CROWDING AND BINDING

Crowding and binding: Not all feature-dimensions behave in the same way

Amit Yashar¹⁻⁴, Xiuyun Wu¹, Jiageng Chen¹ & Marisa Carrasco^{1,2}

¹Department of Psychology, New York University, New York, NY, USA

²Center for Neural Science, New York University, New York, NY, USA

³Department of Special Education, University of Haifa, Haifa, Israel

⁴The Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israel

Corresponding Author:

Amit Yashar, Faculty of Education, University of Haifa, 199 Abba Khoushy Ave, Haifa, 3498838, Israel, Email: amit.yashar@edu.haifa.ac.il, Phone: +972 (4) 8249297

Keywords:

Crowding; Feature binding; Spatial integration; Color; Spatial frequency; Orientation; Estimation

Running title: Crowding and binding

CROWDING AND BINDING

Abstract

Humans often fail to identify a target because of nearby flankers. The nature and stage(s) at which this “crowding” occurs are unclear, and whether crowding operates via a common mechanism across visual dimensions is unknown. Using a dual estimation report, we quantitatively assessed the processing of each feature alone and in conjunction with another feature both within and between dimensions. Crowding emerged due to confusion between orientations or colors of target and flankers, but averaging of their spatial frequencies (SFs). Furthermore, crowding of orientation and color were independent, but crowding of orientation and SF were interdependent. This qualitative difference of crowding errors across dimensions revealed a tight link between crowding and ‘illusory conjunctions’ (mis-binding of feature dimensions). These results and a computational model suggest that crowding and illusory conjunction in the visual periphery are due to pooling across a joint coding of orientation and spatial frequencies but not of color.

CROWDING AND BINDING

Introduction

Recognition of peripheral objects is fundamentally limited by their spacing, not just by the visibility of their features. A target that can be easily identified when presented alone becomes unrecognizable when presented along nearby flankers, i.e. ‘crowded’[1-3]. This breakdown in object-recognition[4] corresponds to the increase in positional uncertainty in the periphery, due to larger receptive-fields[5-7]. Crowding affects various perceptual tasks and hinders the identification of basic stimuli such as letters[2] and faces[8] (see review[9]).

The neural mechanisms underlying crowding are still unknown (reviews by[3,9]). Recent studies have posited that crowding occurs either at an early visual stage, such as V1 or V2 [7,10], or at multiple stages of visual processing and object representation [11-13]. Proposed models explain crowding as reflecting either substitution of objects [14-18] or pooling of features [7,19-23]. Substitution models predict confusion errors, i.e. miss-reporting flankers items instead of the target [14,18]. Pooling models typically predict features averaging errors, i.e. reporting a combination of target and flanker features [23], but recent pooling models attempt to explain both averaging and confusion errors [7,19,21].

Crowding models implicitly assume that the crowding mechanism is in principle the same across various types of stimulus (reviewed by[3,9]); and therefore each proposed model presumably predicts the same type of crowding errors regardless of the types of feature-dimensions, such as orientation, color and spatial frequency (SF). However, whether crowding errors are qualitatively the same across different feature-dimensions and their conjunctions, i.e. feature-binding, is unknown. Most studies have investigated crowding errors within a particular dimension (e.g., orientation [13,14,19,22-25]), and the few investigations of errors within various dimensions could not distinguish averaging and substitution errors, nor could they test for qualitative differences across dimensions [26-28]. Thus, it is still unknown how each of the basic feature-dimensions behaves under crowding and whether they behave interdependently, or independently from each other.

This study had two main goals. First, to test the assumption that the same crowding mechanism applies to different features [9,26]. We investigated the contribution of each flanker to crowding for orientation, SF, and color. Second, we investigated the processing stage at which crowding occurs. Specifically, we asked whether crowding occurs before or after features are bound into an object. To do so, we employed a feature estimation technique that enabled us to simultaneously characterize the pattern of crowding errors within and between feature-dimensions. Thus, we were able to quantitatively assess not only the accuracy of feature perception under crowding conditions, but also the accuracy of feature binding.

Results

Experiment 1: Orientation and Color

Observers performed an orientation and color estimation task of a peripheral (7° of eccentricity) colored T- shaped target (**Figure 1A**; see Methods). The target could be either alone –*target-*

CROWDING AND BINDING

alone condition— or flanked by two similar T-shaped items —*flanker conditions*. Target and flankers were radially aligned on the horizontal meridian axis, either on the left or right hemifield. The center to center distance between the target and the flankers was 2.1°, within the region of

Figure 1. Orientation and color estimation tasks, stimuli, conditions and results in Experiment 1. **A.** Trial sequence. The order of reports (orientation and color) alternated across blocks and was counterbalanced for each observer. The response display In the Supplementary Experiment requires a simultaneous report of color and orientation. Observers had to report both color and orientation by pointing and clicking on the T-shape with the same color and orientation as the target. **B.** Distribution of errors relative to target feature values in the target alone condition. **C.** Distribution of errors in flanker conditions. Error distributions are plotted as a function of the deviation between the estimation report and the target's feature value for orientation (top), color (right), and their conjunction (heat map).

CROWDING AND BINDING

The x-axis is in units of standard deviation (SD) of the error distribution in target-alone trials. Data from asymmetrical flankers were aligned to 60/-90 (as pictured) during analysis. Solid lines indicate the response probabilities predicted by the standard misreport model.

radial crowding [29,30]. The orientation and color of the target were each independently selected at random from two circular parameters spaces. Target orientation was randomly selected out of 180 values evenly distributed between 1°-360°. Target color was randomly selected out of 180 values evenly distributed along a circle in the DKL color space [31,32]. Flankers' orientation and color were selected out of the same circular parameter spaces as the target but differed in values from those of the target. Each flanker had a unique relation to the target in both parameters spaces (negative or positive). This design enabled us to track the direction and distance of estimation errors in both feature-dimensions relative to each flanker. We analyzed the error distributions (the estimated value minus the true target value) of each feature-dimension by fitting probabilistic models (**see Methods**).

Misreporting of orientations or colors

For each feature dimension, the estimation error was defined as the deviation between the reported and the actual feature value of the target (**Figure 1B** and **1C**). We analyzed the error distributions by fitting five probabilistic mixture models to the individual data (**see Methods**). All models included a von Mises (circular) distribution to describe the probability density of precision errors for the target's feature, and a uniform component to capture the guessing in estimation. We compared between these basic mixture models (models S and B), models that also include a misreports component to describes the probability of reporting one of the flankers to be the target (models SM and BM), and a model that also includes an 'educated guess' component to describe the proportion of guessing the target based on flankers values (model EG). We compared models by calculating Akaike Information Criterion with correction (AICc) values for the individual model fits (**Figure 2**). **Table S1 (Supplementary results)** shows fitted model parameters in each condition and dimension.

To compare crowding for orientation and color, we converted each feature dimension values to units of standard deviations (SD) of the error distribution in target-alone trials. So that in each feature dimension target-flanker distance is presented in relation to observer's precision in the target-alone trials. These standardized units of target-flanker distance (± 4.60 and ± 6.90 SD for orientation and ± 3.43 and ± 5.15 SD for color) confirmed that the effect of flanker interference is comparable across feature-dimensions (**Figure 1B** and **1C**).

On target-alone trials, the error distributions for both feature dimensions were well-described by a von Mises distribution centered on the target value with an added non-zero uniform distribution (γ) for both orientation and color (**Figure 1B**), $t(13) > 2.98$, $p < .005$, indicating that a small yet significant proportion of the responses were statistically unrelated to the target (i.e. guessing).

For both orientation and color, models with misreports components outperformed the models without the misreport component. That is, a significant proportion of errors were centered on the value of each of the two flankers, impairing the fit of a single von Mises distribution to the data. Importantly, adding an 'educated guess' component to the misreporting model did not improve the fitting, indicating that observers were unaffected by the partial correlation between target and flankers values. Guessing rate was higher in the flanker conditions, $t(13) > 2.19$, $p < .05$

CROWDING AND BINDING

(**Table S1**). The misreporting rate (β) was larger than zero in all flanker conditions, $t(13)>5.22$, $p<.001$. The variability (α) of errors centered on the target increased significantly relative to the target-alone condition, $t(13)>3.32$, $p<.01$, indicating that crowding led to a reduced precision, and increased the guessing rate and misreporting errors.

Next we compared misreporting rates between orientation and color. Orientations were misreported more than color: orientation averaged $\beta = 0.27 \pm 0.03$ vs. color averaged $\beta = 0.086 \pm 0.01$, $t(13)=7.64$, $p<0.001$, indicating a larger crowding effect for estimation of orientation than for estimation of color.

Figure 2. Model comparison. Comparison of individual goodness of fit of the five models to the flanker-conditions data in the orientation and color reports in Experiment 1 and in the orientation and SF reports in Experiment 2. Mean AICc was calculated by subtracting the mean AICc of the best performing model (lowest AICc) from the individual AICc scores of each mixture model. S=standard, B=bias, SM=standard with misreport, BM=bias with misreport, EG = educated guesses. Error bars are 1 SEM.

Observers misreport orientations independently from color

To assess whether orientation and color errors occur before or after orientation and color are bound, we used trial-by-trial correlation between orientation errors and color errors to test the interdependency of crowding errors across feature-dimensions. The joint distributions in each condition are presented in **Figures 1B and 1C** (bottom panels). In the target-alone condition, only 4 out of 14 observers showed a significant Pearson correlation between orientation and color errors, overall mean $r=0.06$, $SE=0.04$. In the flanker conditions (all three conditions collapsed), only 3 observers showed a significant correlation, overall mean $r=0.03$, $SE=0.02$. These findings show that orientation and color estimation were predominantly uncorrelated.

In both feature-dimensions the nature of errors was largely the same: observers reported the orientation or color of a flanker instead of that of the target (misreporting errors). Note that it is unlikely that orientation errors were due to combining target and flanker T-shape parts (e.g. combining target 'steam' with flanker 'hat'), as such combinations would lead to reporting a vast range of orientations which would be reflected by an increase in the uniform distribution rather than by misreporting errors. The misreporting errors in color could not be explained by optical blur, as such blur predicts averaging errors. Importantly, orientation and color misreporting were

CROWDING AND BINDING

independent from each other, suggesting that orientation and color are unbound under crowding conditions.

An alternative explanation is that the separate reports for color and orientation encouraged observers to separately encode color and orientation. Hence, the uncorrelated errors across dimensions could have been due to response strategy rather than unbounded perceptual representation of features. To rule out this alternative explanation, we conducted a control experiment in which observers reported both orientation and color of the target with a single response (**Figure 1A**). This experiment yielded converging results (**Supplementary Experiment and Figure S1 in Supplementary information**): uncorrelated trial-by-trials errors were maintained even when observers simultaneously reported both dimensions. Taken together, these results show that orientation and color crowding errors occur before features are bound into an object.

Experiment 2: Orientation and Spatial Frequency

In this experiment we tested whether the pattern of results obtained in Experiment 1 would also emerge with different stimuli and feature dimensions. Observers viewed sinusoidal gratings (Gabors) and estimated the target's spatial frequency and orientation (**Figure 3A**). These two dimensions are jointly coded by individual neurons in V1 [33-35], thus we hypothesized instead of independent, crowding errors may be interdependent.

Misreporting of orientations but averaging of spatial frequencies

Error distributions differed between orientation and SF. Whereas orientation errors were best-described by misreporting models, SF errors were best-described with a single Gaussian function with an added uniform distribution (bias models) (**Figure 2**). In both feature-dimensions the best-fitted models outperformed the 'educated guess' model. **Table S1** shows best-fitting model parameters for each condition and dimension.

Orientation: As in Experiment 1, significant proportions of orientation error in the flanker conditions were due to misreporting, $t(13) > 3.6$, $p < .004$, and guessing, $t(13) > 3.6$, $p < .003$ (**Figure 3B**). But, the variability of the von Mises distribution between flanker and target-alone conditions were equivalent, $t < 2.1$, $p > .06$. These results indicate that orientation estimation errors were due to increases in guessing and misreporting rate.

SF: The variance of SF errors was larger in the flanker conditions than the target-alone condition (**Table S1**), $t(13) > 3.9$, $p < .002$. The proportion of guesses did not increase compared to the target-alone condition, $t(13) < 2.1$, $p > .05$. The mean (μ) of the Gaussian distribution in the 40/-70 condition was significantly biased, $t(13) = 4.69$, $p < .001$, toward the -70 flanker (negative bias) compared to the target-alone condition within the same target SF range (**Table S1**). No effects of bias were found in ± 40 and ± 70 conditions compared to the target-alone condition, $t(13) < 0.8$, $p > .4$.

CROWDING AND BINDING

Figure 3. Stimuli and results of Experiment 2, simulated data, and analysis of feature-dimension interdependency in Experiment 1, 2 and simulated data. A. Orientation and SF estimation tasks and stimuli in Experiment 2. Illustration of a trial sequence. The order of feature report – SF and orientation – alternated across blocks and was counterbalanced for each observer. **B.** and **C.** Distributions of errors relative to target feature values in Experiment 2 (**B**) and the simulated data for the joint coding model (**C**). For each condition, error distributions are plotted as a function of the deviation between estimation report and target feature values for orientation (top), SF (right), and their conjunction (heat map). Solid lines are model fits for orientation (standard misreport model)

CROWDING AND BINDING

and SF (bias model). The model simulation produced the same distinctive pattern of crowding errors for orientation and SF: misreporting errors for orientation, and increases in variability and mean bias in 40/-70 flanker conditions for SF. **D.** and **E.** Comparisons of misreport rates (top panels) and SD (bottom panels) between consistent and inconsistent trials in Experiments 1 and Experiment 2 (**D**), and in the simulated data (**E**) of the Joint coding model and independent distribution model. The Joint coding simulated model showed the same pattern of consistent effect as Experiment 2. Error bars are 1 SEM [36]. Asterisks indicate $p < .01$

Observers misreport orientations interdependently of averaging of spatial frequencies

Heat maps of the joint distributions in each condition are presented in **Figure 3B**. Orientation and SF estimation errors were uncorrelated in the target-alone condition, only one observer showed a significant linear correlation; mean $r = 0.002$, $SE = 0.02$. However, across flanker conditions there was a significant linear correlation between orientation and SF errors, $r = 0.138$, $SE = 0.03$, $ps < .0001$, indicating that orientation and SF crowding errors were interdependent. Individual trial-by-trial linear correlations showed significant ($ps < .04$) correlation in 10 out of 14 observers ($Mean\ r = 0.14$, $SE = 0.03$).

To further assess the interdependency of crowding errors across features-dimensions and to compare the interdependency in Experiments 1 and 2, we analyzed the distribution of errors in one dimension based on observers' errors in the other dimension. That is, we tested whether the direction (with respect to flanker values) of an error in one feature-dimension was dependent on the direction of the error in the other feature-dimension (see Method). To this aim, trials were divided into two groups: trials in which estimation errors for both features (orientation vs. color or SF) were toward the same flanker (*consistent trials*), and trials in which errors for each feature were towards separate flankers (*inconsistent trials*). **Figure 3D** plots the effects of consistent vs. inconsistent errors across dimension on misreporting rates and estimation variability for color (Experiment 1), orientation (Experiment 1 and 2), and SF (Experiment 2). The effect of consistency was only found between orientation and SF but not between orientation and color (see **Supplementary Results** for detailed results).

Simulation of pooling of a joint population coding can explain orientation and spatial frequency crowding

To test whether the results of Experiment 2 can be explained by a representation in which orientations and SFs are bound, we fitted a biologically plausible computational model that simulates a population of neurons. In the *joint coding* model the populations jointly code orientations and SFs (**Figure 4**, and **Supplementary Method and Table S2**), whereas in the *independent coding* model the populations separately code orientations and SFs (**Table S3**). The models relies on population coding principles and explains crowding as a weighted sum of target and flanker feature values within a receptive field [19,22]. We used a single bivariate Gaussian distribution to simulate the joint coding of orientation and SF and two univariate normal distributions to simulate the separate representation of orientation and color. To explain averaging in SF reports and in orientation reports, the model assumes pooling over a

CROWDING AND BINDING

larger region of space for SF than orientation. This assumption is related to the fact that SF judgments

Figure 4. Simulation of population neural activity that jointly codes for orientation and SF fitted to Experiment 2. The model simulates the firing rates of three populations of neurons with receptive field (RF) locations, orientation preference, and SF preference. Stimuli (Bottom row): A target presented with two flankers (asymmetrical flanker condition 40/-70). Spatial weights (second row from bottom): Normalized Gaussian function determines the population level to a stimulus as a function of its relative distance (compared to other stimuli) from the center of RF. Population (third row from bottom): A bivariate probability function with orientation preference (horizontal) and SF preference (vertical). Orientation and SF arrangement is centered over the target orientation and SF values. Reports (top panel): Report of orientation is based on a single population over the target location. Report of SF is

CROWDING AND BINDING

based on pooling over different locations, i.e. pooling over three population with RF centered over each location).

involve assessing the width and distance of multiple cycles inside the Gabor [37-40], whereas the orientation judgment may involve assessing a more narrow region of space, e.g. a region with higher aspect ratio [41]. In the simulation, coding precisions (the inverse of the SD of the Gaussian response function) and the SD of the Gaussian receptive field, were directly responsible for orientation estimation variability and misreporting rate, respectively. Therefore, these parameters were determined based on the fitting of the probabilistic standard misreport model to the observed orientation data. Both simulated models were fitted to the results of Experiment 2 (both $r^2 = 0.88$). However, unlike Experiment 2 the independent coding model showed no correlation between orientation and SF (mean $r = 0$), and no effect of consistency (**Figure 3E**). The joint coding model, on the other hand, showed the same interdependent pattern as the results of Experiment 2 (**Figure 3C**), including the correlation between orientation and SF (mean $r = 0.23$), and the increase in misreporting rate of orientations and SD of SF in consistent than inconsistent trials (**Figure 3E**). These results support the conclusion that joint coding of orientation and SF underlies the results of Experiment 2.

Discussion

In this study we simultaneously characterized the pattern of crowding errors within and between feature-dimensions. In three experiments we demonstrate variations in the pattern of crowding errors based on the specific feature dimension (orientation, color & SF) and their conjunctions. Crowding is more pronounced for orientation than for color. Crowding reflects misreporting a flanker orientation or color instead of those of the target, but averaging of their SFs. The pattern of results was contingent on the feature dimension type but not the stimulus type: observers misreported the target orientation regardless of whether the stimulus was a colored T or a grating. The distinct pattern of crowding errors in each dimension suggests a distinct representation for each feature dimension.

Crowding errors for orientation and SF were interdependent, but those for orientation and color were independent. These findings were shown by the analysis of the joint distribution of feature-dimension errors and trial-by-trial correlations. Moreover, comparison of model parameters revealed higher orientation misreporting rates when SF and orientation errors were toward the same flanker, but that was not the case for orientation and color. These findings suggest that the spatial integration that underlies crowding operates after orientation is bound with SF but before it is bound with color.

Not all features behave the same under crowding: errors within feature-dimensions

The present study challenges many models' implicit assumption that crowding operates in the same manner across different feature-dimensions (reviewed by [3,9]). Investigations supporting this view had not tested for qualitative differences in the pattern of errors across dimensions [26-28]. Here, within the same display we show both quantitative and qualitative differences in the pattern of errors across different dimensions. First, orientation misreporting is three times more likely than color misreporting, demonstrating that estimation of orientation is more susceptible to

CROWDING AND BINDING

crowding than estimation of color. Second, whereas orientation or color was misreported, SFs were averaged in the SF dimension. This variation between misreporting and averaging occurred within the same stimuli and display (Gabor). It has been proposed that observed errors may be contingent on the target-flankers orientation distance [19,42] (but see [14]). Here, variation between misreporting errors and averaging errors cannot be explained by target-flanker distance. Observers misreported orientation and averaged SF even when the distance in feature-space was effectively the same (**Figure S2**). Thus, this study shows that crowding varies as a function of the feature-dimension being reported.

Crowding and feature-binding: errors between feature-dimensions

Numerous studies demonstrating observers' failure to correctly report the conjunction of feature-dimensions of peripheral items (e.g. color and orientation), attributed this phenomenon known as 'illusory conjunction' to limited attention (for review see [43]). Alternatively, it has been proposed that 'illusory conjunctions' could be due to crowding [1,4,21,27] or to averaging of features [44]. However, studies that investigated crowding and binding have yielded inconsistent results [27,28]. Crowding has been reported to occur either before [27] or after [28] features are bound.

In this study we quantitatively assessed, for the first time, crowding errors both within and between feature-dimensions. The results reveal that under a crowded display color and orientation remain unbound, even when both dimensions are jointly reported. Unlike color, SF remains bound with orientation under crowding displays; observers tended to misreport flanker orientation and average SF with the same flanker. This contingency of binding errors on the specific feature-dimensions may explain why some studies using a set of stimuli suggested that crowding reflects interference in feature-binding [4,27], whereas another study using a different set of stimuli, suggested that crowding follow feature-binding [28].

Conclusions

This study directly links two mostly-independent topics of research –crowding and feature binding– and challenges conventional views in each of them. By testing crowding both within and between feature-dimensions we show that it is not a uniform phenomenon: it reflects different operations depending on the specific feature-dimensions and their conjunctions. Both the data analysis and our model simulation suggest that crowding reflects spatial integration over neural populations that encode both orientation and SF together but color separately.

Acknowledgments

This study was supported by The Israel Science Foundation (grant No.111/15) to AY, and NIH-Grant R01 EY016200 to MC. We would like to thank Denis Pelli and members of the Carrasco lab for their useful comments.

CROWDING AND BINDING

CROWDING AND BINDING

Methods

Experiment 1: Orientation and color

Observers. Fourteen undergraduate and graduate students from New York University participated in this experiment (five females, age range 18 to 29 years old, 21.00 ± 3.28). All observers were naïve to the purposes of the study, and all reported having normal or corrected-to-normal visual acuity and normal color vision. Written informed consent was obtained from all participants before the experiment. The University Committee on Activities Involving Human Subjects at New York University approved experimental procedures.

Apparatus. Stimuli were programmed in Matlab (The MathWorks, Inc., Natick, MA) with the Psychophysics Toolbox extensions [45] and presented on a gamma-corrected 21-in CRT monitor (Sony GDM-5402, with 1280×960 resolution and 85-Hz refresh rate) connected with an iMac. A chin-rest set the 57-cm viewing distance. Colors and luminance were calibrated using SpectraScan Spectroradiometer PR-670 (Photo Research) spectrometer. Eye movements were monitored and recorded by an Eyelink 1000 (SR Research) infrared eye tracker. Observers used the mouse to generate responses.

Stimuli and procedure. Figure 1A illustrates a trial sequence. Each trial began with the presentation of a centered black plus (+) sign (fixation mark) subtending 0.5° , along with two dots subtending 0.05° in radius. The dots were presented on the horizontal meridian one in the left and the other in the right hemifields, each centered at an eccentricity of 7° and indicated the two possible target locations. Following observer fixation (for a random duration between 300-800 ms) the stimulus display appeared for 75 ms. In the stimulus display the target was a T-shaped item, subtending $1.6^\circ \times 1.6^\circ$ and drawn with a 0.3° stroke, presented on the horizontal meridian in either the left or right (randomly selected) hemifields, centered at 7° eccentricity. The orientation and color of the target were each independently selected at random from two circular parameters spaces. Target orientation was randomly selected out of 180 values evenly distributed between 1° - 360° . Target color was randomly selected out of 180 values evenly distributed along a circle in the DKL color space [31,32] (See Supplementary method). Stimuli color and background were equiluminant (56 cd/m^2).

The target could be either alone – *target-alone condition* – or flanked by two similar T-shaped items – *flanker conditions* – centered on the horizontal meridian to the left and to the right of the target (each 2.1° of center to center distance from the target).

The stimulus display was followed by a blank interval of 500 ms, which was then followed by the response displays that appeared until the observer completed both responses. The response displays included an “orientation circle” (a black circle 0.08° thick with an inner radius of 3.8° around the center of the screen) along with a color wheel (1.5° thick and had an inner radius of 2.25°) containing the 180 colors. Observers were asked to estimate the target orientation by pointing and clicking the mouse cursor at a position on the orientation wheel, and to estimate the target color by pointing and clicking the mouse cursor at a position on the color wheel. During report a visual feedback of the selected feature was presented at the location of the target. A letter at fixation (either an O or a C, which stands for orientation first, and color first, respectively) indicated report order, which was counterbalanced across blocks.

CROWDING AND BINDING

Phase 1

To test the appropriateness of the estimation task with the equiluminant DKL colors, we tested the target-alone condition. **Figure S3** illustrates the color estimation values as a function of target color values for the selected eccentricity. The results showed a linear relation between estimation and target color values, ruling out the possibility that the task was mediated by color categories (e.g., red, blue).

Phase 2

Design. There were four conditions: three flanker conditions (**Figure 1C** [Top panels]; ± 60 , 60/-90 and ± 90) and the target-alone condition (**Figure 1B**). Flankers' orientation and color differed from those of the target by either $\pm 60^\circ$ or $\pm 90^\circ$ in orientation space and DKL color space. Each flanker had the same absolute target-flanker difference in both feature-dimensions. Within each feature-dimension, one flanker had a positive and the other negative target-flanker difference (randomly selected), such that each flanker had a unique relation to the target (negative or positive) in both feature-dimensions. This design enabled us to track the direction of estimation errors related to each flanker in both feature-dimensions. The three flanker conditions were based on the combinations of the target-flanker differences with the two flankers: (1) 60 and -60; (2) 90 and -90; (3) 60 and -90, or 90 and -60, which were labeled ± 60 , ± 90 and 60/-90 respectively. Each condition had 200 trials (800 trials overall). Each observer completed 10 blocks of 80 trials over two 50-minute sessions (five blocks per session). In each block there were 20 trials from each of the four conditions. Response display order was counterbalanced across blocks. The experiment began with an 80-trial practice block. Observers were encouraged to take a short rest between blocks.

Models and analyses. We analyzed the error distributions by fitting probabilistic mixture models, which were developed from the standard model [46] and the standard with misreporting model [47]: For each trial we calculated the estimation error for orientation and color by subtracting the estimation value from the true value of the target. In the flanker conditions we aligned the data such that across feature-dimensions one flanker was consistently positive, and the other was consistently negative. In the 60/-90 flanker condition we aligned all data to be 60 and -90. These alignments enabled us to track the effect of each individual flanker on the error distribution of each feature-dimension separately and in conjunction with the other feature dimension. We compared among five models:

- A. *The standard mixture model (Equation 1)* uses a von Mises (circular) distribution to describe the probability density of the pooling estimation of the target's feature, and a uniform component to reflect the guessing in estimation. The model has two free parameters:

$$p(\theta) = (1 - \gamma)f(\theta)_\sigma + \gamma\left(\frac{1}{n}\right) \quad (1)$$

where θ is the value of the estimation error, γ is the proportion of trials in which participants are randomly guessing (guessing rate), and $f(\theta)_\sigma$ is the von Mises distribution with a standard deviation σ (variability; the mean is set to zero), and n is the total number of possible values for the target's feature.

CROWDING AND BINDING

- B. The bias mixture model (Equation 2)** has three free parameters. In addition to the variability and guessing rate, this model adds a free parameter for the mean (μ) of the error distribution:

$$p(\theta) = (1 - \gamma)f(\theta)_{\mu, \sigma} + \gamma\left(\frac{1}{n}\right) \quad (2)$$

- C. The standard misreport model (Equation 3)** has three free parameters. The model adds a misreporting component to the standard mixture model, which describes the probability of reporting one of the flankers to be the target:

$$p(\theta) = (1 - \gamma - \beta)f(\theta)_{\sigma} + \gamma\left(\frac{1}{n}\right) + \beta\frac{1}{m}\sum_i^m f(\theta_i^*)_{\sigma} \quad (3)$$

where β is the probability of reporting a flanker as the target, m is the total number of non-target items (two in the present study), and θ_i^* is the error to the feature of the i^{th} flanker. Notice that $f(\theta)$, the von Mises distribution (for color and orientation) or the Gaussian distribution (for SF) of the estimation errors here describes the distribution when the observer correctly estimated the target's feature, thus its mean is zero; whereas for $f(\theta_i^*)$ the distribution of estimating one flanker, the mean would be the feature distance of the corresponding flanker to that of the target. The variability of the distributions for each stimulus was assumed to be the same.

- D. The bias misreport model (Equation 4)** has four free parameters. In addition to the standard misreport model, there is a free parameter for the mean (μ) of the distribution of estimating the target's feature, to better account for possible pooling and substitution:

$$p(\theta) = (1 - \gamma - \beta)f(\theta)_{\mu, \sigma} + \gamma\left(\frac{1}{n}\right) + \beta\frac{1}{m}\sum_i^m f(\theta_i^*)_{\sigma} \quad (4)$$

- E. The educated guess model (Equation 5)** has four free parameters. The model adds a misreporting component of the guessed stimuli other than the stimuli presented to the standard misreport model. This misreporting component is similar as the misreporting component of the flankers, but has a different probability:

$$p(\theta) = (1 - \gamma - \beta_F - \beta_G)f(\theta)_{\sigma} + \gamma\left(\frac{1}{n}\right) + \beta_F\frac{1}{m}\sum_i^m f(\theta_i^*)_{\sigma} + \beta_G\frac{1}{k}\sum_i^k f(\theta_i^{**})_{\sigma} \quad (5)$$

where β_F is the probability of misreporting a flanker as the target, and β_G is the probability of misreporting a guessed feature other than the flankers presented. This model follows the assumption that the participant may have the information of one feature, then guess the feature of the target based on all possible target-flanker distances. This “educated guess” could result in misreporting one flanker as the target, or misreporting features not presented in the corresponding trial. For example, there are four possible target-flanker distances: -90, -60, 60, and 90. In a 60/-60 trial, for each detected feature, there are five possible feature values (including the stimulus itself) that could be the possible target; therefore, there will be

CROWDING AND BINDING

two flankers (-60, 60; $m = 2$) with the misreporting probability β_F , and ten not presented guessed stimuli (some of them are overlapped, and the non-overlapping feature values are -150, -120, -90, -30, 30, 90, 120, 150; $k = 10$) with the misreporting probability β_G .

We used the MemToolBox [48] for fitting the models, and compared Akaike information criterion with correction (AICc) to assess model fits.

Experiment 2: Orientation and Spatial Frequency

Observers. Fourteen undergraduate students from New York University participated in this experiment (12 females, age ranged from 18 to 21 years old, 19.43 ± 0.85). All observers were naïve to the purposes of the study, and all reported having normal or corrected-to-normal visual acuity. Written informed consent was obtained from all participants before the experiment. The University Committee on Activities Involving Human Subjects at New York University approved experimental procedures.

Apparatus, stimuli, procedure and design. Sequence of events within a trial and sample stimulus displays are presented in **Figure 3**. The apparatus, stimuli, procedure and design were the same as in Experiment 1 except for the following changes: Target and flankers were sinusoidal gratings (Gabor) with a 2D Gaussian spatial envelope (standard deviation 0.325° and 85% contrast). The orientation and SF of the target were each randomly and independently selected from two parameters spaces. The viewing distance was 91 cm. Target and flankers' center-to-center distance was 2.15° . The orientation parameter space corresponded to range of angle 1° - 180° . Stimulus display duration was 200 ms.

Phase 1

To determine the SF values of the estimation task, we tested the target-alone condition with different SF values, linearly spaced. Based on this test we set the SF parameter space to correspond to the range of 1-5 cpd. Because SF discriminability varies across SF values [49], we scaled the 180 unit steps of SF that were used in Phase 2 according to the variation in the estimation task of SF values in Phase 1. To this aim we fit an exponential function to the SD of the estimation data for each SF value in Phase 1 (**Figure S4**).

Phase 2

Design. Flankers' orientation and SF differed from the target by either ± 40 or ± 70 units of orientation and SF units. Each flanker had the same absolute target-flanker difference in both feature-dimensions. Within each feature dimension, one flanker had a positive and the other negative target-flanker difference (randomly selected), such that each flanker had a unique relation to the target (negative or positive) in both feature-dimensions. Target flanker distance within the parameter space was: (1) 40/-40; (3) 70/-70, (3) 40/-70, 70/-40, which were labeled ± 40 , ± 70 and 40/-70 respectively.

In all trials the target orientation was randomly selected between 1° - 180° in a step size of 2° . In target-alone trials, the target SF was randomly selected out of the 180 steps, as determined in Phase 1. However, because SF is not a circular space, in the crowding display flankers' SF

CROWDING AND BINDING

values restricted the range of target SF values, such that the target SF ranged from 41-140, 71-110, and 41-110/ 71-140 SF steps in ± 40 , ± 70 and 40/-70 flankers conditions respectively.

There were two response displays. In the orientation response display observers had to estimate the target orientation by pointing and clicking the mouse cursor at a position on the orientation wheel. In the SF response display observers estimated SF by pointing and clicking the mouse cursor on a centered horizontal line (0.08°-thick two-directional arrow with 11.2° in length). A “-” on one side and “+” on the other side indicated the direction of SF increase. Because SF is not circular, we extended the range of the SF response (0.48-7.35 cpd) beyond that of the target (1-5 cpd).

Models and analyses. Because SF is not circular, in the mixture models we used a Gaussian distribution when fitting to SF and von Mises distribution when fitting to orientation. Because in flanker conditions we had to restrict the range of target SF, when comparing SF flanker and the target-alone conditions we equated the range of target SF.

References

1. Levi DM. Crowding--an essential bottleneck for object recognition: a mini-review. *Vision Research*. 2008;48: 635–654. doi:10.1016/j.visres.2007.12.009
2. Bouma H. Interaction effects in parafoveal letter recognition. *Nature*. 1970;226: 177–178. doi:10.1038/226177a0
3. Pelli DG, Tillman KA. The uncrowded window of object recognition. *Nature Neuroscience*. 2008;11: 1129–1135.
4. Pelli DG, Palomares M, Majaj NJ. Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*. 2004;4: 1136–1169. doi:10.1167/4.12.12
5. Levi DM, Klein SA. Sampling in spatial vision. *Nature*. 1986;320: 360–362. doi:10.1038/320360a0
6. Levi DM, Klein SA, Yap YL. Positional uncertainty in peripheral and amblyopic vision. *Vision Research*. Pergamon; 1987;27: 581–597. doi:10.1016/0042-6989(87)90044-7
7. Freeman J, Simoncelli EP. Metamers of the ventral stream. *Nature Neuroscience*. 2011;14: 1195–1201. doi:10.1038/nn.2889
8. Martelli M, Majaj NJ, Pelli DG. Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*. 2005;5: 59–70. doi:10.1167/5.1.6
9. Whitney D, Levi DM. Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in cognitive sciences*. 2011;15: 160–168. doi:10.1016/j.tics.2011.02.005
10. Nandy AS, Tjan BS. Saccade-confounded image statistics explain visual crowding. *Nature Publishing Group. Nature Publishing Group*; 2012;15: 463.

CROWDING AND BINDING

11. Manassi M, Whitney D. Multi-level Crowding and the Paradox of Object Recognition in Clutter. *Current Biology*. Elsevier Ltd; 2018;28: R127–R133. doi:10.1016/j.cub.2017.12.051
12. Kimchi R, Pirkner Y. Multiple Level Crowding: Crowding at the Object Parts Level and at the Object Configural level. *Perception*. 2015;44: 1275–1292.
13. Pirkner Y, Kimchi R. Crowding and perceptual organization: Target's objecthood influences the relative strength of part-level and configural-level crowding. *Journal of Vision*. The Association for Research in Vision and Ophthalmology; 2017;17: 1–18. doi:10.1167/17.11.7
14. Ester EF, Zilber E, Serences JT. Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *Journal of Vision*. 2015;15: 4–4. doi:10.1167/15.1.4
15. Huckauf A, Heller D. What various kinds of errors tell us about lateral masking effects. *Visual Cognition*. Routledge; 2014;9: 889–910. doi:10.1080/13506280143000548A
16. Strasburger H, Harvey LO, Rentschler I. Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*. Springer New York; 1991;49: 495–508. doi:10.3758/BF03212183
17. Strasburger H, Malania M. Source confusion is a major cause of crowding. *Journal of Vision*. 2008;13: 1–20. doi:10.1167/13.1.24.Introduction
18. Ester EF, Klee D, Awh E. Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology: Human Perception and Performance*. American Psychological Association; 2014;40: 1022.
19. Harrison WJ, Bex PJ. A unifying model of orientation crowding in peripheral vision. *Current Biology*. 2015;25: 3213–3219. doi:10.1016/j.cub.2015.10.052
20. Greenwood JA, Bex PJ, Dakin SC. Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences*. 2009;106: 13130–13135. doi:10.1073/pnas.0901352106
21. Keshvari S, Rosenholtz R. Pooling of continuous features provides a unifying account of crowding. *Journal of Vision*. 2016;16: 1–15. doi:10.1167/16.3.39
22. van den Berg R, Roerdink JBTM, Cornelissen FW. A neurophysiologically plausible population code model for feature integration explains visual crowding. Kording KP, editor. *PLoS Comput Biol*. 2010;6: e1000646. doi:10.1371/journal.pcbi.1000646
23. Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*. Nature Publishing Group; 2001;4: 739–744.
24. Yashar A, Chen J, Carrasco M. Rapid and long-lasting reduction of crowding through training. *Journal of Vision*. The Association for Research in Vision and Ophthalmology;

CROWDING AND BINDING

- 2015;15: 1–15. doi:10.1167/15.10.15
25. Greenwood JA, Bex PJ, Dakin SC. Crowding changes appearance. *Current Biology*. 2010;20: 496–501. doi:10.1016/j.cub.2010.01.023
26. van den Berg R, Roerdink JBTM, Cornelissen FW. On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. *Journal of Vision*. 2007;7: 14–14. doi:10.1167/7.2.14
27. Pöder E, Wagemans J. Crowding with conjunctions of simple features. *Journal of Vision*. 2007;7: 23–12. doi:10.1167/7.2.23
28. Greenwood JA, Bex PJ, Dakin SC. Crowding follows the binding of relative position and orientation. *Journal of Vision*. 2012;12: 18–18. doi:10.1167/12.3.18
29. Pelli DG. Crowding: A cortical constraint on object recognition. *Current opinion in neurobiology*. Elsevier; 2008;18: 445–451.
30. Toet A, Levi DM. The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*. Elsevier; 1992;32: 1349–1357.
31. Derrington AM, Krauskopf J, Lennie P. Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of physiology*. Wiley Online Library; 1984;357: 241–265.
32. Fuller S, Carrasco M. Exogenous attention and color perception: Performance and appearance of saturation and hue. *Vision Research*. Pergamon; 2006;46: 4032–4047. doi:10.1016/j.visres.2006.07.014
33. Carandini M. Do we know what the early visual system does? *Journal of Neuroscience*. 2005;25: 10577–10597. doi:10.1523/JNEUROSCI.3726-05.2005
34. Mazer JA, Vinje WE, McDermott J, Schiller PH, Gallant JL. Spatial frequency and orientation tuning dynamics in area V1. *Proceedings of the National Academy of Sciences*. 2002;99: 1645–1650. doi:10.1073/pnas.022638499
35. De Valois RL, Albrecht DG, Thorell LG. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*. 1982;22: 545–559. doi:10.1016/0042-6989(82)90113-4
36. Morey RD. Confidence intervals from normalized data : A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*. 2008;4: 61–64. doi:10.3758/s13414-012-0291-2
37. Banks MS, Sekuler AB, Anderson SJ. Peripheral spatial vision: Limits imposed by optics, photoreceptors, and receptor pooling. *JOSA A*. Optical Society of America; 1991;8: 1775–1787.
38. Howell ER, Hess RF. The functional area for summation to threshold for sinusoidal gratings. *Vision Research*. Elsevier; 1978;18: 369–374.

CROWDING AND BINDING

39. Robson JG, Graham N. Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*. Elsevier; 1981;21: 409–418.
40. Tolhurst DJ, Thompson ID. On the variety of spatial frequency selectivities shown by neurons in area 17 of the cat. *Proc R Soc Lond B. The Royal Society*; 1981;213: 183–199.
41. Goris RLT, Simoncelli EP, Movshon JA. Origin and Function of Tuning Diversity in Macaque Visual Cortex. *Neuron*. 2015;88: 819–831. doi:10.1016/j.neuron.2015.10.009
42. Mareschal I, Morgan MJ, Solomon JA. Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision*. 2010;10: 13–13. doi:10.1167/10.8.13
43. Di Lollo V. The feature-binding problem is an ill-posed problem. *Trends in cognitive sciences*. 2012;16: 317–321. doi:10.1016/j.tics.2012.04.007
44. Golomb JD, L'Heureux ZE, Kanwisher N. Feature-binding errors after eye movements and shifts of attention. *Psychological Science*. 2014;25: 1067–1078. doi:10.1177/0956797614522068
45. Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C. What's new in Psychtoolbox-3. *Perception*. Alezzo; 2007;36: 1.
46. Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. *Nature*. 2008;453: 233–235. doi:10.1038/nature06860
47. Bays PM, Catalao RFG, Husain M. The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision. The Association for Research in Vision and Ophthalmology*; 2009;9: 1–11. doi:10.1167/9.10.7
48. Suchow JW, Brady TF, Fournie D, Alvarez GA. Modeling visual working memory with the MemToolbox. *Journal of Vision*. 2013;13: 9–9. doi:10.1167/13.10.9
49. Caelli T, Brettel H, Rentschler I, Hilz R. Discrimination thresholds in the two-dimensional spatial frequency domain. *Vision Research*. Elsevier; 1983;23: 129–133.

CROWDING AND BINDING

Supplementary Information

Table S1. Best-fitting parameters for each condition in Experiment 1, Experiment 2 and Supplementary experiment. For orientation and color, parameters are from the standard mixture model for the target-alone condition, and the standard with misreport model for the flanker conditions. For SF data, parameters are from the bias model.

Feature Dimension	Condition		Parameters (mean $\pm$ SD)			
			Bias (μ)	Guessing rate (γ)	Swapping rate (β)	Variability (σ)
Experiment 1						
Orientation	Target- alone		N/A	.03 $\pm$.04	N/A	13.03 $\pm$ 2.77
	Symmetrical Flankers	± 60	N/A	.22 $\pm$.10	.21 $\pm$.12	15.79 $\pm$ 3.95
		± 90	N/A	.18 $\pm$.12	.34 $\pm$.11	14.83 $\pm$ 2.91
	Asymmetrical Flankers	60/-90	N/A	.21 $\pm$.11	.27 $\pm$.12	15.94 $\pm$ 4.28
	Color	Target- alone		N/A	.02 $\pm$.02	N/A
Symmetrical Flankers		± 60	N/A	.04 $\pm$.03	.12 $\pm$.08	20.09 $\pm$ 3.96
		± 90	N/A	.04 $\pm$.03	.07 $\pm$.05	19.25 $\pm$ 2.33
Asymmetrical Flankers		60/-90	N/A	.03 $\pm$.04	.07 $\pm$.05	21.59 $\pm$ 5.68

CROWDING AND BINDING

Experiment 2

Orientation	Target- alone		N/A	.09 ± .09	N/A	16.10 ± 3.54
	Symmetrical Flankers	±40	N/A	.16 ± .17	.15 ± .16	20.60 ± 9.38
		±70	N/A	.28 ± .21	.18 ± .16	18.49 ± 6.33
	Asymmetrical Flankers	40/-70	N/A	.25 ± .24	.19 ± .15	17.72 ± 4.39
SF	Target- alone	Full range	N/A	.01 ± .02	N/A	33.00 ± 8.49
		Range as ±40	-1.14 ±10.42	.005 ± .01	N/A	25.95 ± 8.69
		Range as ±70	0.56 ±10.42	.001 ± .01	N/A	23.38 ±9.11
		Range as 40/-70	-7.43 ±5.55	.004 ± .01	N/A	26.78 ± 8.14
	Symmetrical Flankers	±40	-2.73 ± 2.37	.003 ± .01	N/A	34.11 ± 9.68
		±70	-1.66 ±3.80	.004 ± .01	N/A	32.71 ± 9.57
	Asymmetrical Flankers	40/-70	-12.95 ± 3.73	.002 ± .01	N/A	32.02 ± 7.96

Supplementary experiment

Orientation	Target- alone		1.48 ±	.02 ± .02	N/A	17.34 ± 4.41
	Symmetrical Flankers	±60	0.98 ±	.24 ± .14	.23 ± .16	19.19 ± 6.65
		±90	3.36 ±	.18 ± .09	.29 ± .14	20.22 ± 8.08
	Asymmetrical Flankers	60/-90	-0.15 ±	.14 ± .08	.35 ± .11	19.04 ± 6.40
Color	Target- alone		5.02 ± 3.47	.01 ± .01	N/A	19.23 ± 3.50
	Symmetrical Flankers	±60	-0.50 ± 2.40	.05 ± .07	.09 ± .11	24.51 ± 5.67
		±90	-4.53 ± 3.75	.04 ± .05	.07 ± .09	24.56 ± 6.48
	Asymmetrical Flankers	60/-90	-0.71 ± 2.61	.05 ± .04	.06 ± .08	23.03 ± 4.68

CROWDING AND BINDING

Supplementary Experiment

In Experiment 1 and 2 observers reported each feature separately, which may have encouraged observers to decode each feature independently during response, which may have led to estimation errors that are uncorrelated. We addressed this alternative explanation in this Supplementary Experiment by using only one response display and having observers simultaneously report both color and orientation.

Method

Observers. Fourteen undergraduate and graduate students from New York University participated in this study (10 females, age range 20 to 27 years old, 21.79 ± 2.04). All observers were naïve as to the purposes of the study, and all reported having normal or corrected-to-normal visual acuity and normal color vision. Written informed consent were obtained from all observers before the experiment. The University Committee on Activities Involving Human Subjects at New York University approved experimental procedures.

Stimuli and procedure. The stimuli and procedure were the same as in Experiment 1 except for the following changes. Target orientation was randomly selected out of 24 values evenly distributed between 1° - 360° . Target color was randomly selected out of 24 values evenly distributed along the circle in the DKL color space as defined in Experiment 1. The sets of 24 orientations and 24 colors were independently selected randomly for each observer. The response display was a 24×24 matrix of T-shape items, which comprised all the possible conjunctions of target orientation and color (**Figure S1**), and were ordered according to their orientation (vertical order) and color (horizontal order) values. Observers were instructed to report the target by pointing and clicking the mouse cursor over the item with the orientation and color of the target. To facilitate report, the matrix was the same for each observer, and observers were allowed to saccade during the response.

Results

Distributions of errors within each feature-dimension

Figure S1a plots the distribution of estimation errors for orientation and color in target-alone trials. In both feature-dimensions the error distribution was accurately described by a von Mises function with an added uniform distribution. One participant was excluded from the analysis due to a large variability in orientation errors and poor fits for all models. For orientation errors, the von Mises function centered on 1.48 ± 2.15 away from target value with a variability (σ) of 17.34 ± 4.41 could describe well the error distributions, and for color errors the circular Gaussian function centered on 5.02 ± 3.47 away from target value, with a variability of 19.23 ± 3.50 , could describe well the error distributions. The probability of the uniform distribution was significantly higher than zero for orientation ($\gamma_o = 0.02 \pm 0.02$, $t(12) = 3.43$, $p < .01$), and marginally so for color ($\gamma_c = 0.01 \pm 0.01$, $t(12) = 1.90$, $p = .08$).

CROWDING AND BINDING

Figure S1. Supplementary Experiment distribution of errors relative to target feature values. Frequency of response is plotted as a function of the deviation between estimation report and target feature values (x-axis units are SD in target-alone trials): for orientation (top), color (right), and conjunction of both features (heat map) in the (A) target-alone condition, (B) ± 60 , (C) 60/-90 and (D) ± 90 flankers conditions. Solid lines indicate the response probabilities predicted by the bias with misreport model.

Figure S1b, S1c, S1d plot the distribution of estimation errors for orientation and color in ± 60 , 60/-90 and ± 90 flanker conditions respectively. As in Experiment 1 for both orientation and color the models with misreport outperformed the models without the misreport component suggesting that in both feature dimensions observers confused flanker feature with that of the target. We then analyzed the parameters in the individual fitting of the bias with misreport model between conditions. Compared to the target-alone condition, guessing rate was higher in the crowded conditions as indicated by the increase in the uniform distribution, $t_{s(12)} > 2.73$, $p < .02$. Also, a significant proportion of the errors concentrated (i.e. von Mises distributed) around the value of one of the two flanker items (misreporting errors). The misreporting rate was

CROWDING AND BINDING

significantly larger than zero in all flanker conditions, $ts(12) > 2.60$, $ps < .03$. In all flanker conditions, the variability of the color errors concentrated around the target or the flanker showed significant increase compared to the target-alone condition, $ts(12) > 4.59$, $ps < .001$, while the variability of the orientation errors showed no significant increase in ± 60 and $60/-90$ conditions, $ts(12) < 2.06$, $ps > .06$, but was significantly increased in the ± 90 condition, $t(12) = 2.23$, $p < .05$. For color errors, when compared to the target-alone condition, all flanker conditions showed significant difference in mean (more toward the negative direction), $ts(12) > 5.67$, $ps < .001$; when compared to zero, however, only in the $60/-90$ condition the mean of the von Mises distribution was significantly different (smaller than zero), $t(12) = -4.35$, $p < .002$. For orientation errors, when compared to the target-alone condition, no flanker conditions showed significant difference in mean, $ts(12) < 1.34$, $ps > .21$; when compared to zero, only in the $60/-90$ condition the mean of the von Mises distribution was significantly different (larger than zero), $t(12) = 2.85$, $p < .02$.

Joint distributions of errors

Similar to Experiment 2, only 3 observers had significant Pearson correlation between orientation and color estimations under the target-alone and the flanker conditions (collapsed), overall mean $r = 0.04 \pm 0.03$ and $r = 0.02 \pm 0.02$, respectively.

CROWDING AND BINDING

Figure S2. Experiment 2 error distributions for orientation ± 40 flanker condition (top panel) and SF ± 70 flanker condition (bottom panel). Even when the flanker feature distance in SD was similar between orientation and SF, observers misreport flanker orientation but averaged SF.

CROWDING AND BINDING

Supplementary Results

Analyses of consistent vs. inconsistent trials. We conducted a two-way (3 flanker conditions x 2 trial types) repeated measures Analysis of Variance (ANOVA) on the parameters of the best fitted models (standard misreport for color and orientation, bias for SF). There were no main effects of consistency on misreporting rate, guessing rate or variability in Experiment 1 in either orientation or color, $F_s < 2.2$, $p_s > .1$. In Experiment 2, however, the proportion of orientation misreporting errors was higher in consistent than inconsistent trials, $F(1,13)=11.20$, $p < .01$. There was no significant effect of consistency on the variability of errors around the target or guessing rates in orientation, all $p_s > .4$. In SF, the variability of errors around the target was higher in consistent than inconsistent trials, $F(1,13)=19.86$, $p < .002$. There was no significant main effect of consistency on mean bias or guessing rate, $p_s > 0.07$.

CROWDING AND BINDING

Supplementary Method

Population coding model

Stimulation of population coding of orientation and SF

Table S2: Equations and parameters of the simulation of joint coding of orientation and SF

Model component	Equation	Description
Report SF	$re(x_{SF}) = \frac{\sum_{i=1}^m rp_i(x_{Or}, x_{SF})}{m}$	In each trial, SF report is the outcome of pooling of readouts from multiple populations. x_{Or} and x_{SF} are orientation and SF values respectively. $re(x_{SF})$ is the SF report in a given trial. rp_i is the sample from the i population; there are three populations: $m = 3$.
Report orientation	$re(x_{Or}) = rp_1(x_{Or}, x_{SF})$	In each trial, $re(x_{Or})$ (orientation report) is the outcome of a readout from the population with receptive-field (RF) over the center of the target: rp_1 .
Readout values	$rp_i(x_{Or}, x_{SF}) \sim C_i(x_{Or}, x_{SF})$	C_i is a population that jointly codes orientation and SF. Populations are represented as a probability function that determine readout values (18).
Bivariate mixture models	<p>Target-alone condition</p> $C_i(x_{Or}, x_{SF}) = P_1(x_{Or}, x_{SF}) \left(w_{i,1} - g \right) + g \frac{1}{k^2}$ <p>Flanker conditions</p> $C_i(x_{Or}, x_{SF}) = P_1(x_{Or}, x_{SF}) \left(w_{i,1} - g \right) + g \frac{1}{k^2} + \sum_{j=2}^n P_j(x_{Or}, x_{SF}) w_{i,j}$	In the target-alone condition, C_i is a mixture of bivariate normal distribution over the target values P_j and a bivariate uniform distribution: $\frac{1}{k^2}$. $k = 180$ is the number of possible values in each feature-dimension. In flanker conditions additional bivariate normal distributions were added over the flankers' values: P_j . $w_{i,j}$ is the probability of reading-out item j from

CROWDING AND BINDING

		population i . g is the probability for reading out a value that is unrelated to the target or the flankers, determined by orientation guessing rate in the standard with misreport model fitted to the data.
Coding precision	$P_j(x_{Or}, x_{SF}) = \frac{1}{2\pi\sigma_{Or}\sigma_{SF}} e^{-\frac{(x_{Or}-\mu_{Or_j})^2}{\sigma_{Or}^2} - \frac{(x_{SF}-\mu_{SF_j})^2}{\sigma_{SF}^2}}$	σ_{Or} is the inverse precision of orientation representation, and was determined by the fitting of the standard with misreport model to the orientation data. σ_{SF} is a free parameter that was fitted to the data. It is the SD of SF coding before pooling. It is different than the SD of the data. In the fitted simulation $\sigma_{SF} = 56.10$. μ_{Or_j} and μ_{SF_j} are the true orientation and SF value of the j item.
Item probability	$w_{i,j} = \frac{\tilde{w}_{i,j}}{\sum_{j=1}^n \tilde{w}_{i,j}}$	$\tilde{w}_{i,j}$ is a spatial weight determined by the RF Gaussian function.
Spatial weights	$\tilde{w}_{i,j} = \frac{1}{\sigma_r \sqrt{2\pi}} \int_{s_j-1}^{s_j+1} e^{-\frac{1}{2} \left(\frac{x-\mu_{r_i}}{\sigma_r} \right)^2} dx$	$\mu_{r_i} = [0, -d, d]$. $d = 4.28$ is a free parameter that was fitted to the data. s_j is the center of the item j in respect to the center of the RF. The size of s_j was 2° and the weight of s_j was the integral of the RF Gaussian distribution over the area between $s_j - 1$ and $s_j + 1$. The SD of RF Gaussian distributions was determined such that flankers' weights with RF over the target correspond to the misreporting rate in the orientation data: $\sigma_r = 0.88$. For simplicity we assumed the same σ_r for all populations.

CROWDING AND BINDING

Table S3: Equations and parameters of the simulation of the independent coding of orientation and SF. The model is the same as the joint coding model except for the following changes.

Model component	Equation	Description
Readout values	$rp_i(x_{Or}) \sim C_i(x_{Or})$ $rp_i(x_{SF}) \sim C_i(x_{SF})$	C_i is a population that independently codes orientation or SF.
Bivariate mixture models	<p>Target-alone condtion</p> $C_i(x_{Or}) = P_1(x_{Or})(w_{i,1} - g) + g \frac{1}{k^2}$ $C_i(x_{SF}) = P_1(x_{SF})(w_{i,1} - g) + g \frac{1}{k^2}$ <p>Flanker conditions</p> $C_i(x_{Or}) = P_1(x_{Or})(w_{i,1} - g) + g \frac{1}{k^2} + \sum_{j=2}^n P_j(x_{Or}) w_{i,j}$ $C_i(x_{SF}) = P_1(x_{SF})(w_{i,1} - g) + g \frac{1}{k^2} + \sum_{j=2}^n P_j(x_{SF}) w_{i,j}$	<p>In the target-alone condition, C_i is a mixture of normal distributions over the target values P_j and a uniform distribution: $\frac{1}{k}$. In flanker conditions additional normal distributions were added over the flankers' values:</p>
Coding precision	$P(x_{Or}) = \frac{1}{\sigma_{Or} \sqrt{2\pi}} e^{-\frac{(x_{Or} - \mu_{Or})^2}{2\sigma_{Or}^2}}$ $P(x_{SF}) = \frac{1}{\sigma_{SF} \sqrt{2\pi}} e^{-\frac{(x_{SF} - \mu_{SF})^2}{2\sigma_{SF}^2}}$	<p>σ_{Or} is the inverse precision of orientation representation, and was determined by the fitting of the standard with misreport model to the orientation data. σ_{SF} is a free parameter that was fitted to the data. It is the SD of SF coding before pooling. It is different than the SD of the data,. In the fitted simulation $\sigma_{SF} = 56.10$. μ_{Or_j} and μ_{SF_j} are the true orientation and SF value of the j item.</p>

Coding model fitting

CROWDING AND BINDING

To fit the model we simulated 2800 trials (same as in the experiment data for all observers together). In each of the four experimental conditions. As in Experiment 2 data, in the target-alone condition we fitted the standard mixture model, and in the flanker conditions we fitted the standard with misreport model to the orientation report and the with bias model to the SF report. To fit the simulations to the data, we minimized the sum of squared errors between the simulated data and Experiment 2 data using the Matlab function 'patternsearch'. Several starting points for the parameters were used to check the consistency of the model fits. We compared between two

CROWDING AND BINDING

DKL colors

The DKL spherical color space is defined by three orthogonal axes: two opponent chromatic cone dimensions (L-M and S-(L+M)) and one luminance (L + M). Here we specified DKL color by calibrating the monitor using the Spectroradiometer PR-670 and the PsychoPy toolbox (<http://www.psychopy.org/general/colours.html#dkl1984>).

Specifying coordinates in DKL requires selection of a reference cone activation vector, usually derived from the background of the visual display and representing the adaptation state. To simplify the calculations for this experiment, we chose cone values corresponding to the achromatic (white) points at the background luminance (56 cd/m²). Following the convention of Derrington et al. (1984), DKL values were expressed in spherical coordinates. Luminance corresponded to elevation angle, which was zero for all colors here. Azimuth angles controlled hue, with the zero azimuth representing activation along the L - M axis in the direction of positive L-M values. Radius controlled distance from the white point, or saturation. To generate the color wheel we picked 180 hue values (angles), which were evenly distributed between 0°-359°, with a fixed radius of 0.9.

CROWDING AND BINDING

Figure S3. Color estimation performance. Scatter plot of the color estimation (y-axis) as a function of target color values (x-axis) in the target-alone condition. The results show color estimates and the true color value are along the diagonal, indicating that colors were accurately estimated and rules out the possibility that the task could have been mediated by color categories (e.g., red, blue).

CROWDING AND BINDING

Figure S4. Phase 1 of Experiment 2. SD for the SF estimation task (y-axis) as a function of SF value (x-axis). An exponential function fit the data and was used to scale the SF steps in Phase 2.