Positional averaging explains crowding with letter-like stimuli

John A. Greenwood[a,1], Peter J. Bex[a,b], and Steven C. Dakin[a]

[a]UCL Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, United Kingdom; and [b]Schepens Eye Research Institute, Harvard Medical School, Boston, MA 02114

Visual crowding is a breakdown in object identification that occurs in cluttered scenes, a process that represents the principle restriction on visual performance in the periphery. When crowded objects are presented experimentally, a key finding is that observers frequently report nearby flanking items instead of the target. This observation has led to the proposal that crowding reflects increased noise in the positional code for objects; although how the presence of nearby objects might disrupt positional encoding remains unclear. We quantified this disruption using cross-like stimuli, where observers judged whether the horizontal target line was positioned above or below the stimulus midpoint. Overall, observers were poorer at judging position in the presence of crowding flankers. However, offsetting horizontal lines in the flankers also led observers to report that the horizontal line in the target was shifted in the same direction, an effect that held for subthreshold flanker offsets. In short, crowding induced both random and systematic errors in observers' judgment of position, with or without the detection of flanker structure. Computational modeling reveals that perceived position in the presence of flankers follows a weighted average of noisy target- and flanker-line positions, rather than a substitution of flanker-features into the target, as has been proposed previously. Together, our results suggest that crowding is a preattentive process that uses averaging to regularize the noisy representation of position in the periphery.

peripheral visual field | texture recognition | context | lateral interaction

When multiple objects are presented in close proximity, their identities can become obscured through a process known as crowding (1, 2). This impairment in identification occurs across a wide range of visual modalities (3–6) and increases in magnitude when target items are presented peripherally, making crowding the principle restriction on our representation of the peripheral visual field (7–9). Although a range of theories have been proposed to explain crowding (reviewed recently in ref. 8), there is little consensus regarding their plausibility, and a scarcity of computational models that can account for the effects of crowding with complex stimuli such as letters.

One of the most striking features of crowding is that observers' identification errors frequently arise from them reporting the identity of a flanking item instead of the target (10–14). Such errors are particularly apparent with multifeature stimuli such as letters (15), where the subjective experience of crowding is that features from each object become "jumbled" (9). This observation has led to the proposal that crowding results from noise in the encoding of position, leading to an intermixing of either the feature positions within each object (12, 16), or the gross position of objects as a whole (14, 17). These proposals are supported by the correlation between the magnitude of crowding—which increases as objects move further into the peripheral visual field (1)—and the decline in position encoding in the periphery (18–20). Similarly, errors for both absolute judgments of position (21) and saccades to eccentric targets (22) exhibit the same asymmetric pattern as crowding (23, 24), with greater errors

along the radial axis from fixation than along the tangential (i.e., orthogonal) axis. It follows that the intermixing of target and flanker features under crowding could simply be the result of the positional uncertainty inherent to peripheral vision.

However, despite the apparent relationship between positional uncertainty and crowding, the precise effect of crowding on the perceived position of objects and their features remains unclear. Although crowding has been shown to affect judgments of relative position in vernier tasks, with thresholds rising sharply in the presence of additional flanking lines (25, 26), vernier acuity also relies upon luminance contrast and orientation information (27, 28), making it unclear whether poor performance reflects positional uncertainty per se. The range and selectivity of these interactions also differs from patterns established with crowding in other domains (9), suggesting that the effects on vernier acuity may arise from processes more akin to masking than crowding. The aim of the present study was thus to directly examine the influence of crowding on the perceived position of object features.

Of the many theories proposed to explain crowding, 2 clear predictions emerge for the effect of crowding on position encoding. First, the disproportionate amount of flanker-related errors (10–14) suggests that crowding arises from the features of flankers (or flankers in their entirety) being substituted into the target. This substitution has been attributed to both low-level processes such as positional uncertainty (12, 13, 16, 17), and higher-order operations such as unfocused spatial attention (10), a distinction that follows a more general debate regarding the locus of crowding (8). A second prediction emerges from the view that crowding is a preattentive and compulsory averaging of target and flanker features. For example, crowded Gabor elements can contribute to a global estimate of orientation (29), although the exact combination of elements may follow more directed pooling schemes according to the modality involved (5, 30). Were this model to apply in the positional domain, the position of target and flanking objects would be averaged, an effect that may serve to simplify the appearance of noisy positional structure in the periphery.

To examine the influence of crowding on position encoding and so distinguish between "averaging" and "substitution" hypotheses, observers were presented with cross-like stimuli, and judged whether the horizontal target line was above or below the stimulus midpoint (see *Methods* and Fig. 1*A*). Stimuli were "letter-like" in that they were multifeatured and had proportions equivalent to those of Sloan letters. An additional experiment required observers to make a continuous (i.e., nondiscrete) estimate of the position of target features, allowing a more

---

A   Example stimuli (two sets)

B   Example data (JAG)
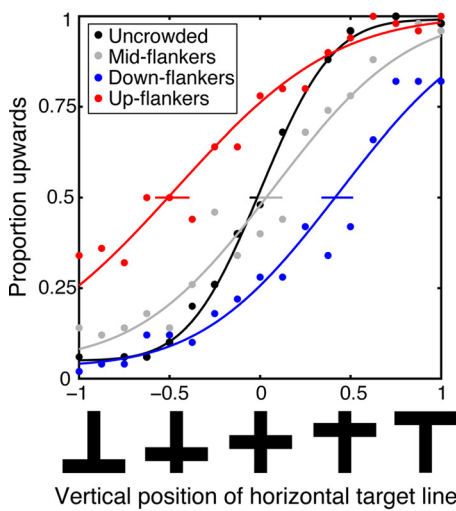
C   Midpoint shifts (bias)

D   Threshold elevation

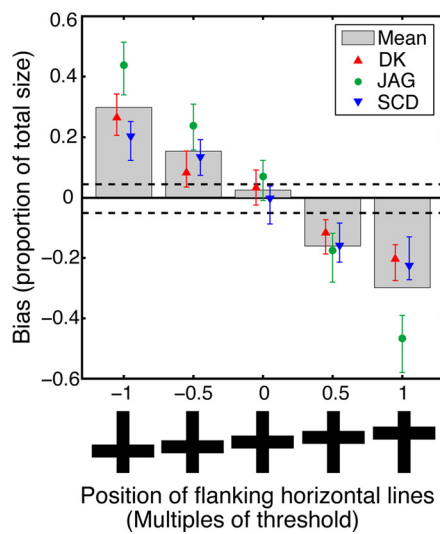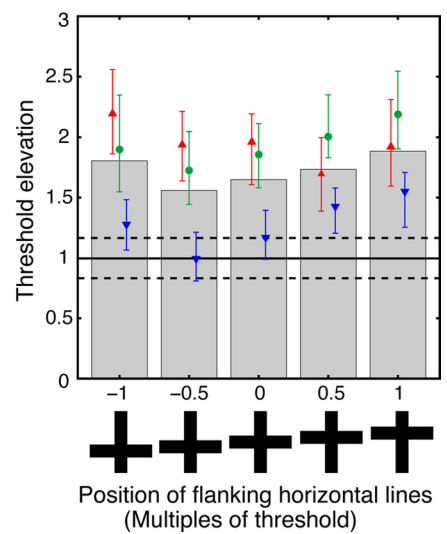**Fig. 1.** The effects of crowding on position encoding. (*A*) Example stimuli, with the middle target-cross depicted with the horizontal line centered along the vertical axis. Flankers on the left have horizontal lines positioned above the midpoint (up-flankers), with the converse down-flankers on the right. The effect of crowding can be seen by fixating near the bottom of the figure caption: Target-horizontals should appear shifted toward the flanker-horizontals. (*B*) Data from 1 observer, with the proportion of 'upward' responses plotted as a function of the vertical position of the target-horizontal (shown schematically on the abscissa), and fit with psychometric functions. 95% confidence intervals are depicted around the midpoints, obtained from bootstrapping. Data are presented for targets in isolation (uncrowded), and crowded conditions with midflankers and up- and down-flankers offset by threshold-level displacements. Crowding both increases the slope and shifts psychometric functions to a degree that depends on flanking structure. (*C*) Midpoints of the psychometric functions (bias) for 3 observers in each condition, expressed as a proportion of stimulus size (scaled to ± 1) and plotted as a function of the vertical position of flanker-horizontals (expressed in multiples of the uncrowded threshold for position discrimination). Mean data are presented as gray bars, with individual points superimposed; error bars (and hatched lines for uncrowded data) depict 95% confidence intervals. Positive shifts in flanker feature positions (even subthreshold ones) produce negative shifts in bias, indicating an increase in 'upward' responses with up-flankers, and vice versa. (*D*) As in 1*C* for threshold elevation. The presence of flankers produced a strong elevation of positional thresholds, regardless of their structure.

detailed comparison of the predictions of averaging and substitution models to human performance.

Results

In the absence of flankers, observers' judgments of the position of a horizontal target line (above or below the midpoint of a cross-like element; Fig. 1*A*) were unbiased, that is, their reports were symmetrically distributed around the stimulus midpoint (Fig. 1*B*, black curve). When 2 flankers were present, each with horizontal lines positioned at the midpoint (midflankers), psychometric functions become shallower (Fig. 1*B*, pale-gray curve), indicating that larger shifts were needed to report the position of the target-horizontal reliably (i.e., position discrimination thresholds were elevated). The midpoint of functions remained unchanged (i.e., centered target-horizontals still appeared centered when crowded). Shallow functions are also evident in conditions where flanking features were displaced above or below the midpoint (up- and down-flankers, red and blue curves in Fig. 1*B*, respectively) by a threshold-level displacement (i.e., a positional shift that was just noticeable for uncrowded elements). Furthermore, up-flankers also shift the entire curve leftward (indicating more 'up' responses), which results in a negative midpoint value as the subjective midpoint

(where observers were equally likely to say 'up' or 'down') is now located at a negative (downward) position on the cross. The opposite pattern occurred with down-flankers, where an increased probability of 'down' responses results in a rightward shift of the psychometric function.

These midpoint/bias values are plotted in Fig. 1*C*, highlighting the finding that flankers with both subthreshold- and threshold-level displacements below the stimulus midpoint produced an increase in 'downward' responses, and a corresponding positive shift in bias. This result demonstrates that judgments were increasingly made in a direction consistent with flanking structure, which is also true of the up-flanker conditions, where negative bias is evident for both displacements. Additionally, the bias induced by subthreshold feature displacements (0.5× threshold) demonstrates that crowding can alter position encoding even when the flankers are difficult to distinguish from those that do not produce bias (i.e., midflankers). That is, although offsets in the subthreshold flanking-horizontals could be correctly identified as offset from the midpoint at only 13.3% above chance (based on uncrowded performance, making the actual percentage lower because of the crowding of flankers by the target; 31), they produced a significant level of bias in observers' responses to the crowded target. Importantly, the

more commonly observed effects of crowding are also apparent in our data, with all configurations producing an equivalent level of threshold elevation (Fig. 1D). These effects were also apparent when observers judged the position of the vertical line (left/right), albeit with less induced bias and higher threshold elevation (supporting information and Fig. S1), which rules out the possibility that our results were caused by observers inappropriately using the flanker-horizontals as a coaligned reference. The effect of crowding on midpoints and thresholds also declined with increasing target-flanker separation according to the 'Bouma law' of spatial extent (1, 9), with interference zones of 4–6° (0.3–0.4× eccentricity; Fig. S2). This similarity in spatial extent is consistent with both effects arising from a common source related to crowding. In sum, our data show that the crowding of positional information involves both an elevation in thresholds for position discrimination and a systematic bias that shifts the position of target elements to resemble flanking features.

We next sought to determine the precise mechanism underlying these effects. As detailed above, 2 predictions can be made for the encoding of position under crowded conditions. First, substitution processes would result in either the features of flanking items, or flankers in their entirety, being substituted into the target (10, 12, 16) so that the perceived position of target features should alternate between the veridical location and that of the flanking elements, producing a bimodal distribution of perceived feature positions. Second, were averaging (29) to apply in the positional domain, the perceived position of target features should lie at the average of target and flanker feature positions. Although more complex averaging schemes could alter the precise weights assigned to target and flanker positions (5, 30), the perceived target positions should nonetheless show a unimodal distribution centered on an intermediate value. As both models could produce the bias in Fig. 1 (Fig. S3), a more precise examination of *perceived* feature positions is required to differentiate between these accounts.

In Experiment 2, the horizontal target line was again presented at a range of positions, although flanker-horizontals were now positioned *relative to the target* (when crowded), either above or below the target line at either 2 or 4 times the threshold for positional discrimination. This procedure allowed a continuous range of both target and flanker positions, reducing the likelihood that observers would preferentially indicate particular feature positions (as they might were specific positions more likely to be present). Following each trial, a response cross-element appeared near fixation and observers adjusted the horizontal-line position until it matched their percept of the target. Results are presented in Fig. 2A, pooled across observers. With single targets (*Left*), the perceived position of the horizontal feature was largely veridical (lying along the white line), although responses are clearly repelled away from the stimulus midpoint to more extreme positions (i.e., observers rarely reported the target as a '+'). This *midpoint repulsion* appears similar to subjects' general reluctance to report the reference in discrimination tasks (31, 32). In our experiment, the stimulus centroid (33, 34) is likely to have served as the reference. Importantly, because this repulsion occurs with uncrowded elements, it cannot be explained by decision processes such as the 'signed-max' rule that depend on stimulus set size (35), nor can it explain the effects of crowding. In the presence of flankers (*Center* and *Right* of Fig. 2A), responses were pulled toward that of the flanking features. Because the results were identical for both up-flankers and down-flankers, the latter results were inverted and pooled with those of the up-flankers to compute the data in Fig. 2A. With a small offset in flanking features (*Center*), responses fell both on and between the flanking and target positions, whereas with large offsets (*Right*) these responses clearly lie between the target and flanker positions. The latter is

consistent with observers relying on the *average* of the target and flanker feature positions, as opposed to the probabilistic substitution of flanker features. However, because of the effects of positional uncertainty and midpoint repulsion, we also sought to test this proposition formally using simulation.

The predictions of averaging and substitution models of crowding were compared using a 3-stage model of each process (depicted in Fig. 2B, see *Methods*). Both began by registering feature positions with added positional noise, reflecting raised positional uncertainty in the periphery (18–20). This stage was unaffected by crowding, consistent with the invariance of contrast-detection thresholds for crowded stimuli (9, 36). The second stage involved the computation of either a weighted average of target and flanker positions or a probabilistic substitution of flanker features into the target. Both the strength of weighting and the probability of substitution were free to vary. Finally, crowded target-position estimates were pushed away from the midpoint to simulate the observed midpoint repulsion. For each model, the best-fitting parameters were selected as those that minimized the least-squares error between the predicted responses and those plotted in Fig. 2A.

These simulated responses for both averaging and substitution are plotted in Fig. 2C. In the absence of flankers (*Left*), both models correctly predict a percept of the targets that is close to veridical, albeit subject to positional uncertainty and repulsion away from the stimulus midpoint. Under crowded conditions, the averaging model produces estimates that initially lie on a line running parallel between the target and flanker positions (see Fig. S4), with the final-stage repulsion then pushing these responses away from the midpoint. With a small feature offset (*Center*), the final predicted responses lie along the target and flanker lines, because of the repulsion of the averaged values. Larger offsets (*Right*) produce responses that are intermediate between the target and flanking positions, closely replicating the pattern in our data. By contrast, predictions from the substitution model were initially distributed bimodally between the target and flanker positions, before the midpoint repulsion in the final stage. Although this process produces a pattern that approximates our data with small flanker offsets (*Center*), the predicted values for large flanker offsets (*Right*) fall at considerably more extreme positions than those perceived by our observers. Substitution is thus inconsistent with the data in this condition, in comparison with the close match produced by a weighted average of the noisy target- and flanker-feature positions.

It is important to note that neither of these models adds any noise in the second (crowding) stage, each adding only bias from the flanking features. Indeed, the addition of a second noise parameter to these simulations did not produce a significant improvement in the predicted responses (see *Methods*). At first glance, this lack of additional noise would appear to conflict with our finding that crowding induces both bias and an elevation of thresholds (Fig. 1). However, model simulations of Experiment 1 demonstrate that threshold elevation does arise from the combination of a noisy representation of the target position with an equally noisy representation of the flanker feature positions (see Fig. S3). An averaging of target and flanker positions can thus account for both the bias and the threshold elevation observed in the perceived position of crowded target features.

Discussion

We have demonstrated that crowding has 2 main effects on the perceived position of object features. First, crowding elevates thresholds for the discrimination of feature positions, consistent with the performance impairments seen for vernier acuity (25, 26) and across a range of other modalities (3–6). Second, crowding produces a systematic shift in the perceived position of object features, bringing them into alignment with those of
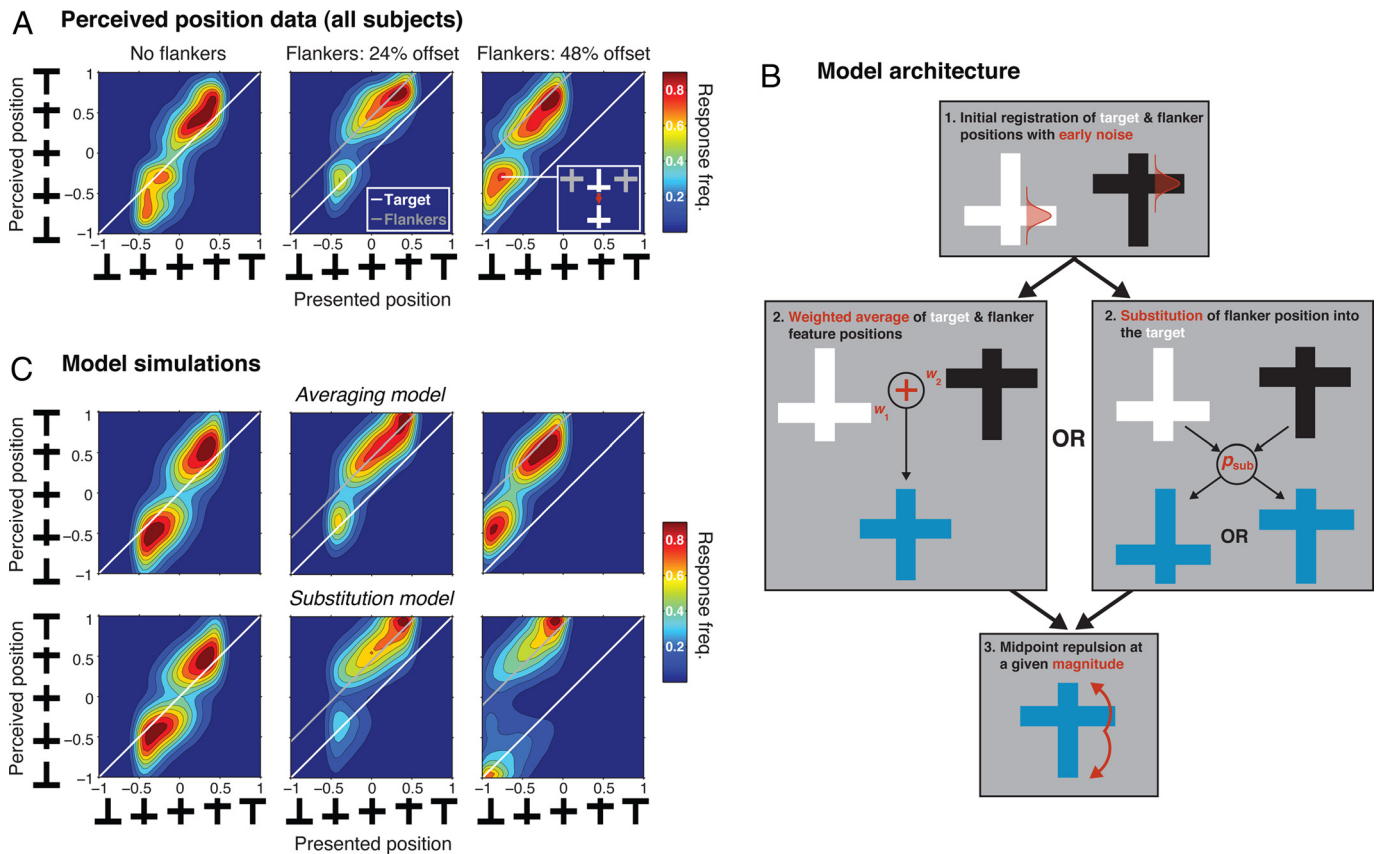
Fig. 2. The perceived position of horizontal line components in our stimuli, alongside model simulations. (*A*) Data for the perceived position task, plotted according to the presented position on the *x* axis and the perceived position on the *y* axis (as indicated schematically). The frequency of each response is shown using the associated color map. White lines in each plot indicate veridical target reports, with relevant flanker positions shown as a gray line. Data are shown for uncrowded stimuli (*Left*), and for crowded conditions with small (*Center*) and large (*Right*) positional offsets in the flanker-horizontals. Data have been pooled to be consistent with ''up–flankers'', with a schematic in the *Right* depicting the effects of crowding on perceived target position: the perceived position of target features is pulled toward that of the flankers. Averaging processes are most evident with large flanker offsets, where perceived positions cluster around values in-between the presented target positions and those of the flankers. (*B*) Two models of crowding. Both share an identical first stage, where the positions of target and flanker features (if present) are corrupted by Gaussian positional uncertainty (red distributions). In the averaging model, a second stage consists of a weighted average of target and flanker positions, although the substitution model involved a probabilistic switching between target and flanker positions. Both models converge on an identical final stage, where positional estimates are subject to midpoint repulsion. The 3 free parameters for each model are indicated in red. (*C*) Simulated responses from the 2 models, depicted in an identical fashion to those in panel *A*. Responses of the averaging model closely approximate those of our observers in all conditions, while the substitution model predicts perceived positions with large flanker offsets that are considerably more extreme than those reported.

flanking objects. When the perceived position of these crowded features was examined, in conjunction with computational modeling of these processes, the observed responses were predicted by a weighted average of position estimates, as opposed to the substitution processes more commonly discussed in the positional domain. This finding is broadly consistent with the averaging observed when oriented Gabor elements are crowded (29), although the weights incorporated into our averaging model can also account for the modulatory effects of target-flanker similarity (4, 9) and target-flanker separation (23). The necessary inclusion of midpoint repulsion in our model is also consistent with prior research suggesting that a simple averaging approach is unlikely to completely encapsulate crowding in other domains (5, 30).

Although substitution processes are typically invoked to explain phenomena such as the high proportion of flanker-related errors in crowding tasks (10–14), an averaging of feature positions between target and flanker objects would also give rise to these effects. This outcome is particularly likely given the nonlinear weighting in our model that can produce effects akin to substitution with an extreme weighting toward the flankers. These simulations also demonstrate the difficulty in separating

predictions from averaging and substitution models, with the combination of midpoint repulsion and high positional uncertainty producing identical predictions in the presence of flankers that differ from targets by small positional offsets. Both models made identical predictions for the results of Experiment 1, demonstrating that both a fine response scale and large positional offsets are required to disambiguate averaging and substitution accounts of crowding. This procedural issue could explain the lack of averaging observed with oriented Gabor stimuli using a coarse response scale (31), which would make the presence of near-averaged stimuli difficult to determine. It is only with large positional offsets and a fine response scale that the predictions from the substitution model can be shown to be untenable.

Importantly, neither of the models examined added any noise to the operations associated with crowding, each adding only bias from the flanking features on any given trial. Both the systematic bias and random effects associated with crowding could thus be attributed to the combination of a noisy representation of the target feature position with noisily represented flanker feature positions. In other words, the deleterious effects of crowding on position encoding can be accounted for by simple recourse to the

PSYCHOLOGY

inherent positional uncertainty in the peripheral visual field (18–20). This aspect of our model is similar to a recent model of feature binding (37), where the high positional noise in the periphery causes target features to be erroneously associated with adjacent elements within coarse feature maps. Additionally, because there are also intrinsic levels of uncertainty in other domains affected by crowding (3–6), such as orientation (38) and direction (39), our model raises the intriguing possibility that crowding could involve the averaging of features across all affected modalities, as opposed to the addition of noise suggested by the historic focus on error rates in crowding tasks.

As a whole, our results are most consistent with a preattentive basis for crowding. The observation that crowding can occur with flanker offsets below the threshold for positional discrimination (Fig. 1) in particular suggests that crowding does not require an awareness of flanking structure. Although we cannot rule out the possibility that this bias arose from the veridical flanker displacements reaching awareness on a small proportion of trials (e.g., 40), the nature of this bias is also inconsistent with the mechanisms typically proposed to link attention with crowding. Attentional theories proposing that crowding causes the loss of target identities from awareness (41) would predict random responses (i.e., threshold elevation; 42) without a bias toward flanker identities, whereas theories with an unfocused attentional spotlight that causes the substitution of flankers for the target (10, 11) would follow the failed predictions of our substitution model (Fig. 2). It is nonetheless conceivable that the attentional spotlight could perform a local averaging of crowded elements, and the observed averaging of feature positions could thus be attributed to either low-level or attention-based mechanisms.

Finally, the averaging processes observed in the present study appear remarkably similar to processes of contour integration (43), whereby spatially distinct elements are grouped into extended edges. In our stimuli in particular, averaging would have the effect of pulling the position of all feature elements in line with extended contours, consistent with a recent suggestion that crowding may underlie many reported phenomena in contour integration (44), including the insensitivity to curved contour structure in the periphery (45). The strong bias induced by collinear (Fig. 1C) compared with parallel features (Fig. S1) is also consistent with the tuning of contour integration for 'snakes' vs. 'ladders' (43). These interactions are likely to serve a beneficial role, with pooling along contours able to relieve the positional uncertainty inherent in the peripheral visual field (18). This finding raises the possibility that crowding processes may actively simplify image representation in the peripheral visual field by generating extended contours in an attempt to overcome both the low spatial sampling and high positional uncertainty within these regions.

Materials and Methods

Observers. Two of the authors (J.A.G. and S.C.D.) and 1 naïve observer (D.K.) participated in each experiment. All had normal or corrected-to-normal visual acuity.

Apparatus. Experiments were programmed using MATLAB (MathWorks, Ltd.) running on a Macintosh computer with PsychToolbox software (46). Stimuli were presented on a CRT monitor (LaCie Electron Blue) (22), with a resolution of 1152×870 pixels and a 75-Hz refresh rate. The monitor was calibrated using a photometer and linearized in software, giving a mean and maximum luminance of 50 and 100 cd/m^2, respectively. Stimuli were viewed with the dominant eye from 57 cm, with responses made using the keyboard (Experiment 1) or mouse (Experiment 2).

Stimuli and Procedures. Stimuli were white 'cross-like' elements (50% Weber contrast above mean luminance; Fig. 1A), with a stroke width equal to one-fifth the line length (equivalent to Sloan letters) and a total size set to twice the size-acuity thresholds for each observer. Targets were presented at 15 ° eccen-

tricity in the upper visual field, with a 0.7×0.7 ° cross presented at fixation. Under crowded conditions, 1 flanker was presented to the left and 1 to the right of targets. The centre-to-centre separation of the target and flankers was 2.5 ° in Experiment 1 [well within the tangential region of interference; (9, 23)], with additional target-flanker separations tested separately (Fig. S2). Test stimuli were presented for 200 ms, followed by a dense 20 ° × 5 ° masking array of cross stimuli with randomized feature positions for a further 200 ms.

Size-acuity thresholds were first estimated by having subjects judge the orientation (4AFC) of a cross-element configured to resemble a 'T' in 1 of 4 cardinal orientations. Stimulus size was determined using QUEST (47), which converged on 62.5% identification. This procedure was repeated 5 times for each observer, with the mean of these thresholds being 0.8 ° for both JAG and SCD, and 0.75 ° for DK. Stimuli were subsequently presented at twice these sizes.

Next, baseline thresholds for position discrimination were measured using a single cross. The x-position of the vertical line was fixed at the midpoint of the horizontal line (as in upright/inverted Ts), with the horizontal line of targets presented at 1 of 17 positions along the vertical line. Observers judged whether the presented position was above or below the stimulus midpoint (2AFC). Each position was repeated 10 times, with 5 blocks run according to the method of constant stimuli. Raw data were fit with cumulative Gaussian functions to determine the midpoint (50% upward) and threshold (75% upward) for each run. Thresholds for uncrowded position displacements were 0.14 ° for DK and SCD, and 0.18 ° for JAG. Midpoints did not differ significantly from 0, indicating a lack of bias in unflanked judgments of position.

To assess the effect of crowding on position encoding (Experiment 1), judgments of the vertical position of the horizontal line were made in the presence of 2 flanking elements. The horizontal lines of flankers were presented at 1 of 5 positions: the vertical midpoint of the stimulus (midflankers), and above or below the midpoint with either threshold or half-threshold displacements (up-flankers and down-flankers, respectively, using the uncrowded position thresholds obtained earlier). Unflanked thresholds were monitored throughout the experiment (allowing calculation of the frequency that subthreshold flanker displacements were likely to have been detected), resulting in 6 interleaved conditions. As before, the horizontal target line appeared at each position 10 times, with each block repeated 5 times. Data were again fit with cumulative Gaussian functions, with 95% confidence intervals determined for midpoints and thresholds using 1,000 repetitions of a bootstrap.

To determine the perceived position of target features (Experiment 2), horizontal target lines were again presented at a range of positions along the vertical line, either in isolation or crowded. Parameters in this experiment were equated across observers, with both positional displacements and stimulus size set to the maximum values of Experiment 1 (0.18 ° and 1.6 °, respectively) and the target-flanker separation decreased to 1.75 ° to maximize crowding. Under crowded conditions, flanking horizontal lines were positioned either above or below the target position by a small (2× threshold, 0.36 °) or large (4× threshold, 0.72 °) displacement. To avoid positions exceeding stimulus dimensions, target positions varied between ± 0.4 ° of the stimulus midpoint in uncrowded and small flanker-offset conditions, or were focused in either the upper or lower half of the stimulus (again spanning 0.8 °) with large flanker offsets, depending on the sign of offset. Following the mask, a response cross (identical to experimental stimuli) was presented 2 ° above fixation. Using the mouse, observers adjusted the horizontal line (initially randomly positioned) until it matched their percept of the target. Each run contained 3 presentations of each target position, with 21 positions per condition; 5 repetitions of each run were undertaken. Responses were smoothed with a Gaussian filter (σ = 1 pixel/0.04°, the difference between each presented position). As responses to upward and downward flankers were mirror images of one another, the latter were inverted and both were pooled. Data were also pooled across observers, as all showed an identical pattern of performance.

Model Simulations. Predicted perceived positions (Experiment 2) were simulated using 2 versions of a 3-stage model, 1 employing a weighted average of target/flanker positions and the other a substitution of flanker features (both depicted in Fig. 2B). Each had 3 free parameters. The first and third stages were identical for each model. The first stage consisted of noisy estimates of flanker and target positions:

$$y = s + a\sigma_e \quad \quad [1]$$

where s could represent s_t (the target-horizontal position), or s_{f1} and s_{f2} (flanker-horizontal positions), σ_e represents the Gaussian positional error and a is a free parameter—the magnitude of this error. Feature positions and Gaussian noise were clipped between ± 1, corresponding to the upper and lower extremes of the stimuli. When flankers were present, a second combi-

natorial stage was used, which could take 2 forms. In the weighted average model, combinations of the stimuli (y_c) were:

$$y_c = \frac{(y_t w_t + y_{f1} w_{f1} + y_{f2} w_{f2})}{(w_t + w_{f1} + w_{f2})} \qquad \textbf{[2a]}$$

where w_t, w_{f1} and w_{f2} are the weights for the target and flankers, respectively. These weights summed to 1, with the proportion of the target weight as the second free parameter. The result is an average of the target and flanker positions that can be biased toward or away from the target representation. In contrast, substitution was produced using a conditional probability:

$$\text{if } \{p_{sub} y_f > (1 - p_{sub}) y_t\} \text{ then } \{y_c = y_f\} \text{ else } \{y_c = y_t\} \qquad \textbf{[2b]}$$

Here, p_{sub} is the probability of flanker substitution, which is the second free parameter for the substitution model. Multiplication of this value with the strength of the target and flankers (i.e., their distance from the stimulus midpoint) determined the final report. The output of these 2 stages is shown in isolation in Fig. S4. Finally, the output of both models underwent reference repulsion: an inverse parabolic function pushed the estimated target position

away from the stimulus midpoint. This equation produced our final estimate of target position (y_{final}):

$$y_{final} = y_c \pm r(1 - y_c)^2 + \sigma_d \qquad \textbf{[3]}$$

Here, r is the magnitude of reference repulsion, the third free parameter. Additional decisional noise (σ_d) was also introduced that was not parameterized, as it does not alter the output of the 2 models qualitatively, but was set to be proportional to the degree of repulsion.

Data obtained in Experiment 2 was used to set the 3 free parameters for each model (a, r, and either w or p_{sub}). The final parameter set minimized the least-squares error between model predictions (1,000 trials per point) and psychophysical data. Simulations of the 2AFC data in Experiment 1 were also produced (Fig. S3). Finally, simulations were run with a fourth parameter, which set the magnitude of an additional Gaussian error term in the second stage of both models. The inclusion of this term did not significantly improve model fits in either case, as determined by the Akaike Information Criterion (48). Because in both cases the 3-parameter models produced a smaller AIC than the 4-parameter fits, the inclusion of this second noise term was deemed unnecessary.

1. Bouma H (1970) Interaction effects in parafoveal letter recognition. *Nature* 226:177–178.
2. Flom MC, Weymouth FW, Kahneman D (1963) Visual resolution and contour interaction. *J Opt Soc Am* 53:1026–1032.
3. Wilkinson F, Wilson HR, Ellemberg D (1997) Lateral interactions in peripherally viewed texture arrays. *J Opt Soc Am* 14:2057–2068.
4. Kooi FL, Toet A, Tripathy SP, Levi DM (1994) The effect of similarity and duration on spatial interaction in peripheral vision. *Spat Vis* 8:255–279.
5. Bex PJ, Dakin SC (2005) Spatial interference among moving targets. *Vision Res* 45:1385–1398.
6. Butler TW, Westheimer G (1978) Interference with stereoscopic acuity: Spatial, temporal, and disparity tuning. *Vision Res* 18:387–392.
7. Pelli DG, Tillman KA (2008) The uncrowded window of object recognition. *Nat Neurosci* 11:1129–1135.
8. Levi DM (2008) Crowding - An essential bottleneck for object recognition: A mini-review. *Vision Res* 48:635–654.
9. Pelli DG, Palomares M, Majaj NJ (2004) Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *J Vis* 4:1136–1169.
10. Strasburger H (2005) Unfocussed spatial attention underlies the crowding effect in indirect form vision. *J Vis* 5:1024–1037.
11. Huckauf A, Heller D (2002) What various kinds of errors tell us about lateral masking effects. *Vis cogn* 9:889–910.
12. Krumhansl CL, Thomas EAC (1977) Effect of level of confusability on reporting letters from briefly presented visual displays. *Percept Psychophys* 21:269–279.
13. Nandy AS, Tjan BS (2007) The nature of letter crowding as revealed by first- and second-order classification images. *J Vis* 7:5.1–26.
14. Strasburger H, Harvey LO, Rentschler I (1991) Contrast thresholds for identification of numeric characters in direct and eccentric view. *Percept Psychophys* 49:495–508.
15. Martelli M, Majaj NJ, Pelli DG (2005) Are faces processed like words? A diagnostic test for recognition by parts. *J Vis* 5:58–70.
16. Wolford G (1975) Perturbation model for letter identification. *Psychol Rev* 82:184–199.
17. Chastain G (1982) Feature mislocalizations and misjudgments of intercharacter distance. *Psychol Res* 44:51–65.
18. Levi DM, Klein SA (1986) Sampling in spatial vision. *Nature* 320:360–362.
19. Rentschler I, Treutwein B (1985) Loss of spatial phase relationships in extrafoveal vision. *Nature* 313:308–310.
20. Bennett PJ, Banks MS (1991) The effects of contrast, spatial scale, and orientation on foveal and peripheral phase discrimination. *Vision Res* 31:1759–1786.
21. White JM, Levi DM, Aitsebaomo AP (1992) Spatial localization without visual references. *Vision Res* 32:513–526.
22. van Opstal AJ, van Gisbergen JAM (1989) Scatter in the metrics of saccades and properties of the collicular motor map. *Vision Res* 29:1183–1196.
23. Toet A, Levi DM (1992) The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Res* 32:1349–1357.
24. Bex PJ, Dakin SC, Simmers AJ (2003) The shape and size of crowding for moving targets. *Vision Res* 43:2895–2904.
25. Levi DM, Klein SA, Aitsebaomo AP (1985) Vernier acuity, crowding and cortical magnification. *Vision Res* 25:963–977.
26. Westheimer G, Hauske G (1975) Temporal and spatial interference with vernier acuity. *Vision Res* 15:1137–1141.
27. Carney T, Klein SA (1999) Optimal spatial localization is limited by contrast sensitivity. *Vision Res* 39:503–511.
28. Hess RF, Barnes G, Dumoulin SO, Dakin SC (2003) How many positions can we perceptually encode, one or many? *Vision Res* 43:1575–1587.
29. Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M (2001) Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci* 4:739–744.
30. Solomon JA, Felisberti FM, Morgan MJ (2004) Crowding and the tilt illusion: Toward a unified account. *J Vis* 4:500–508.
31. Petrov Y, Popple AV (2007) Crowding is directed to the fovea and preserves only feature contrast. *J Vis* 7:8.1–9.
32. Stocker AA, Simoncelli EP (2008) A Bayesian model of conditioned perception. In *Advances in Neural Information Processing Systems*, eds Platt JC, Koller D, Singer Y, Roweis S (MIT Press, Cambridge, MA), Vol 20, pp 1409–1416.
33. Westheimer G, McKee SP (1977) Integration regions for visual hyperacuity. *Vision Res* 17:89–93.
34. Watt RJ, Morgan MJ (1983) Mechanisms responsible for the assessment of visual location: Theory and evidence. *Vision Res* 23:97–109.
35. Gheri C, Baldassi S (2008) Non-linear integration of crowded orientation signals. *Vision Res* 48:2352–2358.
36. Levi DM, Hariharan S, Klein SA (2002) Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *J Vis* 2:167–177.
37. Neri P, Levi DM (2006) Spatial resolution for feature binding is impaired in peripheral and amblyopic vision. *J Neurophysiol* 96:142–153.
38. Vandenbussche E, Vogels R, Orban GA (1986) Human orientation discrimination: Changes with eccentricity in normal and amblyopic vision. *Invest Ophthalmol Vis Sci* 27:237–245.
39. McKee SP, Nakayama K (1984) The detection of motion in the peripheral visual field. *Vision Res* 24:25–32.
40. Swets JA, Tanner WP, Birdsall TG (1961) Decision processes in perception. *Psychol Rev* 68:301–340.
41. He S, Cavanagh P, Intriligator J (1996) Attentional resolution and the locus of visual awareness. *Nature* 383:334–337.
42. Cass J, Bex PJ, Watt RJ, Dakin SC (2007) Equivalent noise reveals that visual crowding is not an attentional effect. *J Vis* 7:339a (abstr).
43. Field DJ, Hayes A, Hess RF (1993) Contour integration by the human visual system: Evidence for a local ''association field''. *Vision Res* 33:173–193.
44. May KA, Hess RF (2007) Ladder contours are undetectable in the periphery: A crowding effect? *J Vis* 7:9.1–15.
45. Hess RF, Dakin SC (1999) Contour integration in the peripheral field. *Vision Res* 39:947–959.
46. Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
47. Watson AB, Pelli DG (1983) QUEST: A Bayesian adaptive psychometric method. *Percept Psychophys* 33:113–120.
48. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19:716–723.

PSYCHOLOGY