

# SCIENTIFIC REPORTS



OPEN

## Visual crowding is a combination of an increase of positional uncertainty, source confusion, and featural averaging

William J. Harrison<sup>1,2</sup> & Peter J. Bex<sup>3</sup>

Received: 28 November 2016

Accepted: 28 February 2017

Published: 05 April 2017

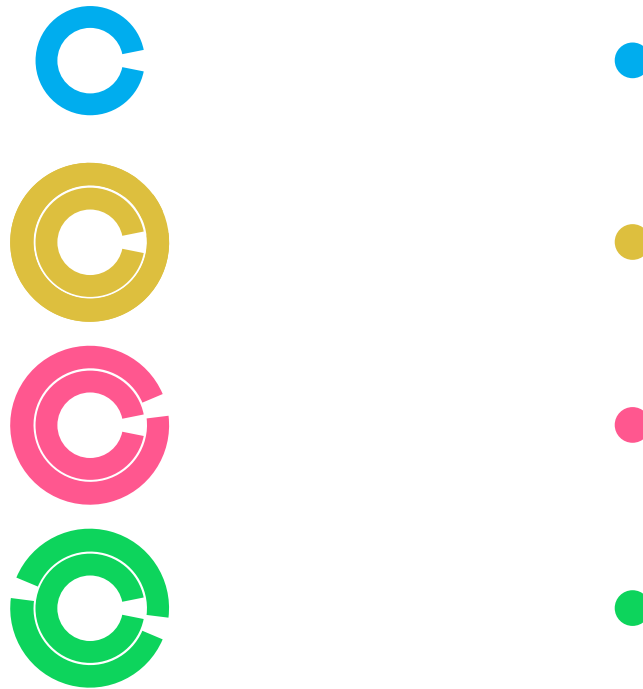
Although we perceive a richly detailed visual world, our ability to identify individual objects is severely limited in clutter, particularly in peripheral vision. Models of such “crowding” have generally been driven by the phenomenological misidentifications of crowded targets: using stimuli that do not easily combine to form a unique symbol (e.g. letters or objects), observers typically confuse the source of objects and report either the target or a distractor, but when continuous features are used (e.g. orientated gratings or line positions) observers report a feature somewhere between the target and distractor. To reconcile these accounts, we develop a hybrid method of adjustment that allows detailed analysis of these multiple error categories. Observers reported the orientation of a target, under several distractor conditions, by adjusting an identical foveal target. We apply new modelling to quantify whether perceptual reports show evidence of positional uncertainty, source confusion, and featural averaging on a trial-by-trial basis. Our results show that observers make a large proportion of source-confusion errors. However, our study also reveals the distribution of perceptual reports that underlie performance in this crowding task more generally: aggregate errors cannot be neatly labelled because they are heterogeneous and their structure depends on target-distractor distance.

Throughout the entire visual field, vision is constrained by multiple bottlenecks in visual processing that limit the information reaching our awareness. Initially, information is lost to physiological factors such as the eyes’ optics and retinal nerve fiber density, and neural selective sensitivity to spatio-temporal patterns<sup>1,2</sup>. However, our ability to identify even a simple object, such as a letter or an oriented grating, is far worse than predicted from these factors when the object is surrounded by clutter<sup>3,4</sup>. These identification failures, referred to as “crowding”, occur even though adaptation after-effects demonstrate that the object’s features have been encoded, at least in primary visual cortex<sup>5–7</sup>. Thus, our ability to consciously access the identity of an object is constrained by information processing capacity, not simply by retinal physiology or sensitivity limitations of the visual system.

Crowding, the inability to recognise an object in visual clutter, influences many aspects of vision. It is generally agreed to occur across the entire visual field<sup>8</sup>, although it is markedly more difficult to measure at the fovea<sup>9</sup>. As discussed in a review by Pelli and Tillman<sup>4</sup>, crowding affects all basic object recognition tasks, predicts reading speed and dyslexia, and is diagnostic of foveal deficits present in amblyopia<sup>10</sup>. Furthermore, it limits visibility of naturalistic images<sup>11–13</sup>, and interacts with saccadic and smooth pursuit eye movements in non-trivial ways<sup>14–19</sup> (but see ref. 20). There have been several recent reviews of crowding that summarize very well its ubiquity<sup>4,21–23</sup>, as well as examples in which object recognition is seemingly unaffected by parameters that cause crowding in other instances e.g. ref. 24.

The spatial extent of crowding is quite similar across paradigms. Perceptual errors increase with eccentricity and decrease as the distance between target and distractor increases. The precise target-flanker distance at which crowding is alleviated at a given eccentricity, often referred to as Bouma’s constant, is somewhat variable across studies<sup>8</sup>, and changes dynamically according to the duration and relative timing of target and distractors<sup>19,25–29</sup>. Despite this variability, the term “crowding” is typically taken to refer to any target identification interference that depends on target-distractor proximity<sup>30</sup>. However, there are also several studies demonstrating that a metric of

<sup>1</sup>Department of Psychology, University of Cambridge, Cambridge, UK. <sup>2</sup>Queensland Brain Institute, The University of Queensland, Brisbane, Australia. <sup>3</sup>Department of Psychology, Northeastern University, Boston, USA. Correspondence and requests for materials should be addressed to W.J.H. (email: willjharri@gmail.com)



**Figure 1.** Examples of stimuli in our experiments that produce qualitatively different perceptual outcomes. From top to bottom, rows represent the unflanked, no-gap flanker, one-gap flanker, and two-gap flanker conditions. Stimuli are drawn to scale, but were white on a grey background in the experiment. Data in later figures follow the colour conventions shown here.

target-distractor distance alone fails to predict target visibility. For example, the extent of crowding varies according to: colour, shape, or polarity differences between target and flankers<sup>31</sup>; the duration of the crowded display<sup>28</sup>; and perceptual grouping of the flankers<sup>32,33</sup>. To the best of our knowledge, no model has been advanced to attempt to account for all of these effects.

Although there are many fascinating aspects of crowding, we focus on fundamental findings. By viewing Fig. 1, the reader can experience firsthand three phenomenal aspects of crowding that arise with stimuli similar to those used in our experiments. Following standard convention, we term these phenomena: 1) positional uncertainty<sup>34,35</sup>, 2) feature averaging<sup>36</sup>, and 3) source confusion<sup>37,38</sup>. In this figure we present the same target stimulus, a modified Landolt C, in a series of distractor conditions. An observer's goal is to locate the orientation of the gap section as accurately as possible. If the reader fixates the spots in succession from top to bottom, they may note that the apparent clarity of the correspondingly-coloured target orientation is affected differently in each condition. The target gap is clearest in the top row, but its position is less clear when fixating the yellow spot below, perhaps because the solid ring distractor adds noise to the positional mechanisms encoding the target orientation<sup>39</sup>. When fixating the pink spot, the target and distractor gaps may perceptually blend together, shifting the perceived target orientation toward the distractor orientation e.g. ref. 40. When fixating the green spot, it is not immediately clear which of the multiple gaps is the target, and it may be easy to confuse a distractor gap for a target gap<sup>41</sup>. The changes in target visibility while viewing the yellow, pink, and green stimuli demonstrate, in order, positional uncertainty, feature averaging, and source confusion.

The perceptual phenomena experienced in cluttered displays are typically revealed across studies employing different methodologies. Changes in positional uncertainty and feature averaging are found in experiments in which an observer is required to make a spatial judgment about a continuous property of the target, such as its orientation or relative position, for examples, see refs 40 and 42. Feature substitutions – mistaking a distractor element for a target – are mostly found in paradigms in which the observer is required to report the categorical identity of target such as a letter; trials in which the observer reports a distractor identity instead of the target reveal source confusions<sup>37,38,43</sup>, which may be independent of an increase in positional uncertainty<sup>44–46</sup>.

It is important to note that distinct categories of errors, such as *averaging* and *substitution* errors, are descriptors of results, not descriptors of a mechanism per se. Indeed, even the term crowding refers to the result of some visual process and not a mechanism. The underlying cause of crowding has previously been explained by various computational models<sup>11,23,40–42,47–49</sup> and higher-level mechanistic hypotheses<sup>24,50</sup>. Population code models, in which all visual features probabilistically contribute to perceptual reports, can produce a wide variety of data<sup>51</sup>, including so-called averaging and substitution errors<sup>41</sup>. We have thus argued that the different classes of errors reported across the crowding literature are actually arbitrary categories of the output of a single mechanism. In the present report, therefore, we use the terms “substitution” and “averaging” as a convenient way to describe patterns in our data, but not to indicate hypothesised mechanisms. Our aim in the present study is to shed further light on the causes of crowding using a single paradigm that produces multiple perceptual phenomena. Here we

use experiment and modelling to quantify changes in positional uncertainty, averaging, and source confusions in visual clutter.

## Methods

This experiment accorded with the protocols reviewed and approved by the Northeastern University Institutional Review Board. We tested three highly experienced psychophysical observers, including the two authors, all of whom gave informed consent. In figures, we refer to the participant naïve to the specific purposes of this experiment as N1, and to the authors as A1 (PJB) and A2 (WJH). All observers previously participated in two similar crowding experiments<sup>41</sup>.

An observer sat 57 cm from the display with their head stabilised by a chin and headrest. The display was a CRT monitor (1280 × 1024 resolution, 85 Hz). We programmed the experiment with the Psychophysics Toolbox Version 3<sup>52,53</sup> in MATLAB (MathWorks). Stimuli were white (100 cd/m<sup>2</sup>) on a gray (50 cd/m<sup>2</sup>) background. The target was centered 10° to the right of the fixation spot, had a 2° diameter, and a line width of 0.4°. The gap width, measured at the midpoint of the line width, was 0.4°. In the one-gap and two-gap flanking conditions, the size of the gaps remained constant across flanker diameters. For all flanking conditions, the line width remained constant (0.4°). A flankers' outer edge was separated from the target's outer edge by 0.4°, 0.92°, 1.62°, 2.58°, 3.9°, or infinity (ie. no flanker). We express each target-flanker condition as the flanker radius as a proportion of the eccentricity of its centre, giving 0.14  $\phi$ , 0.19  $\phi$ , 0.26  $\phi$ , 0.36  $\phi$  and 0.49  $\phi$ , where  $\phi$  is the flanker eccentricity (10°). The flanker condition was selected randomly from trial-to-trial.

The orientations of flankers were constrained in the following way to produce maximal crowding effects<sup>41</sup>. In the one-gap and dual-gap flanker conditions, the flanker orientation was drawn from a normal distribution, centred on the target orientation and with a standard deviation of 22.5°. Within this range, we expect maximum levels of crowding. For the dual-gap flanker condition, a second flanker gap was drawn from a normal distribution centred 180° from the first flanker gap, with a standard deviation of 22.5°.

Each trial began when an observer pressed the space bar, which triggered the display of a small spot in the centre of the screen and the target (with or without flanks) for 500 ms. Immediately following the offset of the target, a Landolt C was presented in the centre of the display, and observers could rotate this clockwise or anti-clockwise by pressing the right or left arrow key respectively. To report an orientation, they pressed the space bar, and the next trial would begin. Observers were instructed to be as accurate as possible without rushing. They could take a break at any time by withholding a report. All observers completed 320 trials per session (20 repetitions of each target-flanker combination), for five sessions, giving a total of 1600 trials, or 100 trials per target-flanker condition. Each session took approximately 15–20 minutes.

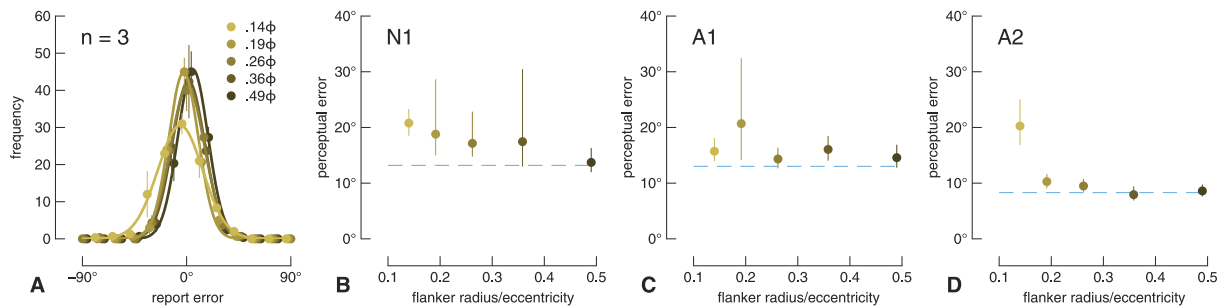
**Alternative analyses.** Prior to performing the main analyses for the one-gap and two-gap flanker conditions described in the Results, we performed three other analyses in order to categorise errors. First, we performed simple linear regression, as described in the Results below. Second, we applied a mixture model used in recent visual working memory papers<sup>54,55</sup>. This model uses maximum likelihood to estimate the proportion of trials in which an observer reported a target versus a distractor, as well as the precision with which reports are made. Third, we added two additional parameters to the mixture model in order to estimate reports corresponding to a weighted average of target and flanker<sup>40</sup>, in addition to pure target and substitution reports. Finally, we used a Monte Carlo analysis as described in Supplementary Figure A2. Before applying these analyses to our data, we tested them on simulated datasets with known distributions. When we simulated data consisting of reports corresponding to the target orientation and substitution errors corresponding to the flanker orientation (plus Gaussian noise; 100 trials as per the experiment), both the mixed modelling and Monte Carlo analyses returned approximately the correct underlying proportions of report types. However, the weighted average model was unreliable, likely due to the number of free parameters. The standard mixture model and Monte Carlo analyses estimating target and substitution reports are further outlined in Supplemental Figure A2.

## Results

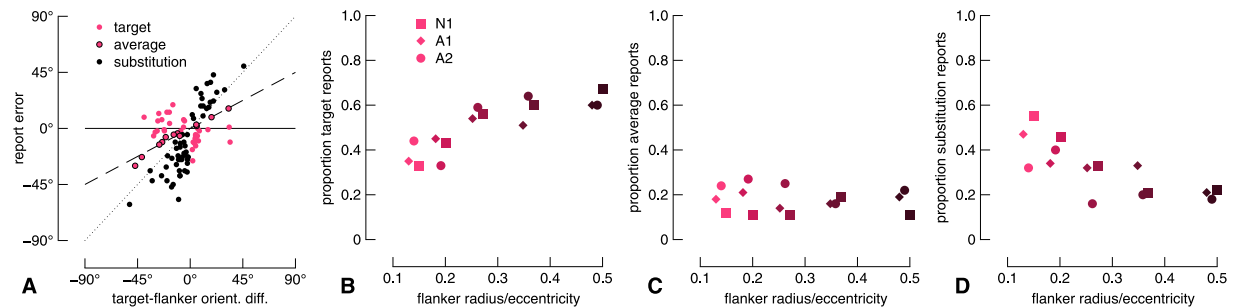
For all conditions, report error is defined as the difference between the reported orientation and the actual target orientation, with positive errors indicating a report that was more clockwise than the target. Because in our experiments increasing the flanker radius increases target-flanker separation, we use the terms “flanker radius” and “target-flanker separation” interchangeably. For consistency with our previous paper, figures show the flanker radius as a proportion of the eccentricity of the target center.

**No-gap flanker condition.** Because the no-gap flanker has no features that could be substituted or averaged with the target, results from this condition allow us to examine clearly changes in positional uncertainty. We rule out the contribution of more classic forms of masking in the Discussion. In the no-gap flanker condition, observers' report errors clustered around 0° for all target-flanker separations (Fig. 2A). We used the Circular Statistics Toolbox (<http://uk.mathworks.com/matlabcentral/fileexchange/10676-circular-statistics-toolbox-directional-statistics->) to find the circular standard deviation of observers' reports. We refer to this measure as perceptual error, which is plotted in degrees separately for each observer in Fig. 2B–D as a function of the flanker radius. Consistent with the crowding literature, all participants' perceptual error was higher than the unflanked condition (dashed lines) for the two smallest flanker radii. Observers N1 and A2 in particular show the characteristic relationship between performance and increasing flanker radius. Performance in the farthest flanker condition was similar to unflanked performance for all observers.

**One-gap flanker condition.** In the one-gap flanker condition, we have previously shown that observers' reports correspond to the distribution of weighted responses to the target and flanker orientations<sup>41</sup>. Thus, reports may include a proportion of responses at the target orientation, the flanker orientation and their mean



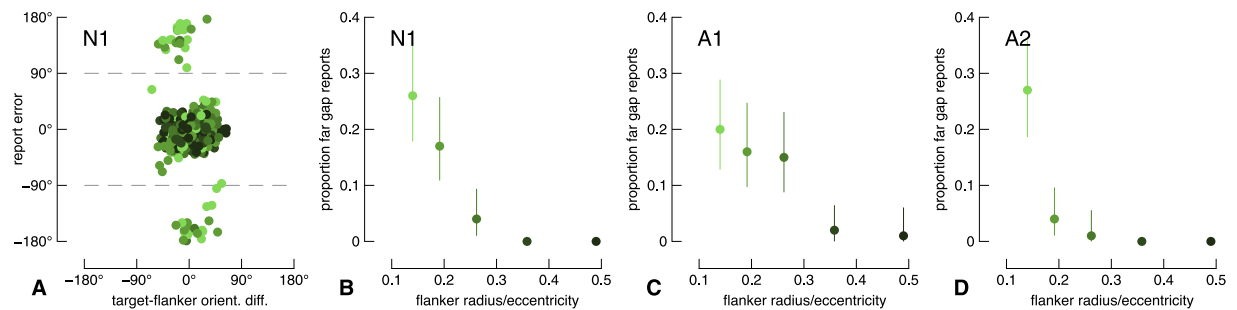
**Figure 2. Results from the no-gap flanker condition.** (A) Average frequencies of each report error size (25 bins; bin width = 15°), fit with circular distributions. To improve visibility, the data have been offset slightly along the x-axis, and the x-axis has been truncated to  $\pm 90^\circ$ . Only a total of 4 errors greater than  $\pm 90^\circ$  were made by all observers across all no-gap flanker conditions. The inset legend shows the target-flanker separation condition. Error bars show one standard error. (B–D) Perceptual error (the circular standard deviation of report errors) plotted as a function of flanker radius for each participant. Target-flanker separation is expressed as flanker radius divided by the eccentricity of its centre (i.e. in units of Bouma's constant). The dashed line indicates perceptual error for the unflanked condition. Error bars show 95% bootstrapped confidence intervals.



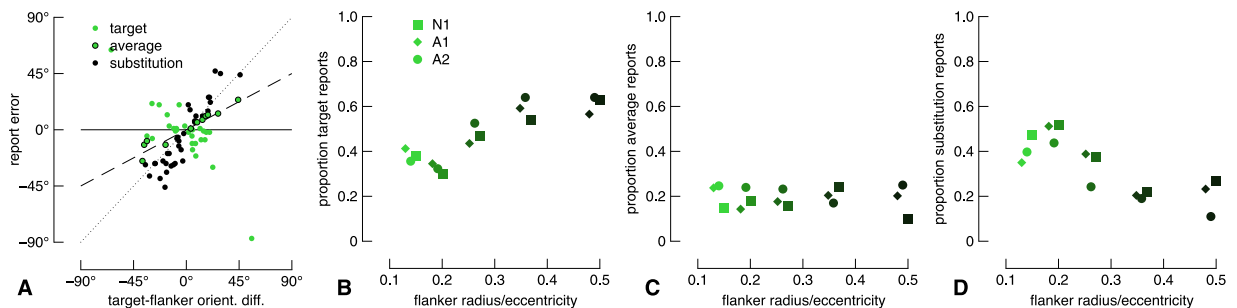
**Figure 3. Results from the one-gap flanker condition.** (A) Raw report errors from the naïve participant for the smallest flanker radius condition. Solid, dashed, and dotted lines show target, averaging, and substitution models, respectively. Data are attributed to each model according to proximity. (B–D) Proportion of report type as a function of flanker radius for all observers. Report type is shown on the ordinate across panels. Data for different observers have been horizontally offset slightly for clarity.

orientation. Quantifying reports with a unimodal circular standard deviation measure as used for the no-gap flanker condition above, is inappropriate in such a case. We express report errors as a function of target-flanker orientation difference (see Fig. 3A and Appendix Figure A1A). Under a first analysis (see Appendix A), we fit linear models to the data: reports at the target orientation have a slope of zero, reports at the flanker orientation fall on the diagonal with unity slope, and reports at the average orientation fall on the diagonal with a slope of 0.5. For all observers for the two most crowded conditions, the slopes were close to 0.5 and reduce to 0 at larger target-flanker separations. However, if these data were composed of noisy target and noisy flanker reports as described above and as has been argued previously<sup>45</sup>, the slope of a linear fit to crowded data may give a spurious interpretation favoring the averaging model. We further applied maximum likelihood mixture modelling as described by Bays *et al.*<sup>54</sup>, as well as Monte Carlo simulations, but these alternative analyses failed to return the true proportions of underlying report types of simulated data with known distributions (see Alternative analyses in Methods). We thus used a simplified approach that labels each datum according to its distance from each model prediction, as described below.

To quantify report errors in the one-gap flanker condition, we measured the distance of all report errors from each of three underlying model predictions that correspond to target reports, averaged reports, or substitution reports, and labelled each datum according to the nearest model (see Fig. 3A). We then quantified report types as a proportion of all trials from each condition. Note that this analysis assumes that three components are necessary to quantify crowding; in Supplementary Figure A2 we show the results from mixture modelling and Monte Carlo analyses assuming reports consist of only target and substitution errors e.g. ref. 45. These proportions are shown for all target-flanker separations in Fig. 3B–D with symbols indicating observers as per the legend. The ordinate labels on the different panels correspond to different model predictions. Note that this analysis fits data simultaneously to all model predictions, and so summing across panels for one flanker radius for a single observer gives 1. The pattern is very similar for all observers: with increasing flanker radius, the proportion of target reports increases, the proportion of average reports is relatively stable, and the proportion of substitution reports decreases. Note that this pattern of results indicates that the response distribution is multi-modal, since



**Figure 4. Report errors in the two-gap flanker condition.** (A) Raw report errors from the naïve participant. Report errors greater than  $90^\circ$  or less than  $-90^\circ$  (dashed lines) were classified as trials in which observers misreported the far-flanker gap rather than the target or near-flanker gap. (B–D) The proportion of such far gap reports as a function of flanker radius for all observers. Errors bars are 95% confidence intervals.

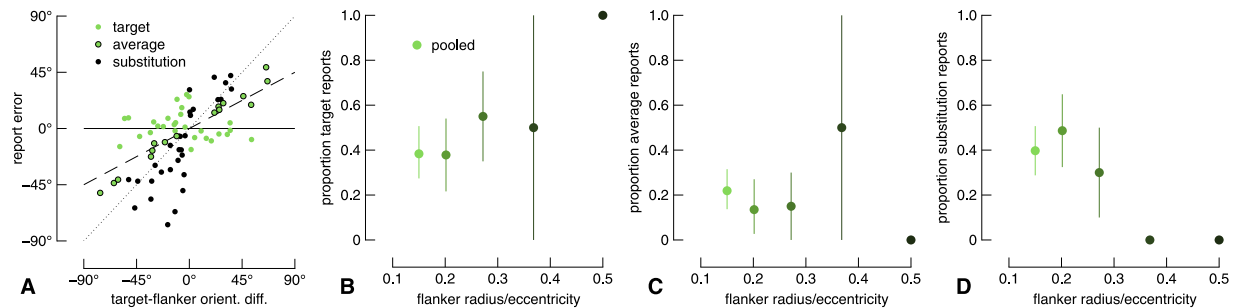


**Figure 5. Results from the two-gap flanker condition.** Data are shown as per Fig. 3. (A) Raw report errors from the naïve participant, focusing on only those errors within  $90^\circ$  of the target. (B–D) Proportion of report type as a function of flanker radius.

the proportions of data for each error type changes non-monotonically. Although the proportion of target reports saturates around 0.6 (Fig. 3B), this is likely an underestimate due to a limitation of our analysis for distributions when the response standard deviation is large and the target-flanker orientation difference is small: our simulations revealed that the modelled proportion of target reports is accurate when the proportion of other model components is high (greater than  $\sim 0.3$ ), but the proportion of target reports is underestimated when the contribution from other model components is minimal. For small flanker radii, the conditions under which we expect relatively poor performance, the proportion of each report type is likely more accurate than for larger flanker radii. Based on our previous work and the crowding literature, it is likely that observers' reports are barely, if at all, influenced by the flanker for the largest flanker condition (for example, see Fig. 2). Indeed, our alternative analyses shown in Supplementary Figure A2 show that the proportion of substitution errors is 0 for the largest flanker radius for all observers when assuming reports consist of only target and substitution reports. Note that mixture modelling<sup>54</sup> also estimates the perceptual error, which we show in Figure A3 in comparison with the no-gap flanker condition. When modelled in this way, perceptual error is vastly different across flanker conditions for observer A2, but fairly consistent for N1 and A1. Finally, mixture modelling provides an estimate of the random guess rate, which was most often 0, but never exceeded 7%, in line with standard lapse rates in psychophysical testing e.g. ref. 56.

**Two-gap flanker condition.** In the two-gap flanker condition, one flanker gap orientation was normally distributed around the target orientation (“near gap”; s.d. =  $22.5^\circ$ ) and the second flanker gap orientation was distributed  $180^\circ$  from the first flanker gap (“far gap”; s.d. =  $22.5^\circ$ ). Because of the relatively narrow report error distributions even in the presence of a single flanker gap (e.g. Fig. 3A), we can with some confidence delineate which errors are associated with the near gap and which with the far gap. In Fig. 4A, we show the raw report errors for the naïve observer. Report errors form two clusters: one cluster centred on the y-axis at approximately  $0^\circ$  and another at approximately  $\pm 180^\circ$ . We arbitrarily defined reports with an absolute error greater than  $90^\circ$  as far-gap reports. These reports are shown above and below the top and bottom dashed lines, respectively, of Fig. 4A. The proportion of far gap reports for each flanker radius and each observer are shown in Fig. 4B–D. For all observers, the proportion of far-gap flanker errors is greatest for the smallest flanker radius, and gradually decreases as the flanker radius increases.

We next divided the two-gap flanker condition data into two subsets for further analysis. First, we examined only those report errors within  $90^\circ$  of the target orientation. We performed the same analysis as we did for the one-gap flanker condition. Results are shown in Fig. 5, and are highly similar to the results from the one-gap



**Figure 6. Results from the two-gap flanker condition for report errors greater than 90° from the target with re-centered data (see text).** Data are shown as per Fig. 3. (A) Raw report errors pooled across observers. (B–D) Proportion of report type as a function of flanker radius. Error bars show 95% bootstrapped confidence intervals. Note that in the condition with the second largest flanker radius, there were only two trials. No error bars are shown for the largest flanker condition because there was only a single trial.

flanker condition in Fig. 3. We also performed simple linear fits to the raw data, the results of which (misleadingly) favour an averaging model (see Appendix Fig. A4).

Second, we performed the same modelling on report errors that were greater than 90° from the target orientation. Due to the small number of observations in this analysis (Fig. 4), we pooled observers' data. We re-centered this subset of data by subtracting 180° from the orientation difference between the target gap and far flanker gap, as well as from the report error. The pooled errors corresponding to the far flanker gap are shown in Fig. 6A. Because we re-centered these data, an error of 0° corresponds to a report of the target's polar opposite orientation, whereas data falling on the line of unity are reports following the far flanker gap orientation. As in the results above, with increasing flanker radius, the proportion of target reports increases and the proportion of substitution reports decreases. The proportion of average reports is relatively stable for the three smallest flanker radii, but the proportions vary greatly for the two largest flanker radii, likely because there were only four trials in total for these conditions. In contrast to the results above, the proportion of target reports at the largest flanker radius reached 1. However, there was only a single trial that had an error greater than 90° for the largest flanker condition, so this proportion necessarily had to be one or zero. Similarly, there were only two trials in the condition with the second largest flanker radius, greatly restricting the possible proportions of report types.

## Discussion

We used a method of adjustment to quantify perceptual error in peripheral vision under novel crowded conditions. In all conditions, performance depended on the distance between the target and flanker, in line with the vast crowding literature<sup>4</sup>. Based on the phenomenological responses and appearance of crowded stimuli, three general classes of mechanism have been advanced to account for crowding: 1) positional uncertainty<sup>34</sup>, 2) feature averaging<sup>36</sup>, and 3) source confusion<sup>37</sup>. These models are not necessarily mutually exclusive and image processing based approaches have been advanced that incorporate elements of each of these mechanisms (refs 11 and 13), but it has been difficult to reconcile which best accounts for the data because of the use of different methodologies and stimuli across studies supporting each account. Furthermore, with image processing based models that produce foveal performance deficits with synthetic images that simulate peripheral vision, it is not clear which combinations of these underlying processes accounts for perceptual performance<sup>11,57,58</sup>. The experimental design and complementary modelling employed here provides a novel way to classify the frequency of each error type with the same stimuli and observers. Our results reveal that all error types characterize crowding with Landolt Cs, but their proportions vary with the distance between the target and flanking stimuli.

In the no-gap flanker condition, the crowding we observed at relatively small target-flanker separations is likely caused exclusively by an increase in orientation uncertainty (for example, ref. 34). In this condition, there are no flank features for averaging or substitution to occur, yet, as shown in Fig. 2, we found a reliable increase in the circular standard deviation of perceptual errors with decreasing target-flanker separation, see also ref. 41. This result is unlikely due to a form of classical masking, such as meta-contrast masking, for two reasons. First, the target duration was much greater than required for such masking<sup>59</sup>. Second, meta-contrast masking increases random guessing<sup>60</sup>, but our observers' perceptual reports were not randomly distributed in the presence of close flankers (e.g. Fig. 2A). It is also difficult to account for these errors with an attentional account of crowding, in which observers' attentional resolution is too coarse to individuate the target gap<sup>5</sup>: even in the smallest flanker condition, observers' reports cluster around the actual target position, indicating that they could indeed attend to the target gap, albeit with greater perceptual error that is directly attributable to an increase in orientation variance (Fig. 2). We suggest that this orientation noise can be attributed to the solid flanking ring increasing the bandwidth of a population code that encodes the target orientation (see below and ref. 41).

The results of the one-gap and two-gap flanker conditions also support recent proposals that crowding may best be accounted for by a population code. Rather than conforming neatly to a single report error type, we found errors could be accurate, follow the flanker gap, or some average of the two. These data are thus difficult to reconcile with simple averaging or substitution models. Van den Berg *et al.*<sup>51</sup> showed that many hallmarks of crowding can be explained by a biologically inspired model that simulates the responses of populations of neurons tuned to orientation within a fixed region of space (ie. a receptive field). Crowded stimuli create systematic shifts in the

population code, so that when the population code is decoded, the decoded signal is prone to error. We further showed that an idealised population code can, for a given crowded stimulus, produce accurate, averaged, or substituted report errors in a probabilistic fashion<sup>41</sup>. We thus argued that there is no averaging or substitution mechanism per se, but instead that perceptual reports are drawn from the population response to the stimuli. Such a process is distinct from any single phenomenon such as source confusion, averaging or substitution but instead results from the broad spatial bandwidth of early stage filters. It is important to note that this model does not predict instances in which the magnitude of crowding is greatly modulated by certain configurations of flankers<sup>33,61</sup>. We are not convinced, however, that these limitations argue against a population code; they instead require reconsideration of the way in which the features are weighted within the population code. We are optimistic that a texture-processing model front-end may be sufficient to model configurational effects<sup>49</sup>, but further efforts are required.

Critically, the results from the present study show that averaging of features is not compulsory, in contrast to previous work<sup>36</sup>. On a number of trials in which a flanker gap is present, observers can recover the target orientation with a precision similar to that observed with no flanker gap (Fig. 3). With only a single report, it is impossible to know if, on a single trial, an observer perceived both the target orientation and the flanker orientation, their average, or if the frequencies of these percepts varied across trials. We previously asked participants to report both the target and flanker orientations in an experiment similar to the present report, and found that participants were generally capable of reporting both elements, though they often reversed feature positions<sup>41</sup> (see also ref. 61). Taken together, these data reveal that perceptual reports in clutter are probabilistic, but relatively fine detail can be recovered. Our results are in line with a conceptually similar earlier study by Pöder and Wagemans<sup>62</sup>, who had observers report a target Gabor defined by spatial frequency, colour, and orientation. Flankers were Gabors with various combinations of colours, spatial frequencies, and orientations, allowing the authors to quantify the extent and category of error types in crowding. In brief, they found observers often mistook an entire flanker for a target, or reported some flanker features as belonging to the target, as would be expected from frequent substitution errors as per our study.

It is clear from our data that observers often report an orientation closer to the flanker orientation instead of the target orientation (Fig. 4). However, our findings suggest the proportion of substitution-type errors is substantially greater than the proportion of substitution errors that occur when whole letter stimuli are used<sup>43</sup>. It is likely that this discrepancy can be explained by the task differences. Letter report paradigms limit the response range and so the observer is forced to select the most similar letter, even if the perceived stimulus does not match any of the possible responses. Our response method does not suffer from this limitation.

Since submission of this article, Agaoglu and Chung<sup>61</sup> published a study using a paradigm motivated by our earlier paper that investigates the validity of the present design. They levelled multiple criticisms of the continuous report approach and population response models of crowding. In short, they had two main aims: to test predictions made by an assumption of the population code model and to determine whether the relationship between the target's orientation and its eccentricity affected behavioural reports. Like previous crowding models, our model assumed that target and flanker orientations were weighted according to their mutual distance. This assumption predicts stronger crowding for flankers inside the target than outside. Agaoglu and Chung tested this prediction and found the opposite result, and therefore concluded that this choice of weighting field is incorrect. This striking observation challenges the many models that predict crowding based on the centre-to-centre spacing of target and flankers<sup>4</sup>. As discussed above, this finding does not negate the appropriateness of a population code model, but it does suggest that interference zones in crowding may be nonlinear. Agaoglu and Chung also found that observers' perceptual errors varied according to the absolute orientation of the target and flankers, consistent with eccentricity effects. We reported a similar finding in our original study, and measured these differences across a range of target-flanker separations to model the variable interference zone around the target (see Supplemental Fig. 1 of ref. 41). How these dependencies affect the present data requires future testing. However, they do not undermine the merit of continuous report designs nor do they render the paradigm inappropriate for measuring crowding. In fact, their data demonstrate how this paradigm can be used to provide a rich analysis of perceptual experiences under crowding. These observations raise many important questions about the data obtained in many previous studies that employed similar stimuli but different response methods and analyses.

Our results thus provide new evidence revealing that the component features of visual objects can be individuated even far in the periphery, although their relative positions and orientations may appear noisy and confusable across trials. The level of detail made available by the visual system has been heavily debated both within the crowding literature and more generally. Our data suggest one possible reason for this apparent conflict in the literature. Note that our analyses suggest all observers have very high rates of substitution-type reports under crowded conditions: combining the proportion of far-flanker gap substitutions (Fig. 4) and proportion of near-gap errors (Fig. 5), observers made approximately 50–70% substitution-type errors in the most crowded conditions. Such performance may be misinterpreted in alternative forced choice experiments (AFC), a more common psychophysical paradigm in which an observer is forced to categorise a target as being one of usually two to four alternatives. In these experiments, such a high proportion of substitution-type reports could render performance at or close to chance, leaving it unclear if a participant was randomly guessing, reporting an average stimulus or reporting what they thought was the target on some trials and the flanker on other trials. This is especially true in experiments with simple stimuli such as lines or Gabors, though it is less problematic with letter stimuli e.g. ref. 38, 43.

In conclusion, our findings show that relatively fine featural detail is not necessarily lost during early visual processing, but the precision of each perceptual report is corrupted. The aggregate of errors made when viewing crowded displays cannot be characterised as simply being accurate, averaged or substituted. The variety of report error types that occur within the same paradigm, as demonstrated here, provides a continued challenge to models of visual crowding.

## References

- Campbell, F. W. & Robson, J. G. Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology* **197**, 551–566 (1968).
- Robson, J. G. Spatial and Temporal Contrast-Sensitivity Functions of the Visual System. *Journal of the Optical Society of America* **56**, 1141–1142 (1966).
- Bouma, H. Interaction effects in parafoveal letter recognition. *Nature* **226**, 177–178 (1970).
- Pelli, D. G. & Tillman, K. A. The uncrowded window of object recognition. *Nature Neuroscience* **11**, 1129–1135 (2008).
- He, S., Cavanagh, P. & Intriligator, J. Attentional resolution and the locus of visual awareness. *Nature* **383**, 334–337 (1996).
- He, S. & MacLeod, D. I. Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature* **411**, 473–476 (2001).
- Shady, S., MacLeod, D. I. A. & Fisher, H. S. Adaptation from invisible flicker. *Proceedings of the National Academy of Sciences* **101**, 5170–5173 (2004).
- Whitney, D. & Levi, D. M. Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences* **15**, 160–168 (2011).
- Lev, M., Yehezkel, O. & Polat, U. Uncovering foveal crowding? *Scientific Reports* **4**, 4067 (2014).
- Song, S., Levi, D. M. & Pelli, D. G. A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *Journal of Vision* **14**, 3 (2014).
- Freeman, J. & Simoncelli, E. P. Metamers of the ventral stream. *Nature Neuroscience* **14**, 1195–1201 (2011).
- Wallis, T. S. A. & Bex, P. J. Image correlates of crowding in natural scenes. *Journal of Vision* **12**, 1–19 (2012).
- Balas, B., Nakano, L. & Rosenholtz, R. A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision* **9**, 1–18 (2009).
- Harrison, W. J., Mattingley, J. B. & Remington, R. W. Eye movement targets are released from visual crowding. *Journal of Neuroscience* **33**, 2927–2933 (2013).
- Harrison, W. J., Retell, J. D., Remington, R. W. & Mattingley, J. B. Visual crowding at a distance during predictive remapping. *Current Biology* **23**, 793–798 (2013).
- Wolfe, B. A. & Whitney, D. Facilitating recognition of crowded faces with presaccadic attention. *Front Hum Neurosci* **8**, 103 (2014).
- Lin, H. *et al.* Face Recognition Increases during Saccade Preparation. *PLoS ONE* **9**, e93112 (2014).
- Harrison, W. J., Remington, R. W. & Mattingley, J. B. Visual crowding is anisotropic along the horizontal meridian during smooth pursuit. *Journal of Vision* **14** (2014).
- Harrison, W. J. & Bex, P. J. Integrating retinotopic features in spatiotopic coordinates. *Journal of Neuroscience* **34**, 7351–7360 (2014).
- Ağaoğlu, M. N., Ögmen, H. & Chung, S. T. L. Unmasking saccadic uncrowding. *Vision Research* **127**, 152–164 (2016).
- Levi, D. M. Crowding—an essential bottleneck for object recognition: a mini-review. *Vision Research* **48**, 635–654 (2008).
- Strasburger, H., Rentschler, I. & Jüttner, M. Peripheral vision and pattern recognition: a review. *Journal of Vision* **11**, 13 (2011).
- Pelli, D. G. Crowding: a cortical constraint on object recognition. *Current Opinion in Neurobiology* **18**, 445–451 (2008).
- Herzog, M. H. & Manassi, M. Uncorking the bottleneck of crowding: a fresh look at object recognition. *Current Opinion in Behavioral Sciences* **1**, 86–93 (2015).
- Chung, S. T. L. Spatio-temporal properties of letter crowding. *Journal of Vision* **16**, 1–20 (2016).
- Huckauf, A. & Heller, D. On the relations between crowding and visual masking. *Percept Psychophys* **66**, 584–595 (2004).
- Ng, J. & Westheimer, G. Time course of masking in spatial resolution tasks. *Optometry and Vision Science* **79**, 98–102 (2002).
- Tripathy, S. P., Cavanagh, P. & Bedell, H. E. Large crowding zones in peripheral vision for briefly presented stimuli. *Journal of Vision* **14**, 1–11 (2014).
- Greenwood, J. A., Sayim, B. & Cavanagh, P. Crowding is reduced by onset transients in the target object (but not in the flankers). *Journal of Vision* **14**, 1–20 (2014).
- Pelli, D. G., Palomares, M. & Majaj, N. J. Crowding is unlike ordinary masking: distinguishing feature integration from detection. *Journal of Vision* **4**, 1136–1169 (2004).
- Kooi, F., Toet, A., Tripathy, S. & Levi, D. M. The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision* **8**, 255–279 (1994).
- Banks, W. P., Larson, D. W. & Prinzmetal, W. Asymmetry of visual interference. *Percept Psychophys* **25**, 447–456 (1979).
- Pachai, M. V., Doerig, A. C. & Herzog, M. H. How best to unify crowding? *Current Biology* **26**, R352–R353 (2016).
- Levi, D. M., Klein, S. A. & Yap, Y. L. Positional uncertainty in peripheral and amblyopic vision. *Vision Research* **27**, 581–597 (1987).
- Levi, D. M. & Klein, S. A. Sampling in spatial vision. *Nature* **320**, 360–362 (1986).
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. & Morgan, M. Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience* **4**, 739–744 (2001).
- Strasburger, H. & Malania, M. Source confusion is a major cause of crowding. *Journal of Vision* **13** (2013).
- Nandy, A. S. & Tjan, B. S. The nature of letter crowding as revealed by first- and second-order classification images. *Journal of Vision* **7**, 5.1–26 (2007).
- van den Berg, R., Johnson, A., Martinez Anton, A., Schepers, A. L. & Cornelissen, F. W. Comparing crowding in human and ideal observers. *Journal of Vision* **12** (2012).
- Greenwood, J. A., Bex, P. J. & Dakin, S. C. Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences* **106**, 13130–13135 (2009).
- Harrison, W. J. & Bex, P. J. A Unifying Model of Orientation Crowding in Peripheral Vision. *Current Biology* **25**, 3213–3219 (2015).
- Dakin, S. C., Cass, J., Greenwood, J. A. & Bex, P. J. Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *Journal of Vision* **10**, 14 (2010).
- Freeman, J., Chakravarthi, R. & Pelli, D. G. Substitution and pooling in crowding. *Attention, Perception & Psychophysics* **74**, 379–396 (2012).
- Greenwood, J. A., Bex, P. J. & Dakin, S. C. Crowding follows the binding of relative position and orientation. *Journal of Vision* **12**, (2012).
- Ester, E. F., Klee, D. & Awh, E. Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology: Human Perception and Performance* **40**, 1022–1033 (2014).
- Ester, E. F., Zilber, E. & Serences, J. T. Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *Journal of Vision* **15**, 1–12 (2015).
- Nandy, A. S. & Tjan, B. S. Saccade-confounded image statistics explain visual crowding. *Nature Neuroscience* **15**, 463–469 (2012).
- Chaney, W., Fischer, J. & Whitney, D. The hierarchical sparse selection model of visual crowding. *Front. Integr. Neurosci.* **8**, 73 (2014).
- Harrison, W. J. & Bex, P. J. Reply to Pachai *et al.* *Current Biology* **26**, R353–R354 (2016).
- Strasburger, H. Dancing letters and ticks that buzz around aimlessly: On the origin of crowding. *Perception* **1–13** (2014).
- van den Berg, R., Roerdink, J. B. T. M. & Cornelissen, F. W. A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology* **6**, e1000646 (2010).
- Brainard, D. H. The Psychophysics Toolbox. *Spatial Vision* **10**, 433–436 (1997).
- Pelli, D. G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision* **10**, 437–442 (1997).



54. Bays, P. M., Catalao, R. F. G. & Husain, M. The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision* **9**, 7.1–11 (2009).
55. Suchow, J. W., Brady, T. F., Fournie, D. & Alvarez, G. A. Modeling visual working memory with the MemToolbox. *Journal of Vision* **13** (2013).
56. Wichmann, F. A. & Hill, N. J. The psychometric function: I. Fitting, sampling, and goodness of fit. *Attention, Perception & Psychophysics* **63**, 1293–1313 (2001).
57. Wallis, T. S. A., Bethge, M. & Wichmann, F. A. Testing models of peripheral encoding using metamerism in an oddity paradigm. *Journal of Vision* **16**, 4 (2016).
58. Keshvari, S. & Rosenholtz, R. Pooling of continuous features provides a unifying account of crowding. *Journal of Vision* **16**, 39 (2016).
59. Alpern, M. Metacontrast. *Journal of the Optical Society of America* **43**, 648–657 (1953).
60. Agaoglu, S., Agaoglu, M. N., Breitmeyer, B. & Ogmen, H. A statistical perspective to visual masking. *Vision Research* **115**, 23–39 (2015).
61. Agaoglu, M. N. & Chung, S. T. L. Can (should) theories of crowding be unified? *Journal of Vision* **16**, 10 (2016).
62. Pöder, E. & Wagemans, J. Crowding with conjunctions of simple features. *Journal of Vision* **7**, 23.1–12 (2007).

## Acknowledgements

This work was supported by NIH grant R01EY021553 (P.J.B.) and a National Health and Medical Research Council of Australia CJ Martin Fellowship (APP1091257; W.J.H.).

## Author Contributions

Both authors designed the experiments and analyses. W.J.H. collected the data and wrote the manuscript. P.J.B. and W.J.H. analysed the data, and P.J.B. revised the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Harrison, W. J. and Bex, P. J. Visual crowding is a combination of an increase of positional uncertainty, source confusion, and featural averaging. *Sci. Rep.* **7**, 45551; doi: 10.1038/srep45551 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017