$

and

$$p(\mathrm{error}|\boldsymbol{\varphi},\mathbf{c},\boldsymbol{\theta},\mathrm{PC}) = 1 - \frac{1}{5}p(\mathrm{IC}|c_{\mathrm{F}},\boldsymbol{\theta}),$$

where $\boldsymbol{\theta} = (a, t, \mu_{\mathrm{IC}}, \sigma_{\mathrm{IC}})$ is a vector containing all model parameters. Combining both cases, the total probability of an error response in Experiment 3 equals

$$p(\mathrm{error}|\boldsymbol{\varphi},\mathbf{c},\boldsymbol{\theta}) = p(\text{no PC}|\boldsymbol{\varphi},\mathbf{c},\boldsymbol{\theta})p(\mathrm{IC}|c_{\mathrm{T}},\boldsymbol{\theta}) + p(\mathrm{PC}|\boldsymbol{\varphi},\mathbf{c},\boldsymbol{\theta})\left(1 - \frac{1}{5}p(\mathrm{IC}|c_{\mathrm{F}},\boldsymbol{\theta})\right). \quad (5)$$

Modeling flanker correspondence rates in Experiment 3

The expected proportion of flanker correspondences (FCs) was modeled as follows. For trials on which no position confusion occurs, the probability of reporting a flanker letter equals the probability of erroneously decoding the target signal to a letter that happened to be the same as one of the flankers. Averaged over all possible target and flanker letter combinations, this probability equals

$$p(\mathrm{FC}|\text{no PC},c_{\mathrm{T}},\boldsymbol{\theta}) = \frac{2}{5}p(\mathrm{IC}|c_{\mathrm{T}},\boldsymbol{\theta}).$$

For trials on which a position confusion occurred, the probability of reporting a flanker equals the probability of correctly decoding the flanker signal plus the probability of incorrectly decoding it to the letter that happens to be the same as the other flanker. Averaged over all possible targets and flankers, this probability equals

$$p(\mathrm{FC}|\mathrm{PC},c_{\mathrm{F}},\boldsymbol{\theta}) = p(\text{no IC}|c_{\mathrm{F}},\boldsymbol{\theta}) + \frac{1}{5}p(\mathrm{IC}|c_{\mathrm{F}},\boldsymbol{\theta}).$$

Combining both cases, the overall probability of a flanker correspondence equals

$$p(\mathrm{FC}|\boldsymbol{\varphi},\mathbf{c},\boldsymbol{\theta}) = \frac{2}{5}p(\text{no PC}|\boldsymbol{\varphi},\mathbf{c},\boldsymbol{\theta})p(\mathrm{IC}|c_\mathrm{T},\boldsymbol{\theta}) + p(\mathrm{PC}|\boldsymbol{\varphi},\mathbf{c},\boldsymbol{\theta})\left[p(\text{no IC}|c_\mathrm{F},\boldsymbol{\theta}) + \frac{1}{5}p(\mathrm{IC}|c_\mathrm{F},\boldsymbol{\theta})\right].$$

Model fits to Experiment 3

Predicted versus empirical error and flanker correspondences rates

With the model parameters, $\boldsymbol{\theta} = (a,\ t,\ \mu_\mathrm{IC},\ \sigma_\mathrm{IC})$, fixed to the values estimated from Experiments 1 and 2, we computed the expected proportions of errors and flanker correspondences in Experiment 3. We found that both the proportion of errors and the proportion of flanker correspondences are strongly underestimated (Figure 7a). Specifically, we found that this model explains 44% ± 5.2% of the errors in Experiment 3 (Figure 7b). Hence, whereas the sensory uncertainty measured in Experiments 1 and 2 can explain a major part of the crowding effect, it also leaves a large part unexplained.

If we assume that subjects behave as ideal observers in this task, then our finding that only half of the errors are accounted for by the perceptual uncertainty measured in Experiments 1 and 2 suggests that the addition of the flankers caused an increase in uncertainty about stimulus positions and/or identities. To examine whether this could be the case and to disentangle the effects of spatial and identity uncertainty, we adapted the model to test the following three hypotheses: (1) Flanker proximity affects positional uncertainty; (2) Flanker proximity affects identity uncertainty; and (3) Flanker proximity affects both positional and identity uncertainty.

Hypothesis 1: flankers proximity affects positional uncertainty

To test whether a model in which flanker proximity affects positional uncertainty can account for the data from Experiment 3, we let parameter a in the equation that relates stimulus contrast and eccentricity to position uncertainty, $\sigma(\varphi,c) = a\varphi c^t$, depend on stimulus spacing. We did this using a power law function, because it satisfies the following conditions: (1) For very large spacing, position uncertainty converges to the case without flankers, as measured in Experiment 1; (2) The relation between position uncertainty and stimulus spacing is monotonic, but not otherwise strongly constrained; and (3) The number of additional model parameters is small. Specifically, we let a be the following function of two parameters, α and β, and target-flanker spacing, $\varphi_\mathrm{OUT} - \varphi_\mathrm{T}$:

$$a = a_\mathrm{I}\left[1 + \alpha(\varphi_\mathrm{OUT} - \varphi_\mathrm{T})^\beta\right],$$

where a_I is the value of a for isolated stimuli as estimated from Experiment 1. For negative $\beta \leq 0$, positional uncertainty converges to a_I when spacing is large.

Model parameters were fitted using maximum-likelihood fitting (see Appendix). We fitted the parameters to explain performance accuracy (correct/incorrect) rather than the actual responses (the letters A through F). As a result, flanker correspondence rates, the computation of which depends on the actual responses, were not available during model fitting. Hence, the flanker correspondence rates can be considered independent model predictions.

As shown in Figure 8 (top row, left), the mean error rates in Experiment 3 can be accounted for by assuming

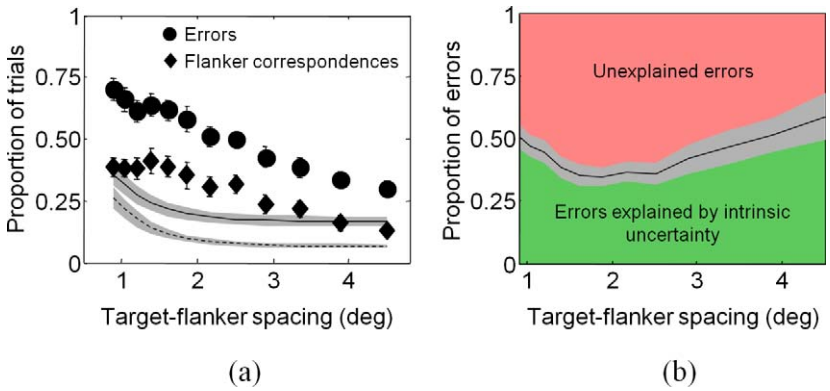

Figure 7. Predicted versus actual error rates in Experiment 3. (a) The ideal-observer model underestimates both the error and flanker correspondence rates (solid curve: predicted error rate; dashed curve: predicted flanker correspondence rate). (b) The ratio between the predicted and actual error rates (black curve). The model explains approximately half of the errors (green area; the mean area size across subjects was 44% ± 5.2%). The gray-shaded areas in both plots represent 1 *SEM* of model predictions.

that flanker proximity affects uncertainty about stimulus position. Indeed, a two-way ANOVA with the factors observer type (model vs. human) and spacing (10 values in the range 0.9 to 4.5) showed no main effect of observer type ($F[1, 120] = 0.81$, $p = 0.37$). However, this model predicts flanker-correspondence rates that are significantly different from those of the human observers ($F[1, 120] = 48.8$, $p < 0.001$). The identification thresholds following from this model (Figure 8, top row, right) are also significantly different from the human observer thresholds ($F[1, 104] = 19.0$, $p < 0.001$). These results argue against the hypothesis that the crowding behavior in Experiment 3 was caused solely by an increase in position uncertainty.

Hypothesis 2: flanker proximity affects identity uncertainty

To test the hypothesis that flanker proximity affects identity uncertainty, we manipulated the probability of an IC through parameter μ_{IC} in Equation 4:

$$\mu_{IC} = \mu_I \left[1 + \gamma(\varphi_{OUT} - \varphi_T)^{\delta}\right],$$

where μ_I is the value of μ_{IC} for isolated stimuli, as estimated from Experiment 2. We let the width of the cumulative Gaussian in Equation 4 scale with its mean by fixing the ratio between μ_{IC} and σ_{IC} to μ_I/σ_I.

There is no significant difference between the error rates in the subject data and those fitted by the model ($F[1, 120] = 0.47$, $p = 0.49$), which suggests that it can accurately account for those rates (Figure 8, center row, left). However, this model predicts flanker-correspondence rates that are significantly different from those in the human data ($F[1, 120] = 52.4$, $p < 0.001$). In addition, the identification thresholds following from this model (Figure 8, center row, right) are significantly different from the human observer thresholds ($F[1, 104] = 15.7$, $p < 0.001$). These results show that it is unlikely that the crowding behavior observed in Experiment 3 resulted solely from an increase in uncertainty about the identities of the stimuli.

Hypothesis 3: flanker proximity affects both position and identity uncertainty

Finally, we tested a model in which flanker proximity affects both position and identity uncertainty. This model produces behavior indistinguishable from the human data (Figure 8, bottom row). An ANOVA such as the ones previously described revealed no main effects of observer type on error rate ($F[1, 120] = 0.008$, $p = 0.93$), contrast thresholds ($F[1, 104] = 0.22$, $p = 0.64$), or predicted flanker-correspondence rates ($F[1, 120] = 1.420$, $p = 0.24$). These results suggest that—if observers were optimal in Experiment 3—then one way to explain crowding effects is by assuming that uncertainty about both stimulus positions and identities depends on flanker proximity.

Model comparison

To compare the likelihoods of the three hypotheses in a more rigorous way, we computed the likelihood of the subject data under each of the models. Because not all the models have the same number of parameters, we used Bayesian model comparison (MacKay, 2003) to compute and compare these likelihoods (see Appendix for details). The results confirmed that the model in which the proximity of flankers affects both position and identity uncertainty describes the data better than the models in which only one type of uncertainty depends on flanker proximity. Averaged over subjects, the log-likelihood difference with the model in which flanker proximity does not affect the level of positional or identity uncertainty is 603 ± 117; the difference with the model in which flanker proximity affects positional uncertainty is 28.8 ± 16.6; and the difference with the model in which it affects identity uncertainty is 27.0 ± 9.3.

Discussion

The main aim of this paper was to examine to what extent ideal observers suffer from crowding effects and to compare this with human observers. By means of simulations, we found that an ideal observer who performs target identification in the presence of flankers and sensory noise will show a critical spacing qualitatively reminiscent of critical spacing in humans. We performed three experiments to quantitatively compare crowding in ideal and human observers. The results showed that about half of the errors in a typical crowding task could be accounted for by the perceptual uncertainty that also exists for isolated targets. The remainder of the errors could be accounted for by a model in which uncertainty about both the position and identity of objects (i.e., the widths of their internal noise distributions) is affected by flanker proximity. Models in which only one of either of these types of uncertainty depended on flanker proximity were insufficient to explain crowding—a finding that is in line with earlier studies (Petrov & Popple, 2007; Sun, Chung, & Tjan, 2010). Our finding that the hallmark of crowding—namely, a critical spacing that scales with viewing eccentricity—is an ideal-observer property corroborates previous proposals that have suggested using ideal-observer theory to explain crowding (Solomon & Dayan, 2011; Sun et al., 2010).

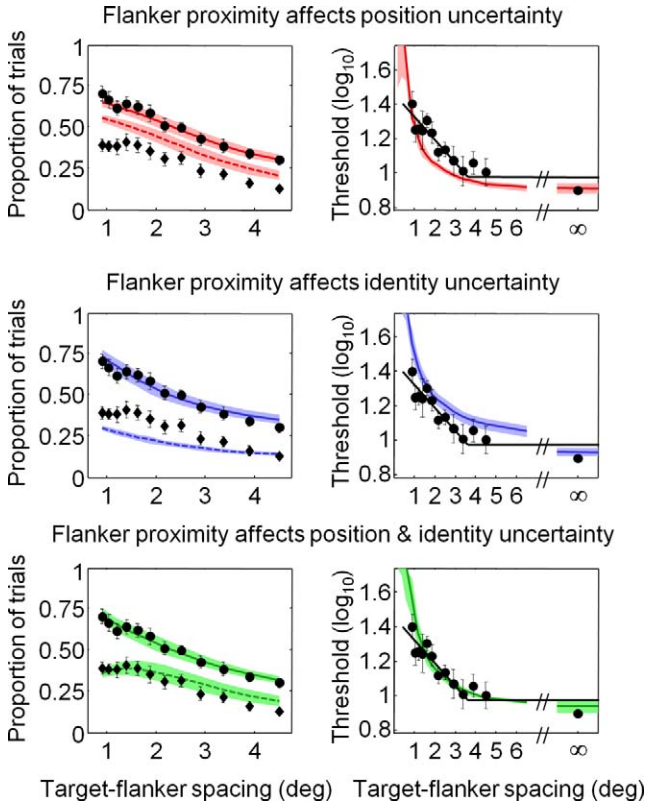

Figure 8. Model fits to the data from Experiment 3. Top row, left panel: The model in which flanker proximity only affects positional uncertainty can accurately fit the human error rates (black markers), but predicts flanker-correspondence rates that are significantly different from the human data. Right panel: Also the identification thresholds are poorly fit by this model. Center row: The same is true for the model in which flanker proximity affects uncertainty only about the identities of the stimuli. Bottom row: The model in which both types of uncertainty depend on flanker proximity accounts well for both the flanker correspondence rates and the threshold elevations. The black solid lines are trilinear functions, which were fitted to obtain an estimate of the critical spacing. Here, the estimated critical spacing is approximately 3.5°. Markers represent human data; curves represent model fits and predictions. Error bars and shaded areas indicate 1 *SEM*.

Furthermore, because the level of crowding in an ideal-observer model is determined by the variability in perceived stimulus locations and identities, ideal-observer theory may also serve as a basis for explaining the recent finding that crowding is tuned for perceived rather than physical stimulus locations (Dakin, Greenwood, Carlson, & Bex, 2011).

Our study improves on previous crowding studies in a number of ways. First, we have shown how a relatively straightforward model based on ideal-observer theory is able to explain the occurrence of a critical spacing—the hallmark of crowding. Second, we have measured the spatial uncertainty of items in isolation and used this to predict its effect in a common crowding paradigm. Whereas quantifying spatial uncertainty about unflanked stimuli seems to be a necessity for any quantitative theory of crowding, as yet, doing so seems to be the exception rather than the rule. Third, we have explored whether and how additional spacing-dependent uncertainty in both spatial and identity certainty can explain the magnitude and properties of human crowding.

Our results suggest that perceptual uncertainty depends on flanker proximity. An important direction for future work is to examine what the neural origin of such a dependence could be. Theoretical studies may help to provide an answer to this question. Recently, it was noted that the fact that receptive fields are spatially extended causes crowding-like effects in ideal observers (Solomon & Dayan, 2011). The intuition behind this is that, due to being spatially extended, receptive fields of single cells will "see" multiple stimuli when stimulus spacing is reduced sufficiently, which in turn causes correlations in population activity. Such correlations are likely to decrease the amount of information in a population of cells, as quantified through, for example, the concept of Fisher information (Ma, 2010). It would be interesting to examine whether this decrease in information depends on stimulus spacing in a way that is compatible with the relationship between perceptual uncertainty and stimulus spacing in our model.

From a computational standpoint, there is no fundamental difference between uncertainty in the spatial domain and uncertainty in the feature domain to the ideal observer. We would therefore predict that a critical "spacing" also exists in the feature domain with properties similar to the critical spacing in the spatial domain. This prediction could be tested using psychophysical methods similar to those used to measure critical spacing in the spatial domain. One could, for example, think of an experiment in which the target is defined by a feature (rather than location; e.g., "the vertical stimulus"), and the task is to report the location of the stimulus, rather than its identity. Our model predicts that performance thresholds (e.g., 75%-correct target contrast) would depend on feature spacing (i.e., similarity) in a way very similar to how these thresholds depend on physical spacing. Whereas effects of feature similarity on crowding have been noted before (e.g., Andriessen & Bouma, 1976; Kooi, Toet, Tripathy, & Levi, 1994; Levi, Hariharan, & Klein, 2002), we do not know of any proposals that explain these in a way that is completely analogous to explanations of spatial effects on crowding.

The present study has several limitations. First, we used complex stimuli (letters) for which it is difficult to define a proper noise distribution. As a result, we were unable to derive the posterior distribution over letter identity, choosing to use an approximation instead. This problem could be addressed by repeating the study

with simpler, one-dimensional stimuli, such as oriented gratings, for which the noise distribution can be modeled as a circular Gaussian (Von Mises) distribution. Second, to fit the data in the crowding experiment, we assumed that uncertainty regarding stimulus positions and identities was related to stimulus spacing by a power law function. Whereas a power law function was able to fit the data without introducing a large number of additional parameters, other choices may have been possible. Third, in our model, we assumed that position and identity confusions are independent events. However, it is conceivable that some interaction takes place, especially when stimulus spacing is very small. Incorporating such interactions is nontrivial but could improve the model. Finally, the present study modeled the crowding phenomenon purely in terms of computations at the behavioral level. Ultimately, in addition to providing an explanation at the computational or behavioral level, a theory of crowding should provide a description of the neural substrate of these computations. We speculate that spacing effects on perceptual uncertainty at the behavioral level may be explained as a result of signal correlations in neural populations, but this hypothesis needs computational and ideally even experimental testing.

Acknowledgments

Ronald van den Berg and Frans W. Cornelissen are supported by the Netherlands Organisation for Scientific Research (NWO). The authors thank Wei Ji Ma for several helpful discussions and Denis Pelli for useful comments on an earlier draft of the manuscript.

Commercial relationships: none.
Corresponding author: Ronald van den Berg.
Email: nronaldvdberg@gmail.com.
Address: Department of Neuroscience, Baylor College of Medicine, Houston, USA.

References

Andriessen, J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research, 16*, 71–78.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177–178.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Chakravarthi, R., & Cavanagh, P. (2007). Temporal properties of the polarity advantage effect in crowding. *Journal of Vision, 7*(2):11, 11–13, http://www.journalofvision.org/content/7/2/11, doi:10.1167/7.2.11. [PubMed] [Article].

Chung, S. T., & Legge, G. E. (2009). Precision of position signals for letters. *Vision Research, 49*(15), 1948–1960.

Dakin, S. C., Cass, J., Greenwood, J. A., & Bex, P. J. (2010). Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *Journal of Vision, 10*(10):14, 1–16, http://www.journalofvision.org/content/10/10/14, doi:10.1167/10.10.14. [PubMed] [Article].

Dakin, S. C., Greenwood, J. A., Carlson, T. A., & Bex, P. J. (2011). Crowding is tuned for perceived (not physical) location. *Journal of Vision, 11*(9):2, 1–13, http://www.journalofvision.org/content/11/9/2, doi:10.1167/11.9.2. [PubMed] [Article].

Faisal, A. A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9*(4), 292–303.

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences of the United States of America, 106*(31), 13130–13135.

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2010). Crowding changes appearance. *Current Biology, 20*(6), 496–501.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334–337.

Hess, R. F., & Holliday, I. E. (1992). The coding of spatial position by the human visual system: Effects of spatial scale and contrast. *Vision Research, 32*(6), 1085–1097.

Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effects. *Visual Cognition, 9*(7), 889–910.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision, 8*(2), 255–279.

Korte, W. (1923). Über die Gestaltauffassung im indirekten Sehen. *Zeitschrift für Psychologie, 93*, 17–82.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654.

Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: Why bigger is better. *Current Biology, 19*(23), 1988–1993.

Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in

peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision, 2*(2):2, 167–177, http://www.journalofvision.org/content/2/2/2, doi:10.1167/2.2.2. [PubMed] [Article].

Ma, W. J. (2010). Signal detection theory, uncertainty, and Poisson-like population codes. *Vision Research, 50*(22), 2308–2319.

MacKay, D. J. (2003). *Information theory, inference, and learning algorithms*. Cambridge, UK: Cambridge University Press.

Michel, M., & Geisler, W. S. (2011). Intrinsic position uncertainty explains detection and localization performance in peripheral vision. *Journal of Vision, 11*(1):18, 1–18, http://www.journalofvision.org/content/11/1/18, doi:10.1167/11.1.18. [PubMed] [Article].

Nandy, A. S., & Tjan, B. S. (2007). The nature of letter crowding as revealed by first- and second-order classification images. *Journal of Vision, 7*(2):5, 1–26, http://www.journalofvision.org/content/7/2/5, doi:10.1167/7.2.5. [PubMed] [Article].

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12):12, 1136–1169, http://www.journalofvision.org/content/4/12/12, doi:10.1167/4.12.12. [PubMed] [Article].

Pelli, D. G., & Robson, J. G. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences, 2*, 187–199.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*(10), 1129–1135.

Petrov, Y., & Popple, A. V. (2007). Crowding is directed to the fovea and preserves only feature contrast. *Journal of Vision, 7*(2):8, 1–9, http://www.journalofvision.org/content/7/2/8, doi:10.1167/7.2.8. [PubMed] [Article].

Solomon, J. A., & Dayan, P. (2011). Selective Bayes: Attentional load and crowding. *Vision Research, 50*, 2248–2260.

Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11):8, 1024–1037, http://www.journalofvision.org/content/5/11/8, doi:10.1167/5.11.8. [PubMed] [Article].

Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty. Its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology, 53*, 471–477.

Sun, G. J., Chung, S. T., & Tjan, B. S. (2010). Ideal observer analysis of crowding and the reduction of crowding through learning. *Journal of Vision, 10*(5):16, 1–14, http://www.journalofvision.org/content/10/5/16, doi:10.1167/10.5.16. [PubMed] [Article].

van den Berg, R., Roerdink, J. B., & Cornelissen, F. W. (2007). On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. *Journal of Vision, 7*(2):14, 1–11, http://www.journalofvision.org/content/7/2/14, doi:10.1167/7.2.14. [PubMed] [Article].

van den Berg, R., Roerdink, J. B., & Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology, 6*(1), e1000646.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160–168.

Appendix

Maximum likelihood estimation in Experiment 1

The probability that the stimulus on the left is perceived as being closer to fixation than the stimulus on the right equals

$$p(x_L < x_R|\varphi_L, \varphi_R, \sigma_L, \sigma_R) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\varphi_R - \varphi_L}{\sqrt{2(\sigma_L^2 + \sigma_R^2)}}\right),$$

where φ_L and φ_R denote the eccentricities of the stimuli on the left and the right, respectively; x_L and x_R denote the internal representations of the left and right stimuli; and σ_L and σ_R denote the trial-to-trial variability over these internal representations; erf(.) is the error function; and $\sigma(\varphi, c) = a\varphi c^t$. Under this model, and assuming independence between trials, the likelihood of the data obtained in Experiment 1 equals

$$p(\text{data}|a, t) = \prod_i p(\text{data on trial } i|a, t) = \prod_i p(\varphi_{L,i}, \varphi_{R,i}, c_{L,i}, c_{R,i}, r_i|a, t)$$

$$= \prod_i p(r_i|\varphi_{L,i}, \varphi_{R,i}, c_{L,i}, c_{R,i}, a, t) \times p(\varphi_{L,i}, \varphi_{R,i}, c_{L,i}, c_{R,i}|a, t)$$

$$\propto \prod_i p(r_i|\varphi_{L,i}, \varphi_{R,i}, c_{L,i}, c_{R,i}, a, t)$$

$$= \prod_{i=1}^{M} p(r_i = \text{“left”}|\varphi_{L,i}, \varphi_{R,i}, c_{L,i}, c_{R,i}, a, t) \times \prod_{i=M+1}^{N} p(r_i = \text{“right”}|\varphi_{L,i}, \varphi_{R,i}, c_{L,i}, c_{R,i}, a, t)$$

$$= \prod_{i=1}^{M}\left[\frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{\varphi_{R,i} - \varphi_{L,i}}{a\sqrt{2(\varphi_{L,i}^2 c_{L,i}^{2t} + \varphi_{R,i}^2 c_{R,i}^{2t})}}\right)\right] \times \prod_{i=M+1}^{N}\left[\frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{\varphi_{R,i} - \varphi_{L,i}}{a\sqrt{2(\varphi_{L,i}^2 c_{L,i}^{2t} + \varphi_{R,i}^2 c_{R,i}^{2t})}}\right)\right],$$

where we sorted the data in such a way that the first M trials were those on which the subject responded “left” and the subsequent N–M trials were those on which the subject responded “right.” We find the maximum-likelihood values of parameters a and t by maximizing the log of the previous expression:

$$\arg\max_{a,t} \sum_{i=1}^{M} \log\left[\frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{\varphi_{R,i} - \varphi_{L,i}}{a\sqrt{2(\varphi_{L,i}^2 c_{L,i}^{2t} + \varphi_{R,i}^2 c_{R,i}^{2t})}}\right)\right] + \sum_{j=M+1}^{N} \log\left[\frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{\varphi_{R,i} - \varphi_{L,i}}{a\sqrt{2(\varphi_{L,i}^2 c_{L,i}^{2t} + \varphi_{R,i}^2 c_{R,i}^{2t})}}\right)\right].$$

Bayesian model comparison

We applied Bayesian model comparison in Experiment 1 (to choose which of several models best relates subjects’ position uncertainty to stimulus contrast and eccentricity) and in Experiment 3 (to choose which of three models best describes the subjects’ crowding behavior).

Bayesian model comparison is a principled way to decide which of a set of models provides the most likely description of a given set of data. Using this method, one computes for each model the *average* likelihood of the data over its parameter space between, instead of merely the *maximum* likelihood. As a result of the averaging, additional model parameters are automatically punished for. The intuition behind it is that, while the maximum likelihood can never get worse when adding free parameters to a model, the average likelihood will increase only if the additional parameter adds sufficient “likelihood mass” to the entire volume, because adding a parameter will lead to averaging over a larger volume.

Because the mathematics are very similar in both cases, here we explain it only for Experiment 3. For each model M, we calculated the probability of the data D given that model. Denoting model parameters by θ, this probability can be written as

$$p(D|M) = \int p(D|M, \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (6)$$

The conditional probability $p(D|M, \boldsymbol{\theta})$ is calculated by assuming that the data are conditionally independent across trials:

$$p(D|M, \boldsymbol{\theta}) = \prod_{i=1}^{N} p(D_i|M, \boldsymbol{\theta}) = \prod_{i=1}^{M} p(\text{error}|\boldsymbol{\varphi}, \mathbf{c}, \boldsymbol{\theta}) \times \prod_{i=M+1}^{N}\Big(1 - p(\text{error}|\boldsymbol{\varphi}, \mathbf{c}, \boldsymbol{\theta})\Big); \quad (7)$$

with $p(\text{error}|\boldsymbol{\varphi}, \mathbf{c}, \boldsymbol{\theta})$ as in Equation 5, and trials sorted such that trials $1 \ldots M$ are the trials on which the observer made an error and $M + 1 \ldots N$ the trials on which the observer reported correctly. (In the analysis of Experiment 1, we used $p(\text{data}|a,s)$ as defined in the previous section instead of $p(\text{error}|\boldsymbol{\varphi}, \mathbf{c}, \boldsymbol{\theta})$.) It is convenient to take the logarithm of Equation 7:

$$L_M(\boldsymbol{\theta}) = \log p(D|M, \boldsymbol{\theta}) = \sum_{i=1}^{M} \log p(\text{error}|\boldsymbol{\varphi}, \mathbf{c}, \boldsymbol{\theta}) + \sum_{i=M+1}^{N} \log\Big(1 - p(\text{error}|\boldsymbol{\varphi}, \mathbf{c}, \vec{\boldsymbol{\theta}})\Big).$$

We assume a uniform prior distribution over the parameters in Equation 6:

$$p(\boldsymbol{\theta}) = \frac{1}{\text{Vol}_{\boldsymbol{\theta}}}$$

where $\text{Vol}_{\boldsymbol{\theta}}$ is the volume of parameter space. Combining and taking the log, we find the following expression for the log likelihood of model M under data set D:

$$\log p(D|M) = \log\frac{1}{\text{Vol}_{\boldsymbol{\theta}}} \int e^{L_M(\boldsymbol{\theta})} d\boldsymbol{\theta}.$$

We approximated this integral numerically, by evaluating $L_M(\boldsymbol{\theta})$ for a large range of (linearly spaced) parameter values:

$$\log p(D|M) = \log\frac{1}{\text{Vol}_{\boldsymbol{\theta}}} + \log\int e^{L_M(\boldsymbol{\theta})} d\boldsymbol{\theta} = -k \log m + \log\sum e^{L_M(\boldsymbol{\theta})},$$

where k is the number of free parameters of the model and m the number of values at which each parameter is evaluated.

Model fitting in Experiment 3

The model fits shown in Figure 8 were obtained by maximum likelihood estimation, i.e., by maximizing the summed log likelihood function over all trials. Specifically, we computed

$$\arg\max_{\vec{\theta}} \sum_{i=1}^{M} \log p(\text{error}|\vec{\varphi}_i, \vec{\mathbf{c}}_i, \vec{\theta}) + \sum_{j=M+1}^{N} \log\Big(1 - p(\text{error}|\vec{\varphi}_i, \vec{\mathbf{c}}_i, \vec{\theta})\Big),$$

where trials $1 \ldots M$ are the trials on which the observer made an error and trials $M+1 \ldots N$ the trials on which the observer made a correct response, $\boldsymbol{\theta}$ a vector with the model's free parameters.