Dissociable effects of attention and crowding on orientation averaging

Steven C. Dakin
UCL Institute of Ophthalmology, University College London, London, UK
Peter J. Bex
UCL Institute of Ophthalmology, University College London, London, UK, & Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA
John R. Cass
UCL Institute of Ophthalmology, University College London, London, UK, & School of Psychology, University of Sydney, Sydney, NSW, Australia
Roger J. Watt
Department of Psychology, University of Stirling, Stirling, UK

It has been proposed that visual crowding—the breakdown in recognition that occurs when objects are presented in cluttered scenes—reflects a limit imposed by visual attention. We examined this idea in the context of an orientation averaging task, having subjects judge the mean orientation of a set of oriented signal elements either in isolation, or “crowded” by nearby randomly oriented elements. In some conditions, subjects also had to perform an attentionally demanding secondary task. By measuring performance at different levels of signal orientation variability, we show that crowding increases subjects’ local uncertainty (about the orientation of individual elements) but that diverting attention reduces their global efficiency (the effective number of elements they can average over). Furthermore, performance with the same stimulus-sequence, presented multiple times, reveals that crowding does not induce more stimulus-independent variability (as would be predicted by some accounts based on attention). We conclude that crowding and attentional load have dissociable perceptual consequences for orientation averaging, suggesting distinct neural mechanisms for both. For the task we examined, attention can modulate the effects of crowding by changing the efficiency with which information is analyzed by the visual system but since crowding changes local uncertainty, not efficiency, crowding does not reflect an attentional limit.

Keywords: crowding, attention, orientation, context


Introduction

Natural scenes are rich sources of visual information. In this context, the human visual system must filter out irrelevant structure and focus on information that can usefully guide behavior. One way we do this is by moving our head and eyes so that an area of interest falls within the central visual field, where visual spatial resolution is highest, engaging a correspondingly higher proportion of cortical resources. Consequently, when we fixate objects we generally have little difficulty segmenting the object from its surroundings, and, if the object is familiar to us, recognizing it. The process by which the brain achieves this segmentation and recognition is complex but examining conditions under which it fails may serve to reveal how it works. In particular, we are known to be poorer at recognizing isolated objects, such as letters, when they are presented in the peripheral visual field. The principle cause of this is that fewer neurons (at all post-receptoral stages of the visual hierarchy) are devoted to representing the periphery, which in turn leads to lower spatial resolution. Consequently, one should generally be able to overcome this reduction in performance simply by making the object bigger, and indeed one can predict the amount of scaling required based simply on the reduced numbers of cortical neurons representing the periphery (the cortical magnification factor). While this approach works for simple objects (e.g. letters) presented in isolation (Martelli, Majaj, & Pelli, 2005), the same scaling does not work when the objects are surrounded by irrelevant distracters. Such clutter makes recognition so bad that scaling of the...
object cannot overcome its effects. The deleterious influence of irrelevant nearby distractors on target identification is known as visual crowding. It is crowding and not detection that limits our performance in the peripheral visual field, as illustrated by Figure 1.

Much previous research has focused on determining the spatial determinants of crowding. The critical spacing (the smallest target-distractor distance inducing no crowding) is about half the target eccentricity (Bouma, 1970), a result that is remarkably robust across many different experiments (Pelli & Tillman, 2008). Toet and Levi (1992) used letters presented in the periphery to map “interference zones”—the region around the target within which a distractor letter would interfere with target recognition. They showed that interference zones scaled with eccentricity (confirming earlier observations by Bouma, 1970) and also were elongated radially away from fixation. Furthermore these zones are asymmetrical with the more eccentric flanker being more disruptive than the flanker closest to fixation (Bex, Dakin, & Simmers, 2003; Bouma, 1970; Chastain, 1982; Petrov & Popple, 2007). Many of these results can be accommodated by the view that the critical distance for crowding is a fixed cortical distance (Pelli, 2008).

Recently there has been interest in looking at what information is preserved when objects crowd one another. Parkes, Lund, Angelucci, Solomon, and Morgan (2001) showed that crowding did not prevent observers from accessing the average orientation of a cluster of oriented elements. Baldassi, Megna, and Burr (2006) had subjects make an estimate of perceived orientation and concluded that a “signed-Max” rule captured their performance. It is worth noting that although these experiments have been interpreted as indicating that crowding compels the observer to prepare what is essentially a statistical representation of local image structure, i.e. to treat it as texture, what they actually show is that statistical descriptions remain available, not that the preparation of such a description is the cause of crowding. There is an important distinction here between voluntary averaging—our ability to judge the mean attribute of a number of distinct elements (e.g. orientation, Dakin & Watt, 1997)—and involuntary averaging—an unavoidable (frequently undesirable) pooling of stimulus attributes. Parkes et al. (2001) showed that observers’ report of an individual crowded orientation was consistent with involuntary averaging which they inferred must be involved in crowding. Below we describe a noise paradigm that uses a voluntary averaging task to quantify the contribution of involuntary averaging under crowding.

As well as a voluntary averaging of orientation, visual adaptation persists under crowding (He, Cavanagh, & Intriligator, 1996). Generally, prolonged viewing of an oriented grating increases the detection threshold for a subsequently viewed grating at the adapting orientation. He et al. (1996) showed that this adaptation persists even when the target grating is crowded so that the subject is unaware of its orientation. Because adaptation has generally been linked to the operation of neurons in the primary visual cortex, the authors conclude that the site of crowding follows processing in V1 and is therefore, likely linked to attention. However, recently Blake, Tadin, Sobel, Raissian, and Chong (2006) have shown that because adaptation saturates with high contrast adapting patterns, use of such adapters by He et al. (1996) could have masked the effects of crowding. By using lower contrast adaptors Blake et al. (2006) were able to measure robust effects of crowding on adaptation. Therefore it would appear one can neither discount V1 from the neural

![Figure 1](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933556/)

Figure 1. While fixating the red marker in (a), details of individual birds are difficult to make out. Now try fixating the marker in (b) and attending to a pair of the same birds presented in isolation. Much more information about the birds should now be apparent. This disruptive effect of “clutter” on object recognition in the periphery is known as visual crowding.
circuit supporting awareness, nor from the mechanism producing crowding.

He et al.’s (1996) proposal that attention plays a central role in crowding is not new, although the evidence bearing on it is equivocal. There is little influence of pre-cuing target location on subjects’ ability to identify the orientation of Landolt C’s (Nazir, 1992) or of tilted Gabor elements (Wilkinson, Wilson, & Ellemberg, 1997). However pre-cuing the location of a target letter can relieve the effects of crowding in a moving display (Cavanagh & Holcombe, 2007) and in a change-detection task (Freeman & Pelli, 2007). Intriligator and Cavanagh (2001) measured critical dot spacing for subjects’ ability to localize dot targets (using a stepping and tracking task). Their results showed many of the typical features of “interference zones” for crowding (e.g., asymmetries between upper and lower visual fields, radial organization with respect to fixation) which they interpreted as the result of the limited spatial resolution of attention, which is coarser than visual acuity. Freeman and Pelli (2007) criticize such approaches since they are predicated on the notion that performance on such tasks is strictly “attention-limited” with position (compared to other features such as orientation) being “a more reliable source of information… immune to feature degradation.” Similarly, Strasburger and colleagues (Strasburger, 2005; Strasburger, Harvey, & Rentschler, 1991) have shown when subjects incorrectly report the identity of a flanked alphanumeric target, their responses are correlated with the identity of the flanking character. The authors interpret this flanker-substitution as a failure of positional encoding arising from spatially limited attentional resolution. However, recently we have presented evidence that such flanker substitutions might not arise from a failure to encode the position of whole letters but rather from a low-level weighted average of feature-position (Greenwood, Bex, & Dakin, 2009) with no need for an attentional component.

It is implicit in a theory of crowding based on attention that crowded displays effectively impose additional perceptual load, which observers have insufficient attentional resources to deal with. What is already known about the interaction between attentional capacity and perceptual load? Lavie and Tsal (1994) proposed that all of the attentional capacity is always allocated within a visual search paradigm. Lavie (1995) used an elegant response-competition paradigm to show the highly counter-intuitive finding that, in low load conditions, surplus attentional resources are allocated to the distractors. Under high load, no extra capacity is available for the distractors. This idea has received considerable support from fMRI. For example Rees, Frith, and Lavie (1997) had subjects monitor the case (low load) or number of syllables (high load) of a word in the presence of irrelevant background motion. They report that attentional load prevents motion selective areas of the brain responding to the irrelevant motion. If surplus attentional capacity is being used to process distractors, might this have consequences for crowding? If crowding and attentional load are tapping in to a common resource, then this would lead to the prediction that inducing perceptual load, to engage surplus attention, could decrease crowding.

On the assumption that less visible stimuli impose a higher perceptual load than more visible stimuli, Tripathy and Cavanagh (2002) presented evidence contrary to this view. They showed that manipulating presentation time (while equating visibility) could influence the size of the interference zones. However, the idea that reducing visibility is the same as increasing attentional load is contentious. It is already known that reducing visibility does not decrease target interference (Lavie & de Fockert, 2003) leading these authors to differentiate between limits of sensory-data and attentional-resource. With this in mind we sought a more direct measure of the influence of attentional load on crowding.

To be specific, in this paper we seek to determine if the behavioral consequences of crowding and attention are dissociable; i.e., for a given task will crowding produce a pattern of behavior that could reliably be distinguished from a pattern produced by attentional load and what can we infer about mechanism from such a finding? It is important that we be careful about the term dissociation. In the neuropsychology literature, dissociations are used to identify the neural substrate of a particular brain function. For an experiment with two different manipulations (F and G) and two different measures (X and Y), a “double dissociation” is said to have arisen when for example, F influences X but not Y, and G influences Y but not X. In this case X and Y must be supported by independent brain functions (Shallice, 1988). Sternberg (2003) considers how a similar logic can be applied to decompose a given process into its component sub-processes via what he terms a “Within-task double dissociation.” He considers two scenarios (each involving two experimental manipulations as in the neuropsychology case). The first is where the experimenter has access to a single experimental measure receiving contributions from two sub-processes (a “composite” measure). Double-dissociations from such experiments must be inferred and rely on bridging assumptions about how the sub-processes combine. The second scenario involves two experimental measures from a single task. If one observes that (for example) the first experimental measure depends only the first manipulation, and the second measure only on the second manipulation then one has achieved what (Sternberg, 2003) terms a “Within-task double dissociation: Pure measures.” In this case one can infer both that the sub-processes supporting each measure are independent and both are required to perform the task. It is important to note that this dissociation methodology is firmly bound to a particular task. Successfully decomposing a process into sub-processes using a dissociation (e.g., separating the contributions from crowding and averaging to a given task) in no way implies that these processes might not show dependencies on one another for another task. However showing that they can be independent does show, more generally, that they cannot be the same thing. In this paper
we apply a noise analysis, which yields two experimental measures from a single task (orientation averaging), to separate the contribution of two experimental manipulations: crowding and attentional load.

Methodology

In this paper we employ a noise analysis to both address the role of attention in crowding, and to study the specific mechanisms involved in crowding. This method (Barlow, 1956; Pelli & Farell, 1999) has been widely used elsewhere to study integrative processes in blur, spatial disorder, contour integration, etc. The version we use is based on earlier applications looking at the integration of orientation (Dakin, 1999, 2001a, 2001b) and motion direction (Dakin, Mareschal, & Bex, 2005) and is illustrated in Figure 2. Subjects are presented with a number of elements whose orientations are drawn from a Gaussian probability density function (Figure 2a). Subjects are required to judge if the overall orientation of the stimulus is clockwise or anti-clockwise of some reference orientation. We suppose that their strategy for performing this task is based on the estimated orientation of a subset of the elements present, so that performance is limited only by (a) the global sample size and (b) the precision of each local estimate. By estimating subjects’ performance at various offsets of the mean orientation (Figure 2b) one can estimate the orientation discrimination threshold, the smallest offset in mean orientation that subjects can reliably discriminate. Noise experiments estimate such thresholds at various levels of orientation variability (Figure 2c). Plotting threshold as a function of the range of orientations present (gray symbols, Figure 2c) it is evident that observers’ performance is good when orientation variability is low and deteriorates as it increases. Because we are estimating response variability as a function of stimulus variability, noise analysis exploits additivity of variance under convolution to model the data (boxed equation) in terms of external noise (the orientation variability; $\sigma_{\text{ext}}$), and local and global limits on integration. The solid line in Figure 2c shows the noise model fit to the data shown. It reveals that, in terms of the averaging model, presented inset in Figure 2a, the subject is averaging a global pool of 15 elements with each local sample having a precision (s.d.) of 5.5 deg. Although there is good evidence that this averaging strategy is close to the one observers employ rather than e.g. relying on the biggest single orientation offset (Dakin, 2001a, 2001b; Dakin & Watt, 1997) it is important to note that these estimates are not bound to a particular underlying model of performance. Noise analysis unambiguously indicates that the subject is performing as though they are averaging so many elements, each with a particular local precision. This point is key. For a stimulus containing $n$

![Figure 2](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933556/)
elements the effective sample size tends to be around $\sqrt{n}$ (Dakin, 2001a, 2001b) but this does not necessarily mean that some elements are being pooled and the remainder ignored. We have shown that, for moving patterns (Dakin et al., 2005), this dependence of sample size on $n$ emerges naturally from a direction integration system (using either vector averaging or maximum likelihood estimation) operating on a population of neural responses that are corrupted by multiplicative (Poisson) noise. Although the model integrates all the moving elements, increasing $n$ serves to drive up the neural response, which helps in overcoming noise and leads to a higher effective sampling rate. We will return to the notion that increases in effective sampling might be equivalent to a change in the gain of a noisy neural system below. With that proviso in mind, noise analysis is still unique in being able to separate local and global aspects of discrimination tasks. Here we use it to quantify the influence of attention and crowding on local and global orientation processing: do they limit out ability to see individual elements, or compromise our ability to voluntarily pool elements across space?

It is important at this juncture to highlight the limitations of the “averaging with noise” paradigm for exploring interactions between crowding and attention. As we alluded to above, our results can speak to such interactions only in the context of global orientation averaging. To anticipate our findings, we show that attention determines the efficiency of averaging (the effective number of elements averaged) and crowding determines how well the orientation of any one element is registered. We cannot know if, for example, crowding and attention are being evenly distributed across all elements: it could be that attention causes subjects to ignore whole elements. Since we do not know if attention is being deployed to process every element, a finding of independence in no way implies that attention and crowding are not interacting at the level of individual elements.\(^1\) What the paradigm does allow us to determine is what aspects of averaging are affected by crowding and attention, and by extension whether they share a common mechanism for the task at hand. We return to this point in the Discussion.

**Methods**

**Apparatus**

Experiments were run under the MATLAB programming environment (MathWorks Ltd) using software from the PsychToolbox (Brainard, 1997). Stimuli were presented on a CRT monitor (LaCie Electron Blue 22) fitted with a Bits++ box (Cambridge Research Systems) operating in Mono++ mode to give true 14-bit contrast accuracy. The display was calibrated with a Minolta LS110 photometer, then linearized in software using a look-up table, and had a mean (background) and maximum luminance of 50 and 100 cd/m$^2$ respectively.

**Subjects**

Three of the authors (wearing optical correction as necessary) and one subject naïve to the purpose of the experiment, served as observers. The authors are experienced in psychophysical tasks involving crowding and orientation estimation. All subjects conducted practice runs until their performance reached asymptotic levels.

**Stimuli**

Stimuli (illustrated in Figure 3) consisted of a set of orientationally and spatially band-pass filtered noise patches. They were generated by applying a log-Gabor filter (Field, 1987) to patches of white noise. The log-Gabor is defined in the frequency domain as:

$$G(f, \theta) = \exp\left(-\frac{\ln^2(f/f_0)}{2\ln^2(\sigma_f/f_0)}\right) \exp\left(-\frac{(\theta - \theta_0)}{2\sigma_\theta}\right),$$

where $f_0$ is the center spatial frequency of the filter (2.5 c/deg.), $\sigma_f$ is its spatial frequency bandwidth (0.5 octaves), $\theta_0$ is the center orientation of the filter and $\sigma_\theta$ is the orientation bandwidth (10°). The result of filtering a white noise patch was windowed using a raised two-dimensional cosinusoid (total radius, 0.5 deg.; peak-to-trough distance, 0.27 deg.) and its Michelson contrast set to 90%. Six of these target patches were presented on an iso-eccentric arc so that they fell 8 deg. from fixation in the upper visual field, and were separated from one another by 5 deg. of visual angle. We determined in pilot studies, that this separation produced no crowding between (unflanked) patches. For the noise condition, directions were drawn from normal distributions (wrapped at 180°) with standard deviations of 0.5°, 4°, and 32°. In crowded conditions (Figures 3c and 3d) an additional six patches were presented on an “inner” iso-eccentric arc 6 deg. from fixation so the center of each flank fell 2 deg. from its corresponding target.

**Procedure**

The subjects’ task was to judge whether the overall tilt of a collection of static oriented elements was clockwise or counter-clockwise of an internal vertical standard. Stimuli were presented for 150 ms and auditory feedback was given for incorrect responses. In the first experiment, the mean orientation of the stimulus from trial-to-trial was determined using a method of constant stimuli. This
procedure yields the probability that the subjects will report “clockwise” as a function of the mean orientation (e.g. Figure 2b). The range of mean orientation offsets (cues) tested was similar, for a given level of orientation variability, across various conditions of attentional-load and crowding in order to avoid any biasing of results. The cue-range tested was determined by pilot runs using an adaptive staircase procedure (Watson & Pelli, 1983). At least 3 blocks of 272 trials were undertaken for each subject in all conditions. In the second experiment we ran a similar experiment at a fixed orientation offset.

**Attentional load task**

In all conditions, the fixation mark consisted of a sequence of tumbling black T’s, constrained to contain no repeated orientations, running at ≈15 frames per second, and containing an “odd-man-out” white-element. A new sequence was generated every trial. The time when the odd-man-out appeared within the sequence was also randomized from trial to trial but was constrained to fall within 100 ms of the stimulus onset. In conditions of high attentional load we asked subjects to make a judgment about the sequence of tumbling T’s: to identify (4AFC) the orientation of the odd-man-out ‘T’ (“north,” “south,” “east” or “west”). To modulate task difficulty the entire sequence was fixed at 10% Michelson contrast, embedded in white noise, whose contrast was under control of an adaptive staircase (Watson & Pelli, 1983). This procedure, which ran independently from the staircase that controlled target orientation offset, converged on a T-target white noise mask contrast that elicited 75% correct identification. By maintaining task difficulty we ensured a constant level of attentional allocation at the central fixation location. The tumbling T sequence was always present, the only difference between high-load and low-load conditions being that in the former, subjects had to make the T-identification (which they were instructed to make their primary task) before responding to the orientation of the oriented elements. In low-load conditions the T white-noise mask contrast was fixed at 6%.

**Data fitting**

Raw data were fit with wrapped cumulative Gaussian functions using a bootstrapping procedure to derive the psychometric function (fit line in Figure 2b). Error bars on all data graphs indicate 95% confidence intervals on the estimated slope (threshold) of these fits. Thresholds were defined as the standard deviation parameter of the cumulative Gaussian function that best fit the data. Thresholds were measured as a function of the orientation variability of the signal, and were fit with the noise model illustrated in Figure 2c (again using a bootstrapping procedure). Using the same raw trial-by-trial response data used to bootstrap the psychometric functions, we repeatedly generated new data sets, fit each one using the noise model, and recorded the best fitting parameters. This yielded a set of 1024 estimates of local and global noise parameters for a given condition, from which we could estimate 95% confidence intervals. These values are shown as error bars on the plots of best-fitting noise-model parameters.

**Experiment 1: Separable effects of crowding and attentional load on perception of mean orientation**

The first experiment investigated the role of focused visual attention on crowding. If crowding results from a limited attentional resource, then it should increase as attentional resources become less available. To test this

![Figure 3](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933556/ on 10/14/2018)
idea we measured crowding in the presence or absence of a secondary “attentional load” task. Specifically, we employed a $2 \times 2$ design, where subjects were required to judge the overall orientation of a set of iso-eccentric oriented elements, in (a) the presence or absence of randomly oriented crowding elements, and/or (b) the presence or absence of an attentional load task. The secondary load task was to identify the orientation of a white T element within a sequence of tumbling black-Ts, with the whole sequence embedded in a white noise mask of variable contrast. By manipulating the contrast of the mask, we could maintain constant difficulty of the attentional load task and the consequent attentional resource it demanded. In all conditions involving a secondary attentional load task, the subject was instructed that this (and not the orientation judgement) was their primary task; they reported it first, and the overall mean orientation of the oriented elements second. Mean orientation judgements were made with various stimulus orientations (to estimate orientation thresholds), and at various levels of orientation variability (to estimate a noise function).

Figures 4a, 4b, and 4c plot orientation discrimination thresholds for three subjects from the four conditions tested (recall that higher thresholds indicate poorer performance). Solid and dashed lines show the fit of the noise model to the attentional load and crowded conditions, respectively. Note that attentional load (compare thin and thick lines) induces a uniform drop in performance across all orientation ranges tested, for both crowded and un-crowded conditions. This is consistent with a reduction in the effective pool of orientation samples available to the subject (an increase in global noise; long dashed-lines Figure 2c). By contrast the effect of crowding (compare solid and dashed lines in Figure 4) is to elevate thresholds (i.e. reduce performance) at low but not high levels of orientation variability. This is consistent with an increase in local noise, i.e. crowding makes subjects less certain as to the orientation of individual elements (short dashed-lines Figure 4c). Note that a straightforward consequence of this view is that allocation of attentional resources can alleviate the effects of crowding on orientation averaging; at low orientation s.d.s (e.g. 1°) thresholds for the “Crowd” (squares) versus “Crowd+Load” (diamonds) conditions are consistently lower. The effect of not diverting attention is to apparently reduce the effect of crowding on the overall threshold. This is because observers’ averaging performance depends both on global pooling and local precision. Thus, although allocating attention can reduce thresholds under crowding our noise analysis reveals that this is not indicative of a common mechanism but rather that this is achieved by attention influencing a different factor (global efficiency) than crowding (local precision).

The boxed legends inset in Figures 4a, 4b, and 4c give the parameters leading to the best fit of the noise model across the conditions tested. These values are plotted in Figures 4d, 4e, 4f, and 4g for all four subjects tested. Averaging across conditions and subjects, attentional load reduces efficiency—the global pooling of orientation across space—by about 60%. Similarly crowding reduces the precision which subjects can judge the local orientation (of any one element) by around 50%. The independence of these effects is evident comparing the baseline condition (no crowding or load) to simultaneous crowding and load condition. Here both local precision is reduced (on average by around 45%) and global sampling is reduced (by around 57%). The effects of crowding and attentional load are almost perfectly separable in terms of their effects on local and global processing of orientation.

Our finding that elevated attentional load induces effective under-sampling of the stimulus is consistent with a recent report that requiring observers to simultaneously monitor several positions in the visual field (engaging high perceptual load) can induce something akin to “visual neglect,” where some locations are completely ignored (Morgan & Solomon, 2006). It also accords with an earlier report (in abstract form) from one of us that attentional diversion reduced only efficiency on mean orientation judgments (Dakin, 2001a, 2001b). However it is important to understand that a reduction in efficiency reduces the effective number of samples used and does not necessarily mean that whole elements are discounted; for a four-element display an efficiency of $n = 2$ could mean that observers use two elements and ignore two, but equally could mean that half of the information present in all four elements contributes to the pool. In the Discussion we consider how such a reduction in efficiency might be accounted for in terms of reduction in neural activity in an integration system limited by multiplicative noise.

There are different ways in which the raised local noise under crowding could arise. We compared the crowded performance of subject SCD to three models that either (a) replaced crowded target orientations with a weighted average of the flank and target orientation ($0.15f + 0.85t$), (b) probabilistically ($p = 0.02$) replaced target with flank orientations or (c) probabilistically ($p = 0.10$) performed a weighted average of flank and target orientations ($0.5f + 0.5t$). Using baseline internal noise and sampling efficiency from the uncrowded condition, all three models produced behavior very similar to SCD’s crowded discrimination performance. These data do not serve to greatly constrain the specific mechanism underlying the local crowding effect. In the next experiment we sought to determine how reliably flanks do contribute to the crowded percept.

### Experiments 2 and 3: Dual- and multi-pass analysis of crowding

Subjects distracted from performing a task make errors. This would apparently suggest that any errors arising from
Figure 4. Comparison of the effects of crowding and attentional load on judgment of mean orientation. (a–c) Three subjects’ mean-orientation discrimination thresholds are plotted as a function of the orientation variability of the stimulus, under four experimental conditions (indicated by the legend). Fits of the noise model to the crowded (dashed lines) and attentional-load conditions (thicker lines) are also shown with corresponding estimates of local noise (s) and global sampling (n) listed in the legend box. Note that the addition of attentional load, to crowded or uncrowded conditions, has the effect of shifting data upward on log-log axes. Crowding, by contrast selectively impairs performance only at low levels of orientation variability. (d–g) Summarizes these findings for all observers. Crowding reduces subjects’ local precision i.e. at estimating the orientation of individual elements, while attentional load reduces the number of elements they effectively pool to make their judgment.
inattention should be independent of the stimulus because the stimulus is not being processed. While there is direct evidence suggesting that attention can work in this way (Murray, 2008) there is also evidence that attention operates through selective gain change of neural responses (for recent review see Reynolds & Heeger, 2009). If the latter is correct then it follows that attention will be more effective for stimuli that are in some sense more receptive to gain change. In this experiment we attempt to tease apart whether crowding generates the random behavior one would expect from lapses in attention, or whether it is determined by the structure of the flankers.

There is a direct way of assessing the degree to which subjects’ errors depend upon the stimulus being presented and that is to use a dual-pass paradigm where the same sequence of experimental stimuli is presented more than once to the subject (known as a “frozen noise” experiment in the auditory literature; Ahumada & Lovell, 1971; Burgess & Colborne, 1988; Pfafflin & Matthews, 1966). By comparing performance across runs, one can calculate both the percent correct performance (on the task) and also the percent agreement between runs. This procedure can serve to explain how much of the noise on a task is attributable to stimulus-related and -unrelated noise.

To quantify the effect of crowding on response-variability we ran a dual-pass experiment using similar stimuli and procedures to those used in Experiment 1. Orientation s.d. was fixed at 4 degrees. Subjects were presented with two identical runs consisting of 256 trials. They participated in this dual-pass procedure twice: once with crowded (2 × 256 trials) and once with uncrowded stimuli (2 × 256 trials); none of the conditions involved additional attentional load tasks. Results are shown in Figure 5, which plots percent correct performance against percent agreement across the two runs, for two observers under crowded and uncrowded conditions. In order to quantify the similarity of the two data sets we fit the data with a model: pc = m \log_{10}(p_a/100) + 100, where m depends on \text{sum}/\text{next} (Gold, Sekuler, & Bennett, 2004). By bootstrapping the fits of the crowded and uncrowded data separately we derived distributions of model-slopes for each data set, as shown in the lower two panels of Figure 5. Note there is no significant separation between the two distributions indicating that there was no significant difference between the slopes of the best-fitting functions to our data. This demonstrates that crowding does not induce higher levels of stimulus independent error.

If—as our noise analysis indicates—crowding is due to noise generated by local interactions between crowding and target elements, then we would expect that certain target-crowder element pairings would limit performance more effectively than others. If, on the other hand, the increase in local noise induced by crowding were independent of local image structure, then we would predict that the proportion correct achieved by subjects across repeated runs will be determined by simple binomial probability. To address this question we conducted a third experiment in which we presented subjects with 20 identical runs of 128 unique trials (all crowded, no load task). Target orientation s.d. was again fixed at 4°, and the average target offset was set to the uncrowded discrimination threshold (3.12° for SCD and 3.26 for PJB).

The blue histogram in Figure 6 plots the number of stimuli (y-axis) at which a given level of performance was achieved (x-axis); these data are averaged across both observers (i.e. they are based on a total of 40 runs). The red curve shows the mean number of stimuli expected to generate a particular level of performance based on chance alone; the shaded region indicates the 99.99% confidence interval on those estimates. These data were derived using a bootstrap based on 10000 iterations, each iteration generating \(40 \times 128\) responses drawn with a 0.768 probability of a correct response (the mean performance in the psychophysical experiment). If crowding is unrelated to image structure, it is inconceivable that a reasonable proportion of stimuli should fall outside the shaded region. That a large proportion of stimuli do, implies that the local disruption induced by crowding is highly stimulus-specific. For example, the second left-most bar of the histogram indicates that two particular target-crowder configurations elicited an incorrect response 39 out of the 40 times it was presented. Conversely, bars on the far right illustrates that sixteen different target-crowder configuration always (40/40) produced a correct response. The probability of this stimulus-response pairing occurring by chance is vanishingly small. Therefore, crowding-induced increases in local uncertainty (Experiment 1) are, in many cases, determined by local image-driven interactions.

As to the nature of the local interactions that cause crowding, we note two features of error-inducing stimuli (an example is shown as the lower embedded image in Figure 6). First, the mean offset (cue) contained in stimuli will vary probabilistically and, perhaps unsurprisingly, many error-inducing stimuli contain smaller cues. Second, the orientation of the flanks in the error-inducing stimuli tends to agree with the subjects’ report. This is consistent with the notion that the disruptive effect of nearby flankers is due to either compulsory averaging of flanks and target (Parkes et al., 2001) or with signal elements being perceptually substituted by flanker elements (Strasburger, 2005).

**Discussion**

To summarize we have shown:

* An attention-demanding task decreases observers’ efficiency—the effective number of elements sampled—in making a judgment of average orientation of a set of stimuli.
Figure 5. (a, b) Plot of percent correct performance, for two runs with the same (uncrowded) stimuli, against the degree of agreement between runs. (c, d) Same for crowded conditions. The fit-lines in (a–d) are a model from Gold et al. (2004) (dashed lines are 95% confidence intervals for the fit). (e, f) By comparing the distribution of best-fitting slopes, using a bootstrapped data set, one can see that the pattern of percent correct versus percent-agreement is essentially similar with or without crowding. Crowding does not induce subjects to make more stimulus-independent random errors.
The presence of an adjacent crowding distractor decreases the precision of the local orientation sample taken from each location. Variation of attention available for averaging does not affect the local process—crowding—that degrades the local orientation estimate. Similarly, the presence of crowding distractors does not affect the effective number of elements averaged. Errors made by subjects are not random but depend on the structure of the stimulus.

A dissociation of crowding and attention?

We report that attention limits global pooling and that crowding limits estimation of local orientation. It is important to note that while this means the effects of crowding and attentional load are dissociable for orientation averaging it does not (as we explain in the Methodology section above) mean that we have shown that crowding and attention do not interact nor that an experiment could not be created in which this interaction was made explicit. A reasonable interpretation of our results is that crowding affects the fidelity with which the orientation of individual elements are estimated and that attention only affects the pooling of those elements at a later stage. Since we do not know that crowding and attention ever both came to bear on any given element we cannot conclude that our findings indicate a lack of interaction at the local level, which would be a requirement of a dissociation in the sense that it is widely used in the psychological literature. This is an intrinsic limitation of the averaging paradigm that, by its nature, requires the presence of multiple elements. What our results indicate...
is that crowding and attention have different effects on this averaging task and by extension likely are underpinned by different mechanisms. With this knowledge in place we would hope that such paradigms could be extended to examine the issue of local interaction, for example through the use of paired local and global judgements of orientation in such displays.

**Crowding and averaging**

Our results shed some light on earlier proposals that averaging is spared under conditions of crowding (Parkes et al., 2001). There has been some confusion in the literature between the notions (a) that crowding is enforced averaging or (b) that subjects report the average because it is spared the effects of crowding (whereas information about individual elements is not). By changing the orientation variability of our stimuli we have shown that crowding affects observers’ ability to average voluntarily by interfering with the measurement of the local orientation (of individual elements) while preserving observers’ ability to combine multiple measurements. The end-result of this is that crowding only produces measurable effects on voluntary averaging under conditions of low orientation variability (when performance is predominantly governed by observers’ ability to estimate the orientation of individual elements); at high levels of external noise crowding produces little or no effect on performance (since performance is predominantly governed by observers’ ability to pool information which remains unchanged under crowding).

Because crowding seems to operate at the level of local orientation processing, our findings cannot rule out an early cortical locus for the phenomenon. As to the locus of voluntary averaging, we know that subjects can average dichoptically presented elements (Mansouri, Hess, Allen, & Dakin, 2005) indicating that the earliest possible site for element registration is layer 4 of primary visual cortex (Hubel & Wiesel, 1977). Given that crowding also operates under dichoptic presentation, this unambiguously indicates a cortical site for crowding. This in turn is consistent with a recent finding by Blake et al. (2006) who demonstrated that localized orientation-contingent contrast adaptation could be attenuated by the introduction of crowding elements. Because orientation-contingent adaptation is believed to operate at a very early stage of cortical processing, the observation that crowding can attenuate it implies that crowding effects must occur at an equivalently early (or earlier) level of cortical processing.

The high levels of consistency exhibited by subjects in dual- and multi-pass experiments indicate that the effects of crowding are not just generally to elevate the subjects’ uncertainty as to stimulus orientation, as this would manifest as a greater variability in their response to particular stimuli across runs. That the effects of crowding seem to be in no small part stimulus-dependent is encouraging in that it implies there are deterministic rules at work e.g. lateral interactions that have been proposed to underlie both crowding and the tilt illusion. Such stimulus-dependence would also be consistent with involuntary averaging of target and flank structure: either via explicit orientation pooling (Parkes et al., 2001) or by pooling of structure within the excitatory zone of a complex cell (Wilkinson et al., 1997). However several lines of evidence point to crowding requiring more than simple averaging e.g. in the motion domain, surrounds consisting of four drifting Gabors conveying complex motion (e.g. rotation) produce robust crowding even though their averaged motion signal is zero (Bex & Dakin, 2005). Furthermore when subjects make errors on crowded discrimination tasks, they frequently report the flanker identity, which has led some to suggest that flankers may be being substituted for the target under crowding. We have recently shown that a simple weighted averaging model—operating in the spatial position domain—can produce such substitution-like effects and successfully predict errors in discriminating the structure of ‘+’ like patterns. By incorporating measures of structural similarity depending on contour colinearity (Greenwood et al., 2009; May & Hess, 2007) we believe such averaging approaches will be able to explain subtle interactions that underlie crowding between complex stimuli such as letters.

**What does attentional load do?**

There is a long-standing debate in the field of attention as to whether focused attention excludes distractors from processing ("early selection") rather than preventing distractors from influencing behavior ("late selection"). Lavie (1995) noted that evidence favoring either early or late selection occurred under conditions or either low or high attentional load, respectively. To address this directly Lavie and Tsal (1994) had subjects report the identity of a letter surrounded by distractors which could either contain very few elements similar to the target (low load) or many elements similar to the target (high load). Their response-competition paradigm uses the difference in reaction time for responding to the target when the distractors contained an instance of the target (congruent) versus when they did not (incongruent). Simply put, taking longer to identify the stimulus in the presence of an incongruent compared to a congruent distractor indicates that the subject processed the distractor(s). Lavie has shown, using this and other manipulations, that low load allows subjects to see the distractors better. Her conclusion is that attention has a limited capacity (which determines early selection effects) but processes all stimuli (irrelevant as well as relevant). However, what is unclear is whether low load makes subjects poorer at ignoring incongruent items because they can now see individual features better (an improvement in local processing) or whether they can just
see more of them (a global improvement). Using our “distractor-free” displays we can unambiguously resolve this issue: load induces a reduction in sampling efficiency so that whole features are apparently discounted from the pool. Note that this finding does not imply a particular locus for the effect of attentional load; it is possible that the effective sampling rate we report arises not from subjects failing to see whole elements (i.e. early selection) but rather from load inducing a particular pattern of response variability (i.e. late selection). What we can say is that because our experiments employ multiple targets, we have also shown that load influences the processing of task-relevant (as well as irrelevant) items. This confirms that a common attentional resource is affected in such load experiments, one that deals with both targets and distractors.

Morgan and Solomon (2006) have shown that making subjects uncertain as to the location of an oriented target within a field of distracters (a) induces something akin to spatial neglect where whole regions of the visual field are not processed and (b) attenuates the contribution of both signal and noise, leaving signal-to-noise ratio constant. These findings are in close agreement with our view that load (a) affects sampling efficiency, and (b) disrupts a common attentional resource (deployed for both processing of targets and distractors).

Evidence from fMRI indicates that having subjects perform a secondary task while viewing checkerboard patterns, can reduce and even eliminate the visual response in both cortex (Murray, 2008; O’Connor, Manly, Robertson, Hevenor, & Levine, 2004; Schwartz et al., 2005) and LGN (O’Connor et al., 2004). Along with a sizeable body of electrophysiology research (reviewed by Reynolds & Heeger, 2009), these findings indicate that attention serves to regulate the gain of cortical and subcortical neurons. What then is the relationship between the reduced sampling efficiency we observe under conditions of attentional diversion, and the reduction in activity that accompanies inattention in fMRI? As mentioned above, previously we (Dakin et al., 2005) have described an integration model that uses maximum likelihood estimation operating on the response of a population of Poisson-noise limited integration neurons. This model (which estimates motion direction but which can be applied to orientation) emulates psychophysical increases in sampling efficiency we observe as the number of stimulus elements increases. This improvement arises from a simple change in the gain of the neurons pooling direction estimates, providing strong support for the notion that higher sampling efficiency can be explained by modifying gain on the response of integrator neurons.

Crowding and attention

Crowding is broadly inconsistent with contrast gain-change, since the balance of psychophysical evidence unambiguously indicates that crowding is more than simple masking (Pelli, Palomares, & Majaj, 2004; Petrov, Popple, & McKee, 2007). That is not to say stimulus conditions inducing crowding might not also induce activity reduction (Blake et al., 2006) but that we consider it unlikely that these reductions are the prima facie cause of the resulting perceptual degradation of crowded features. Similarly, it is not to say that we do not believe that attentional resources cannot ameliorate the effects of crowding; our own experiments show that allowing subjects to apply undiverted attention to a crowded task improves their performance. However, the application of attention improves performance by improving global sampling and not by reducing the local noise induced by crowding. It has been shown that directing attention to the location of a crowded target can improve its identification (Cavanagh & Holcombe, 2007) or discrimination (Freeman & Pelli, 2007). This, however, in no way implies that crowding results from any kind of attentional limit, merely that extra attention can relieve the effects of crowding. Our proposal is that allocating attention to the target increases the sampling efficiency of any processes employed to recognize it, so increasing the likelihood that one will successfully encode its identity, and allocating attention elsewhere (due to spatial uncertainty, for example) will decrease this likelihood. The distinction we draw here is between the notion that extra attention is overcoming an attentional-limitation that has already caused crowding, or (as our data suggest) that attention can be used to relieve the disruption arising from a completely independent resource limitation we refer to as crowding.

Here we are proposing a specific way in which attention might improve performance: via increased sampling efficiency (i.e. decreasing global noise). It remains an open question as to the spatial scale at which this attentionally modulated efficiency operates. For example, it may occur at the level of entire targets and/or at a level at which local target features (e.g. letter elements) are combined. In the context of motion perception Dakin et al. (2005) have shown that changes in sampling efficiency can be modeled by allowing all moving-elements to contribute equally and to decode the overall direction using a maximum likelihood system operating on a directionally tuned channels that are limited by multiplicative (Poisson) noise. A similar approach could be employed for decoding the orientation of crowded orientation signals. In the context of letter recognition, elevated efficiency could mean that one effectively uses more letter strokes to determine its identity or it could be that multiple letters are falling into a single extended receptive field and allocation of attention serves to increase the contribution of the attended item (possibly through a change in a neural gain-control mechanism (Reynolds & Heeger, 2009). Our results are unable to separate between these alternatives. We believe the level
at which attentional effects arise (either before or after feature integration) to be central to our understanding of crowding.

**Acknowledgments**

This research was supported by a grant from the Wellcome Trust to SCD and PJB.

Commercial relationships: none.
Corresponding author: Steven C. Dakin.
Email: s.dakin@ucl.ac.uk.
Address: UCL Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, U.K.

**Footnote**

1 We are grateful to an anonymous reviewer for pointing this out to us.

**References**


Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with
