Crowding is tuned for perceived (not physical) location

Steven C. Dakin,
Institute of Ophthalmology, University College London, London, UK

John A. Greenwood,
Institute of Ophthalmology, University College London, London, UK, & Laboratoire Psychologie de la Perception, Université Paris Descartes, Paris, France

Thomas A. Carlson, and
Department of Psychology, University of Maryland, College Park, MD, USA

Peter J. Bex
Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA

Steven C. Dakin: s.dakin@ucl.ac.uk; John A. Greenwood: john.greenwood@parisdescartes.fr; Thomas A. Carlson: tcarlson@psyc.umd.edu; Peter J. Bex: peter.bex@schepens.harvard.edu

Abstract

In the peripheral visual field, nearby objects can make one another difficult to recognize (*crowding*) in a manner that *critically* depends on their separation. We manipulated the *apparent* separation of objects using the illusory shifts in perceived location that arise from local motion to determine if crowding depends on physical or perceived location. Flickering Gabor targets displayed between either flickering or drifting flankers were used to (a) quantify the perceived target–flanker separation and (b) measure discrimination of the target orientation or spatial frequency as a function of physical target– flanker separation. Relative to performance with flickering targets, we find that flankers drifting away from the target improve discrimination, while those drifting toward the target degrade it. When plotted as a function of perceived separation across conditions, the data collapse onto a single function indicating that it is perceived and not physical location that determines the magnitude of visual crowding. There was no measurable spatial distortion of the target that could explain the effects. This suggests that crowding operates predominantly in extrastriate visual cortex and not in early visual areas where the response of neurons is retinotopically aligned with the physical position of a stimulus.

Keywords

spatial vision; shape and contour; object recognition; motion—2D

Introduction

Crowding is widely defined as a breakdown in object identification caused by the presence of nearby irrelevant visual structure (Figure 1, and for recent reviews, see Levi, 2008; Whitney & Levi, 2011). These interactions are the primary limitation on peripheral vision,

Corresponding author: Steven C. Dakin. s.dakin@ucl.ac.uk. Address: Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, UK.

affecting a range of visual attributes, e.g. orientation (Wilkinson, Wilson, & Ellemberg, 1997), position (Greenwood, Bex, & Dakin, 2009), motion (Bex & Dakin, 2005), and colour (van den Berg, Roerdink, & Cornelissen, 2007) over large regions of the visual field (Toet & Levi, 1992). However, there is considerable uncertainty regarding the mechanisms that underlie crowding and, accordingly, its precise location within the hierarchy of visual processing.

It is clear that crowding reflects *cortical* processes, at least above the level of monocular neurons in V1, as dichoptic presentation of the target and flanker elements does not reduce the strength of these effects (Flom, Weymouth, & Kahneman, 1963; Kooi, Toet, Tripathy, & Levi, 1994). However, more precise localization has proven elusive. For instance, one finding that places crowding as a late visual process is the lack of effect on contrast detection thresholds for the target element, despite the significant impairments in its identification (Levi, Hariharan, & Klein, 2002; Pelli, Palomares, & Majaj, 2004). This has been taken as evidence that veridical target and flanker signals are present within early feature detection stages of the visual system, prior to their interaction at a later crowding stage that is more involved in identification processes. However, crowding has been shown to elevate contrast detection thresholds when the number of flankers is increased from two to six (Poder, 2008); when crowding is maximized, it may therefore exert its effects at the earliest stages of feature detection. Though it is possible that these effects reflect an interaction with processes more typically ascribed to *masking* (Levi et al., 2002; Pelli et al., 2004), the fact remains that it is difficult to ascertain the locus of crowding based on detection thresholds.

A second line of evidence placing crowding late in the visual hierarchy is the effect of crowding on *adaptation*. Despite impaired identification of target elements, adaptation to crowded Gabor targets produces the same level of threshold elevation for the detection of similarly oriented Gabors as adaptation to uncrowded Gabors (He, Cavanagh, & Intriligator, 1996). This again has been taken to suggest that crowding does not affect the initial stages of feature detection. However, when one reduces target contrast to control for contrast saturation, crowding *does* reduce the degree of orientation-selective adaptation (Blake, Tadin, Sobel, Raissian, & Chong, 2006). This could reflect either suppressive interactions within the early feature detection stages or a change in the target orientation that shifts the peak adaptation away from the test stimuli in these circumstances (Greenwood, Bex, & Dakin, 2010). Either way, it questions the extent to which adaptation can be used to infer the cortical locus of crowding.

Links between attention and crowding have also been used to place crowding as a later-stage integration process. For instance, the division of attention between crowded targets on opposite sides of the vertical midline produces less impairment than divided attention to crowded arrays within the same hemifield (Chakravarthi & Cavanagh, 2009). Similar effects can be induced through having target and flankers span the boundaries between hemifields (Liu, Jiang, Sun, & He, 2009). In addition, while the presentation of targets and flankers at different polarities can reduce the strength of crowding (Hess, Dakin, Kapoor, & Tewfik, 2000; Kooi et al., 1994), this effect disappears with polarity alternations above 6–8 Hz (Chakravarthi & Cavanagh, 2007), a figure that matches the temporal resolution of attention (Verstraten, Cavanagh, & Labianca, 2000). However, attention and crowding have also been shown to have quite dissimilar effects on orientation averaging (Dakin, Bex, Cass, & Watt, 2009)—crowding affects the local noise of each orientation estimate, whereas dividing attention alters the efficiency with which these estimates can be globally pooled. Thus, although performance with crowded stimuli (like performance on many visual tasks) is affected by attentional resources (Mareschal, Morgan, & Solomon, 2010), crowding and attention could interact without being the same mechanism.

Perhaps the clearest feature of crowding is its dependence on the location of target and flanker elements, both in relation to one another and in the visual field. In particular, flankers affect target identification only within a spatial region around the target known as the *interference zone* (Bouma, 1970; Toet & Levi, 1992). The size of this zone grows in the periphery, scaling to approximately 0.5× the target eccentricity with such reliability that it has been termed the *Bouma Law* (Pelli & Tillman, 2008) although the precise value varies depending on stimulus characteristics and task requirements (Whitney & Levi, 2011). More precisely, the shape of these zones is roughly elliptical, with the principal axis extending in a radial direction from the fovea (Toet & Levi, 1992) and a further "centrifugal anisotropy" where flankers nearer to fixation interfere over a smaller range than more eccentric flankers (Bex, Dakin, & Simmers, 2003; Chastain, 1982). This robust scaling with eccentricity implies a relatively constant distance for crowding among cortical receptive fields (Pelli, 2008). It is also clear that the size of these interference zones is independent of target size (Bouma, 1970; Tripathy & Cavanagh, 2002), following instead the center-to-center separation of target and flanker elements (Levi & Carney, 2009; Pelli & Tillman, 2008).

The current "back-pocket" model for crowding is a two-stage account (He et al., 1996; Pelli et al., 2004) where initial feature detection is largely immune to crowding, which instead influences subsequent integration of object features. This account is consistent with the observations above that crowding has little to no effect on adaptation to oriented target elements (Blake et al., 2006; He et al., 1996) nor on contrast detection thresholds (Levi et al., 2002; Pelli et al., 2004) unless the number of elements is high (Poder, 2008). That is not to imply that one can explain crowding as a "single-stage" process since effects of grouping (Livne & Sagi, 2007; Saarela & Herzog, 2009) and (say) object category (e.g., "faceness"; Louie, Bressler, & Whitney, 2007) are unlikely to be mediated in a single "upstream" area. Rather, crowding could encompass a wide range of interference effects operating within multiple visual areas, with the further possibility that interactions between these areas (and feedback from "upstream" areas to lower retinotopic cortical areas) could modulate these effects. The complexity of this system may hamper efforts to identify a single cortical locus for crowding. Indeed, our own efforts to use functional imaging to localize the neural correlates of changes in appearance induced by crowding indicate the involvement of multiple visual areas (Anderson, Dakin, Schwarzkopf, Rees, & Greenwood, in press).

With those caveats in mind, given the reliability of the positional effects of crowding, we wondered whether the latter could be used to even broadly constrain the likely cortical locus of orientation crowding. In particular, our aim was to examine whether it is the *physical* or *perceived* position of flankers that determines the strength of crowding.

To manipulate perceived position, we exploited the *De Valois effect*: Objects with stationary contrast envelopes and moving carriers show a pronounced displacement in their perceived position in the direction of the carrier motion (De Valois & De Valois, 1991). Investigations of the De Valois effect using functional MRI have shown that its magnitude does not correlate with activity in V1 (Whitney et al., 2003) and that it is disrupted only by transcranial magnetic stimulation (TMS) over area MT/V5, with no discernable effects from TMS over V1 (McGraw, Walsh, & Barrett, 2004). Further, visual transients can restore the veridical position of moving elements, suggesting that these veridical signals are maintained within V1 throughout (Kanai & Verstraten, 2006). Positional shifts can nonetheless be induced by crowded motion signals (Whitney, 2005), suggesting that these effects are produced by mid-level mechanisms. Indeed, a dependence on perceived (rather than physical) position is a hallmark of cortical areas such as V3a and V4, as seen with both moving (Maus, Weigelt, Nijhawan, & Muckli, 2010; Sundberg, Fallah, & Reynolds, 2006) and static stimuli (Fischer, Spotswood, & Whitney, 2011). The cortical locus for these motion-induced shifts in position might therefore set a lower limit on the locus of crowding.

Were crowding associated with processing as early as V1, we would expect its magnitude to follow the physical position of flankers. A later-stage process should instead follow the perceived position of these elements.

Methods

Equipment

Experiments were run under the MATLAB programming environment (MathWorks) using software from the PsychToolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented in 14-bit grayscale (achieved using a Bits++ video processor; Cambridge Research Systems) on a LaCie Electron Blue 22″ CRT monitor. The monitor was calibrated with a Minolta photometer and linearized in software using a lookup table. The display operated at a resolution of 1024×768 pixels and a frame refresh rate of 75 Hz and had a mean (background) and a maximum luminance of 50 and 100 cd/m^2, respectively.

Observers

Observers were two of the authors (SCD and JAG) and one naïve subject (DK). All have normal or corrected-to-normal vision and are experienced psychophysical observers.

Stimuli and procedure (crowding)

For the crowding experiments, we measured monocular orientation discrimination thresholds with flickering targets located 8° in the upper visual field. We used a single-interval two-alternative forced-choice (2AFC) procedure using an adaptive staircase procedure (QUEST; Watson & Pelli, 1983) to assess the minimum (threshold) target tilt supporting 75% correct discrimination from vertical (90°). Thresholds are averaged across 2–4 runs. The target was a Gabor ($\sigma_{env} = 0.17$ deg) with a 3.0 c/deg carrier whose contrast counter-phase flickered (4.7 Hz) between 0 and 50% contrast. Stimuli were presented for 750 ms. The test orientation was ramped on with a Gaussian profile peaking at 750 ms (i.e., the end of the sequence) and a σ of 160 ms (Figure 2a). We selected these parameters so that the orientation offset would not be visible to the observer during the buildup of positional distortions arising from the motion of the flankers (i.e., the time course of the De Valois effect), which can take around 500 ms to asymptote (Chung, Patel, Bedell, & Yilmaz, 2007).

Flanking Gabor elements had similar characteristics to the target, except their orientations were fixed at vertical and their carriers either (a) counter-phase flickered at the same rate as the target, (b) drifted (1.6 deg/s) toward the target, or (c) drifted (1.6 deg/s) away from the target. Gabor flankers were positioned between 1.0 and 2.5 deg away from the center of the target, and separate QUEST procedures were run at each separation to determine tilt thresholds with each arrangement. Observers completed five runs for each condition.

Stimuli and procedure (perceived location)

We assessed the perceived separation of similar stimuli (i.e., identical vertically oriented flankers and vertical targets) using a temporal 2AFC (500-ms ISI) and a method of constant stimuli. A randomly selected reference interval contained one of the three classes of stimuli (flickering, inward drift, outward drift) described above, with a separation of 1.5°. The other test interval contained static flickering targets and flankers with variable spacing between 1 and 2.5°. Subjects reported the interval in which elements appeared more widely spaced. We plot the proportion of times they reported elements in the test interval were more widely spaced as a function of the physical displacement of elements, then fit a cumulative Gaussian psychometric function to derive the point of subjective equality (PSE).

Results

Figure 3a plots the results from the perceived position experiment. Raw data (symbols) show the proportion of times observers judged that the spacing of the reference stimulus (which had a fixed spatial offset and whose elements could flicker or drift, depending on condition) was wider than that of the test stimulus (which had a variable spatial offset between elements that always flickered). Symbol color codes the different conditions (i.e., which type of flanker the reference contained). The red symbols indicate that a reference with flankers drifting inward was rarely judged to be more widely spaced than the flickering test, thus shifting the entire data set to the right. Now the test offset leading to 50% performance falls around +0.3 deg (for SCD), indicating that the flickering flankers of the test needed to be pushed inward toward the central element by ~0.3 deg for a perceptual match with the inward-moving reference stimuli. Thus, by fitting these raw data with cumulative Gaussian psychometric functions (curves), we can derive the point of subjective equality (PSE)—the physical flanker offset (within the test) that led observers to report that the spacing of the test stimulus was greater or less than the unchanging reference spacing with equal probability (dashed lines). This PSE is an effective measure of perceived separation. While judgments are largely unbiased with counter-phasing elements (green data), inward-drifting Gabors (red) were seen to be closer to the central Gabor (i.e., requiring a decreased physical separation of elements of the test, a positive/inward bias), while outward-drifting Gabors (blue) appeared more widely separated from the target Gabor (i.e., requiring an increase in physical separation of test elements, giving negative/outward bias). These PSE values shift in each direction by ~0.25 to 0.50 deg, depending on the subject.

Figure 3b plots tilt discrimination thresholds as a function of target–flanker separation. The green symbols/lines indicate performance with flickering flankers; red and blue data are for flankers drifting inward or outward, respectively. In each data set, observe that thresholds become smaller (i.e., performance improves) as the separation between targets and flankers increases. However, at the majority of these target–flanker separations, we also observe substantially more crowding (shaded region) from inward- compared to outward-moving flankers, even though their physical locations are equated. For counter-phase flickering flankers, data are intermediate (2/3 observers) or roughly matched (1/3 observers) to the generally superior performance elicited by outward-moving flankers.

Figure 3c uses the results from the first experiment to reexpress physical target–flanker separation (the *x*-axis in Figure 2b) as *perceived separation*: For each observer, PSE values for each stimulus type were used to horizontally shift the tilt discrimination data from Figure 3b to reflect the perceived (rather than physical) separation of the elements in all three flanker conditions. Performance is now more similar across the different conditions, indicating that it is perceived and not physical position that sets the strength of crowding.

Because crowding affects the identification of a range of visual features besides orientation, including spatial attributes such as spatial frequency (Wilkinson et al., 1997), we wished to know whether these findings were specific to the discrimination of orientation. This also enabled us to address an alternative explanation of our results. In particular, we were concerned that the drifting flankers could have elevated thresholds not because their perceived position shifted but because they caused some compression of visual space around the target. Though it is generally assumed that the drifting carrier of a moving Gabor displaces its perceived location in the direction of motion (top part of Figure 4a), it is possible that they retain their position but distort space around the target location, due to its proximity to either the leading or trailing edge of the flanker motion. For instance, the De Valois effect has been shown to distort the shape of the envelope containing a moving

carrier (Tsui, Khuu, & Hayes, 2007), while moving objects also cause shifts of position in the space surrounding the object (Whitney & Cavanagh, 2000). Were this true in our experiments, in the case of inward-flanking motion (lower part of Figure 4a), space would be compressed around the target, reducing the magnitude of the target's orientation offset (i.e., making it appear more vertical). Similarly, a release from crowding could be effected by an expansion of space from outward-flanking motion with a corresponding increase in the orientation offset size. We would expect that such changes would be accompanied by a change in the perceived spatial frequency of the target, as illustrated in Figure 4: Outward-moving flankers should produce a decrease in perceived spatial frequency, while inward-moving flankers should give an increase. In contrast, a shift in the position of the flanker elements should alter the discrimination of spatial frequency (i.e., threshold elevation, as in the main experiment) without affecting the mean perceived spatial frequency (i.e., judgments should remain unbiased).

To examine this issue, we conducted a third experiment in which observers now had to judge not the orientation but the *spatial frequency* of the central element's carrier, indicating whether it was lower or higher than the flankers, which were separated from the target by a fixed (physical) distance of 1.5°. The spatial frequency of flanker elements was fixed at 3 c/deg, while the target could vary between 1.2 and 6 c/deg according to the method of constant stimuli. By fitting a cumulative Gaussian function to resulting psychometric functions, we could derive two measurements. First, the *perceived spatial frequency* of the test: the SF leading to 50% of reports that the test was higher SF than the reference (i.e., the PSE, or μ parameter from the standard cumulative Gaussian function). Second, the *threshold for SF discrimination* derived from the slope of psychometric functions (specifically, the σ parameter from the cumulative Gaussian). As before, the spatial frequency offset was ramped on, with the same temporal window as used for the orientation offset (see Figures 4b and 4c)—all other experimental details were as before.

Figure 5 plots the psychometric functions from one observer (left column) and the PSE and threshold values from all three observers (right column). PSEs reveal no systematic effect of condition on the matching SF indicating that, for example, the inward condition was not compressing space in a manner that led to any increase in the matched SF of an isolated Gabor (dashed black line). This is consistent with movement direction inducing an offset in the apparent location of the flanker with no distortion of perceived space or spatial frequency around the target. Additionally, the elevation of spatial frequency discrimination thresholds in the presence of inward-drifting flankers and the reduction in thresholds observed with outward-drifting flankers (Figure 5c) demonstrates a difference in the magnitude of crowding that depends on perceived rather than physical location, in good agreement with Experiment 1.

Discussion

Our results demonstrate that crowding, the disruption in object identification caused by clutter, depends on the perceived position of objects rather than their physical positions. To dissociate physical and perceived positions, we used the De Valois effect (De Valois & De Valois, 1991), which causes objects with stationary envelopes and moving carriers to be perceptually displaced in the direction of motion: Crowding stimuli could thus have a physical separation (seen veridically using counter-phasing elements) that was more or less cluttered perceptually (using inward- and outward-moving flankers, respectively). Consistent with a dependence on perceived position, inward-moving flankers produced an elevation of discrimination thresholds (i.e., more *crowding*) than counter-phasing flankers with the same physical separation, and counter-phasing flankers produced worse performance than outward-moving flankers. Using data from a position discrimination task

with the same elements, we were able to align these thresholds on a perceived separation axis, further underscoring the dependence of crowding on perceived position. These effects were demonstrated for the discrimination of both orientation and spatial frequency but can also be seen to have an effect on a form discrimination task (T-orientation discrimination; Figure 6[1]) that is widely used as a proxy for letter recognition in crowding experiments. T-orientation discrimination requires binding of position and orientation information; that illusory displacement of elements can modulate the interaction of the elements of such stimuli suggests that the importance of perceived position is a general phenomenon of crowding.

Our control experiment is consistent with the De Valois effect causing a shift in apparent flanker locations rather than a local distortion of visual space. Shifts in perceived position were not associated with changes in the perceived spatial frequency of crowded elements, as would be expected if the De Valois effect caused the compression or expansion of space (Tsui et al., 2007). This is consistent with models of the De Valois effect that propose asymmetric contrast gain control fields in front of and behind the drifting target: An increase in contrast at the leading edge of the drifting Gabor would cause a shift in the centroid of the element in the direction of motion (Arnold & Johnston, 2003; Whitney et al., 2003). On this basis, one might argue that such a contrast shift could account for the changes in crowding we observe so that inward-moving flankers have a high contrast feature closer to the target than outward-moving flankers. We consider this unlikely given that crowding is relatively insensitive to the contrast of flanker elements (Chung, Levi, & Legge, 2001; Pelli et al., 2004) but more fundamentally that, although the De Valois effect influences local contrast gain control (to uniformly shift the object), it does not do this by simply introducing an asymmetric skewing of contrast energy (Roach, McGraw, & Johnston, 2011). The latter finding also rules out alternative explanations of our SF control experiment, including the possibility that a change in envelope size from these motion-induced distortions could modulate the strength of crowding without an attendant change in carrier SF. Instead, we propose that it is the shift in the perceived centroid of the flanker elements that governs this effect (Levi & Carney, 2009).

Generally, this work is consistent with recent demonstrations that crowding depends on what is perceived rather than the low-level physical properties of the stimulus. As a first example, the magnitude of crowding depends on the number of elements that are *perceived* rather than the number that are *physically* present (Wallis & Bex, 2011), when some of the flankers are temporarily removed from awareness after adaptation (Motoyoshi & Hayakawa, 2010). As a further example, grouping flankers into Gestalt-type "perceptual wholes" (through proximity, similarity, or "good continuation") can alleviate or enhance the effects of crowding (Livne & Sagi, 2007; Saarela, Sayim, Westheimer, & Herzog, 2009), even when it is distant flankers that are manipulated and the immediate target–flanker positions and identities are held constant (Saarela & Herzog, 2009). Such findings point to a complex relationship between the perception of the visual scene and crowding: Although crowding changes the appearance of crowded items (Greenwood et al., 2010), it is also dependent upon our perception of the position and identity of objects in the surround. This interaction would allow crowding to regularize the visual field into simplified texture while also maintaining an important link to the most salient aspects of the visual scene.

More specifically, our finding that crowding depends on perceived and not physical location is consistent with a growing consensus that crowding and contour integration are linked (e.g., Dakin, Cass, Greenwood, & Bex, 2010; Livne & Sagi, 2007; May & Hess, 2007),

---

[1] We note that similar effects of motion on aperture location have been reported with equiluminant dot patterns (Ramachandran & Anstis, 1990).

since it is already known that contour integration depends on perceived and not physical position (Hayes, 2000).

While the cortical locus of crowding has proved difficult to establish, a considerable body of work has examined the locus of the De Valois effect and other associated illusions of position. Retinotopic activity in area V1 is correlated with the physical location of static targets rather than their perceived location (Fischer et al., 2011), while the activity produced by the De Valois effect goes in the opposite direction to the perceptual shifts (Whitney et al., 2003). Similarly, shifts in perceived position are disrupted by TMS over area MT/V5 but not over V1 (McGraw et al., 2004). However, though crowding does disrupt the identification of the direction of moving objects (Bex & Dakin, 2005), it seems unlikely that cortical area MT/V5 could subserve the entirety of these crowding effects, given its primary role in motion processing. Rather, a dependence on perceived (vs. physical) position appears to be a general property of visual cortex beyond V1. For instance, activity in area V3a follows the illusory position of objects that fade in motion (Maus et al., 2010), while the position of receptive fields in area V4 has been shown to shift in a manner consistent with perceived shifts in the flash-lag illusion (Sundberg et al., 2006). Activity within both of these areas follows mislocalization errors with static targets (Fischer et al., 2011), as well as a range of higher order areas including the lateral occipital cortex and the parahippocampal place area.

The cortical regions responsive to perceived position also show many properties that would seem to be implicated in crowding. In particular, neurons within V4 demonstrate properties suitable for the production of "temporal crowding" effects (Motter, 2006) and receptive field sizes consistent with the well-known eccentricity-dependent scaling of crowding (Motter, 2009). There is also less crowding between targets and flankers that span the vertical but not the horizontal meridian (Liu et al., 2009), consistent with the hemifield organization of receptive fields in area V4. As the physiology of cortical areas V1–V3 predicts effects of both the vertical and horizontal meridians (Sereno et al., 1995; Zeki, 2003), this supports V4 as the earliest cortical locus for crowding. Other behavioral evidence suggests an important role for V2 in crowding. Namely, Freeman and Simoncelli (2010) have recently presented results from a texture-based model of crowding based on a modified version of the texture synthesis algorithm of Portilla and Simoncelli (2000). Observers are unable to detect the presence of gross texture substitutions into natural images (see Figure 1). The spatial extent of such texture metamers is consistent with receptive field sizes in mid-ventral areas such as V2 and V4. However, while the role of cortical area V2 in mislocalization illusions such as the De Valois effect is currently unclear, V2 activity does not differentiate between physical and perceived object positions in a position discrimination task (Fischer et al., 2011). Thus, at the very least, our results strongly suggest that crowding is a consequence of visual processing after area V1, with cortical area V4 appearing the most likely candidate.

Of course, this picture would be complicated if the altered position of elements in the De Valois effect were to feed back to earlier cortical stages. While the relatively slow buildup of these positional displacements might argue for some role of feedback (Chung et al., 2007), the physiological data described above do not provide any direct evidence for effects arising within primary visual cortex (McGraw et al., 2004; Whitney et al., 2003). Additionally, behavioral evidence suggests that visual transients can restore the veridical position of moving elements (Kanai & Verstraten, 2006). This suggests that the De Valois effect is a later-stage process in vision, with the veridical signals maintained within the initial processing stages.

The dependence of crowding on the perceived position of objects thus offers strong evidence to favor the two-stage model of crowding (He et al., 1996; Pelli et al., 2004). By this account, the initial stages of feature detection are largely unaffected by crowding, which

exerts an effect only on the subsequent integration and identification of object features. The two-stage account is also consistent with the observations that crowding has little to no effect on adaptation to oriented target elements (Blake et al., 2006; He et al., 1996) nor on contrast detection thresholds (Levi et al., 2002; Pelli et al., 2004) unless the number of elements is high (Poder, 2008). As outlined in the Introduction section however, these prior studies do not clearly distinguish early gain control (Levi et al., 2002; Pelli et al., 2004) from high-level feature identification stages. The present demonstration of a dependence on perceived object positions demonstrates that crowding interactions must take place at a level where perceived object position has been extracted. As the spatial relations between crowded objects are a fundamental determinant of whether or not crowding occurs (Bouma, 1970; Dakin et al., 2010; Toet & Levi, 1992), we consider this a strong demonstration of the high-level nature of crowding. Nonetheless, while crowding is clearly a later-stage process in vision, we note that demonstrations that crowding can be *modulated* by attention (Alvarez & Cavanagh, 2005; Chakravarthi & Cavanagh, 2007, 2009; Mareschal et al., 2010) do not necessarily *equate* crowding with attention (Dakin et al., 2009).

Acknowledgments

References

Alvarez GA, Cavanagh P. Independent resources for attentional tracking in the left and right visual hemifields. Psychological Science. 2005; 16:637–643. [PubMed: 16102067]

Anderson E, Dakin SC, Schwarzkopf DS, Rees G, Greenwood JA. The neural correlates of crowding-induced changes in appearance. Journal of Vision. (in press).

Arnold DH, Johnston A. Motion-induced spatial conflict. Nature. 2003; 425:181–184. [PubMed: 12968181]

Bex PJ, Dakin SC. Spatial interference among moving targets. Vision Research. 2005; 45:1385–1398. [PubMed: 15743609]

Bex PJ, Dakin SC, Simmers AJ. The shape and size of crowding for moving targets. Vision Research. 2003; 43:2895–2904. [PubMed: 14568377]

Blake R, Tadin D, Sobel KV, Raissian TA, Chong SC. Strength of early visual adaptation depends on visual awareness. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103:4783–4788. [PubMed: 16537384]

Bouma H. Interaction effects in parafoveal letter recognition. Nature. 1970; 226:177–178. [PubMed: 5437004]

Brainard DH. The psychophysics toolbox. Spatial Vision. 1997; 10:433–436. [PubMed: 9176952]

Chakravarthi R, Cavanagh P. Temporal properties of the polarity advantage effect in crowding. Journal of Vision. 2007; 7(2):11, 1–13. http://www.journalofvision.org/content/7/2/11. [PubMed: 18217826]

Chakravarthi R, Cavanagh P. Bilateral field advantage in visual crowding. Vision Research. 2009; 49:1638–1646. [PubMed: 19362572]

Chastain G. Confusability and interference between members of parafoveal letter pairs. Perception & Psychophysics. 1982; 32:576–580. [PubMed: 7167357]

Chung ST, Levi DM, Legge GE. Spatial-frequency and contrast properties of crowding. Vision Research. 2001; 41:1833–1850. [PubMed: 11369047]

Chung ST, Patel SS, Bedell HE, Yilmaz O. Spatial and temporal properties of the illusory motion-induced position shift for drifting stimuli. Vision Research. 2007; 47:231–243. [PubMed: 17190608]

Dakin SC, Bex PJ, Cass JR, Watt RJ. Dissociable effects of attention and crowding on orientation averaging. Journal of Vision. 2009; 9(11):28, 1–16. http://www.journalofvision.org/content/9/11/28. [PubMed: 20053091]

Dakin SC, Cass J, Greenwood JA, Bex PJ. Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. Journal of Vision. 2010; 10(10):14, 1–16. http://www.journalofvision.org/content/10/10/14. [PubMed: 20884479]

Dakin SC, Greenwood JA, Bex PJ, Carlson TA. Crowding depends on perceived (not physical) position. Perception. 2008; 37:81.

De Valois RL, De Valois KK. Vernier acuity with stationary moving Gabors. Vision Research. 1991; 31:1619–1626. [PubMed: 1949630]

Fischer J, Spotswood N, Whitney D. The emergence of perceived position in the visual system. Journal of Cognitive Neuroscience. 2011; 23:119–136. [PubMed: 20044886]

Flom MC, Weymouth FW, Kahneman D. Visual resolution and contour interaction. Journal of the Optical Society of America. 1963; 53:1026–1032. [PubMed: 14065335]

Freeman J, Simoncelli EP. Crowding and metamerismin the ventral stream [Abstract]. Journal of Vision. 2010; 10(7):1347, 1347a. http://www.journalofvision.org/content/10/7/1347.

Greenwood JA, Bex PJ, Dakin SC. Positional averaging explains crowding with letter-like stimuli. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:13130–13135. [PubMed: 19617570]

Greenwood JA, Bex PJ, Dakin SC. Crowding changes appearance. Current Biology. 2010; 20:496–501. [PubMed: 20206527]

Hayes A. Apparent position governs contour-element binding by the visual system. Proceedings of the Royal Society of London B, Biological Sciences. 2000; 267:1341–1345.

He S, Cavanagh P, Intriligator J. Attentional resolution and the locus of visual awareness. Nature. 1996; 383:334–337. [PubMed: 8848045]

Hess RF, Dakin SC, Kapoor N, Tewfik M. Contour interaction in fovea and periphery. Journal of the Optical Society of America A, Optics, Image Science, and Vision. 2000; 17:1516–1524.

Kanai R, Verstraten FAJ. Visual transients reveal the veridical position of a moving object. Perception. 2006; 35:453–460. [PubMed: 16700288]

Kooi FL, Toet A, Tripathy SP, Levi DM. The effect of similarity and duration on spatial interaction in peripheral vision. Spatial Vision. 1994; 8:255–279. [PubMed: 7993878]

Levi DM. Crowding—An essential bottleneck for object recognition: A mini-review. Vision Research. 2008; 48:635–654. [PubMed: 18226828]

Levi DM, Carney T. Crowding in peripheral vision: Why bigger is better. Current Biology. 2009; 19:1988–1993. [PubMed: 19853450]

Levi DM, Hariharan S, Klein SA. Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. Journal of Vision. 2002; 2(2):3, 167–177. http://www.journalofvision.org/content/2/2/3.

Liu T, Jiang Y, Sun X, He S. Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. Current Biology. 2009; 19:127–132. [PubMed: 19135367]

Livne T, Sagi D. Configuration influence on crowding. Journal of Vision. 2007; 7(2):4, 1–12. http://www.journalofvision.org/content/7/2/4. [PubMed: 18217819]

Louie EG, Bressler DW, Whitney D. Holistic crowding: Selective interference between configural representations of faces in crowded scenes. Journal of Vision. 2007; 7(2):24, 1–11. http://www.journalofvision.org/content/7/2/24. [PubMed: 18217839]

Mareschal I, Morgan MJ, Solomon JA. Attentional modulation of crowding. Vision Research. 2010; 50:805–809. [PubMed: 20122950]

Maus GW, Fischer J, Whitney D. Perceived positions determine crowding. PLoS ONE. 2011; 6:e19796. [PubMed: 21629690]

Maus GW, Weigelt S, Nijhawan R, Muckli L. Does area V3A predict positions of moving objects? Frontiers in Psychology. 2010; 1:186. [PubMed: 21897824]

May KA, Hess RF. Ladder contours are undetectable in the periphery: A crowding effect? Journal of Vision. 2007; 7(13):9, 1–15. http://www.journalofvision.org/content/7/13/9. [PubMed: 17997637]

McGraw PV, Walsh V, Barrett BT. Motion-sensitive neurones in V5/MT modulate perceived spatial position. Current Biology. 2004; 14:1090–1093. [PubMed: 15203002]

Motoyoshi I, Hayakawa S. Adaptation-induced blindness to sluggish stimuli. Journal of Vision. 2010; 10(2):16, 1–18. http://www.journalofvision.org/content/10/2/16. [PubMed: 20462317]

Motter BC. Modulation of transient and sustained response components of V4 neurons by temporal crowding in flashed stimulus sequences. Journal of Neuroscience. 2006; 26:9683–9694. [PubMed: 16988039]

Motter BC. Central V4 receptive fields are scaled by the V1 cortical magnification and correspond to a constant-sized sampling of the V1 surface. Journal of Neuroscience. 2009; 29:5749–5757. [PubMed: 19420243]

Pelli DG. The VideoToolbox software for visual psychophysics: Transforming number into movies. Spatial Vision. 1997; 10:437–442. [PubMed: 9176953]

Pelli DG. Crowding: A cortical constraint on object recognition. Current Opinion in Neurobiology. 2008; 18:445–451. [PubMed: 18835355]

Pelli DG, Palomares M, Majaj NJ. Crowding is unlike ordinary masking: Distinguishing feature integration from detection. Journal of Vision. 2004; 4(12):12, 1136–1169. http://www.journalofvision.org/content/4/12/12.

Pelli DG, Tillman KA. The uncrowded window of object recognition. Nature Neuroscience. 2008; 11:1129–1135.

Poder E. Crowding with detection and coarse discrimination of simple visual features. Journal of Vision. 2008; 8(4):24, 1–26. http://www.journalofvision.org/content/8/4/24. [PubMed: 18484863]

Portilla J, Simoncelli EP. A parametric texture model based on joint statistics of complex wavelet coefficients. International Journal of Computer Vision. 2000; 40:49–71.

Ramachandran VS, Anstis SM. Illusory displacement of equiluminous kinetic edges. Perception. 1990; 19:611–616. [PubMed: 2102995]

Roach NW, McGraw PV, Johnston A. Visual motion induces a forward prediction of spatial pattern. Current Biology. 2011; 21:740–745. [PubMed: 21514158]

Saarela T, Herzog M. Crowding in multi-element arrays: Regularity of spacing [Abstract]. Journal of Vision. 2009; 9(8):1017, 1017a. http://www.journalofvision.org/content/9/8/1017.

Saarela TP, Sayim B, Westheimer G, Herzog MH. Global stimulus configuration modulates crowding. Journal of Vision. 2009; 9(2):5, 1–11. http://www.journalofvision.org/content/9/2/5. [PubMed: 19271915]

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, et al. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science. 1995; 268:889–893. [PubMed: 7754376]

Sundberg KA, Fallah M, Reynolds JH. A motion-dependent distortion of retinotopy in area V4. Neuron. 2006; 49:447–457. [PubMed: 16446147]

Toet A, Levi DM. The two-dimensional shape of spatial interaction zones in the parafovea. Vision Research. 1992; 32:1349–1357. [PubMed: 1455707]

Tripathy SP, Cavanagh P. The extent of crowding in peripheral vision does not scale with target size. Vision Research. 2002; 42:2357–2369. [PubMed: 12350424]

Tsui SY, Khuu SK, Hayes A. The perceived position shift of a pattern that contains internal motion is accompanied by a change in the pattern's apparent size and shape. Vision Research. 2007; 47:402–410. [PubMed: 17184808]

van den Berg R, Roerdink JB, Cornelissen FW. On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. Journal of Vision. 2007; 7(2):14, 1–11. http://www.journalofvision.org/content/7/2/14. [PubMed: 18217829]

Verstraten FA, Cavanagh P, Labianca AT. Limits of attentive tracking reveal temporal properties of attention. Vision Research. 2000; 40:3651–3664. [PubMed: 11116167]

Wallis TS, Bex PJ. Visual crowding is correlated with awareness. Current Biology. 2011; 21:254–258. [PubMed: 21277208]

Watson AB, Pelli DG. QUEST: A Bayesian adaptive psychometric method. Perception & Psychophysics. 1983; 33:113–120. [PubMed: 6844102]

Whitney D. Motion distorts perceived position without awareness of motion. Current Biology. 2005; 15:R324–R326. [PubMed: 15886084]

Whitney D, Cavanagh P. Motion distorts visual space: Shifting the perceived position of remote stationary objects. Nature Neuroscience. 2000; 3:954–959.

Whitney D, Goltz HC, Thomas CG, Gati JS, Menon RS, Goodale MA. Flexible retinotopy: Motion-dependent position coding in the visual cortex. Science. 2003; 302:878–881. [PubMed: 14500849]

Whitney D, Levi DM. Visual crowding: A fundamental limit on conscious perception and object recognition. Trends in Cognitive Sciences. 2011; 15:160–168. [PubMed: 21420894]

Wilkinson F, Wilson HR, Ellemberg D. Lateral interactions in peripherally viewed texture arrays. Journal of the Optical Society of America A: Optics, Image Science, and Vision. 1997; 14:2057–2068.

Zeki S. Improbable areas in the visual brain. Trends in Neurosciences. 2003; 26:23–26. [PubMed: 12495859]
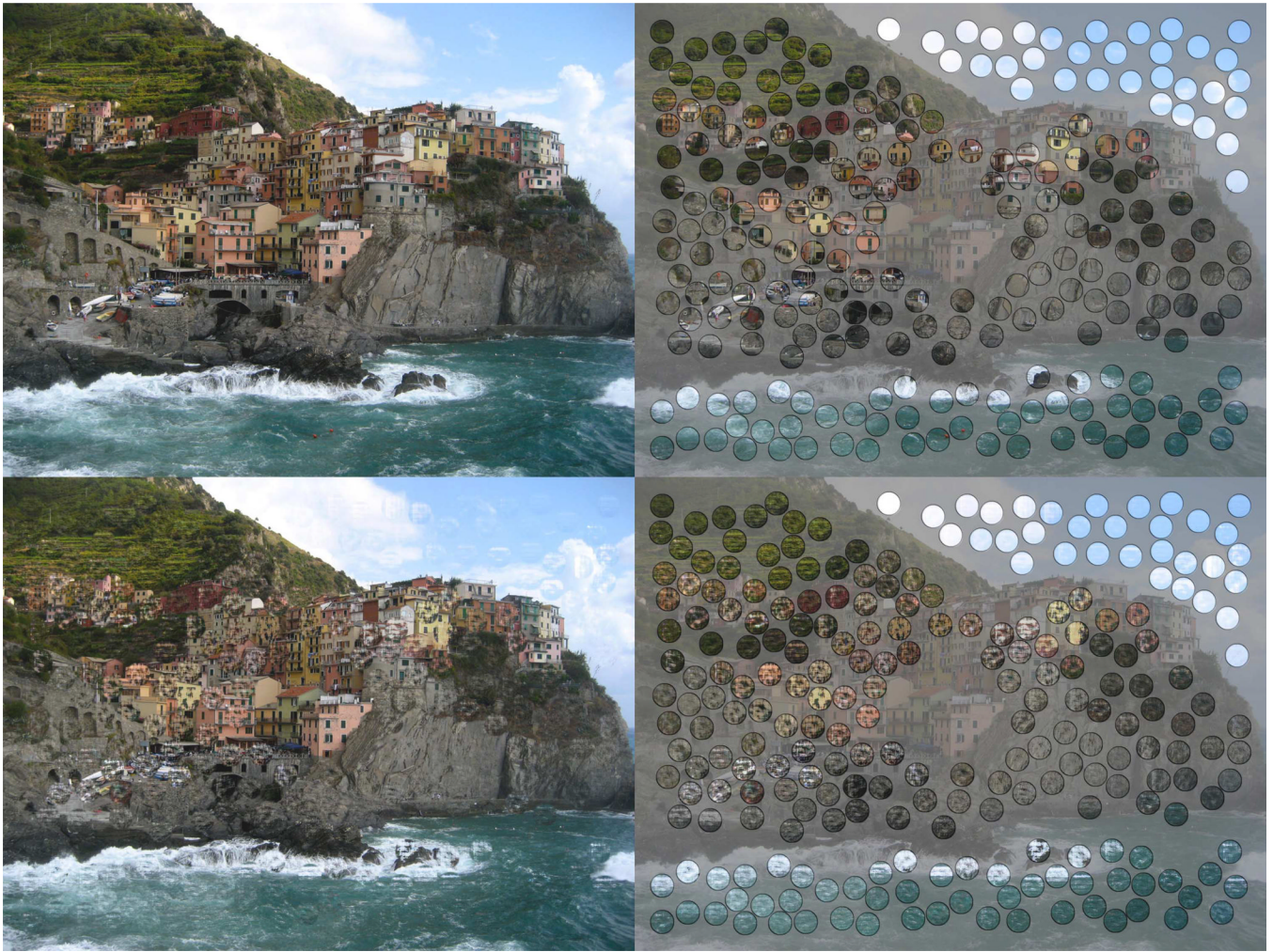
Figure 1.
Crowding in a natural scene. The two images in the left column are hard to tell apart, particularly if one fixates, in turn, the large pink house in the center foreground of each. The lower image actually only contains ~35% of the original information since image structure within a large number of patches (top right) has been phase-scrambled (bottom right) and embedded in the original. Crowding renders this substantial disruption invisible.
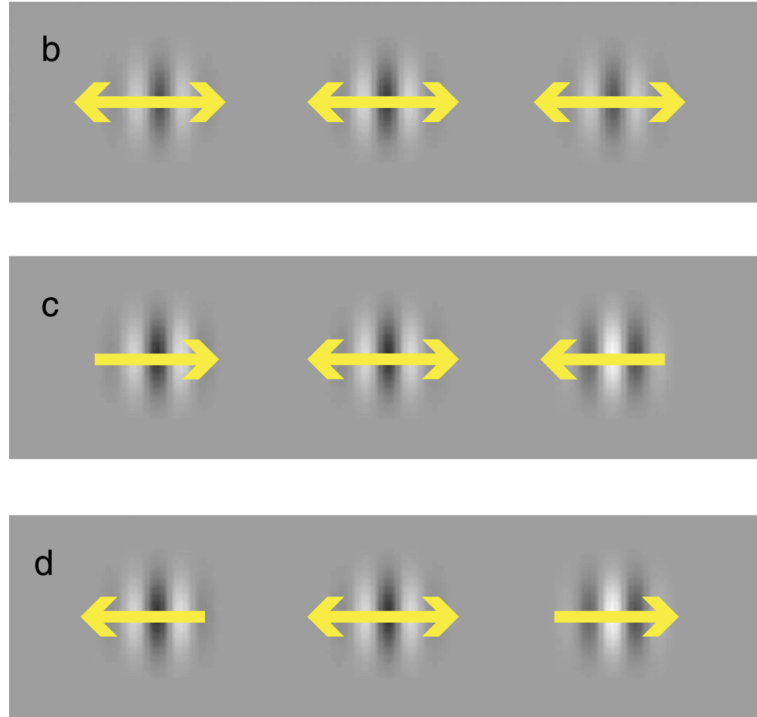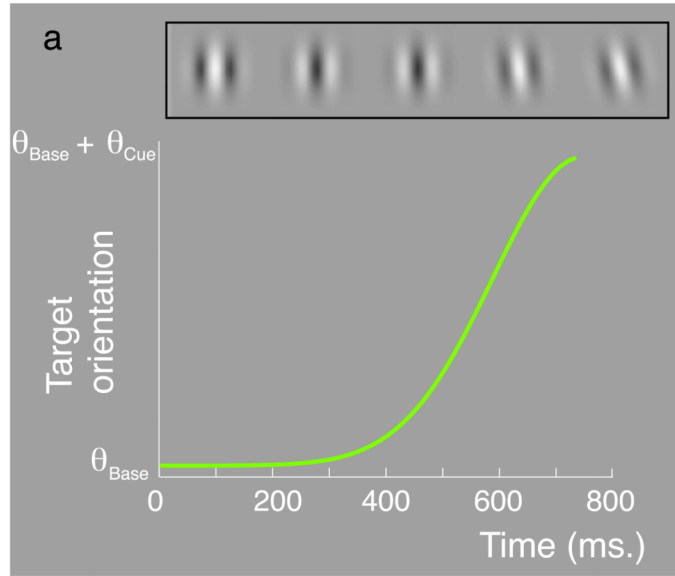
Figure 2.
(a) Stimuli were composed of a counter-phase flickering Gabor target whose orientation changed smoothly over the course of 750 ms (such that the maximum target tilt matched the time course of the *De Valois effect*). This target was flanked to the left and right by two vertical Gabors that either (b) flickered (double-headed arrow), (c) drifted inward, or (d) drifted outward relative to the target location. Observers were required to judge the tilt of the target Gabor.
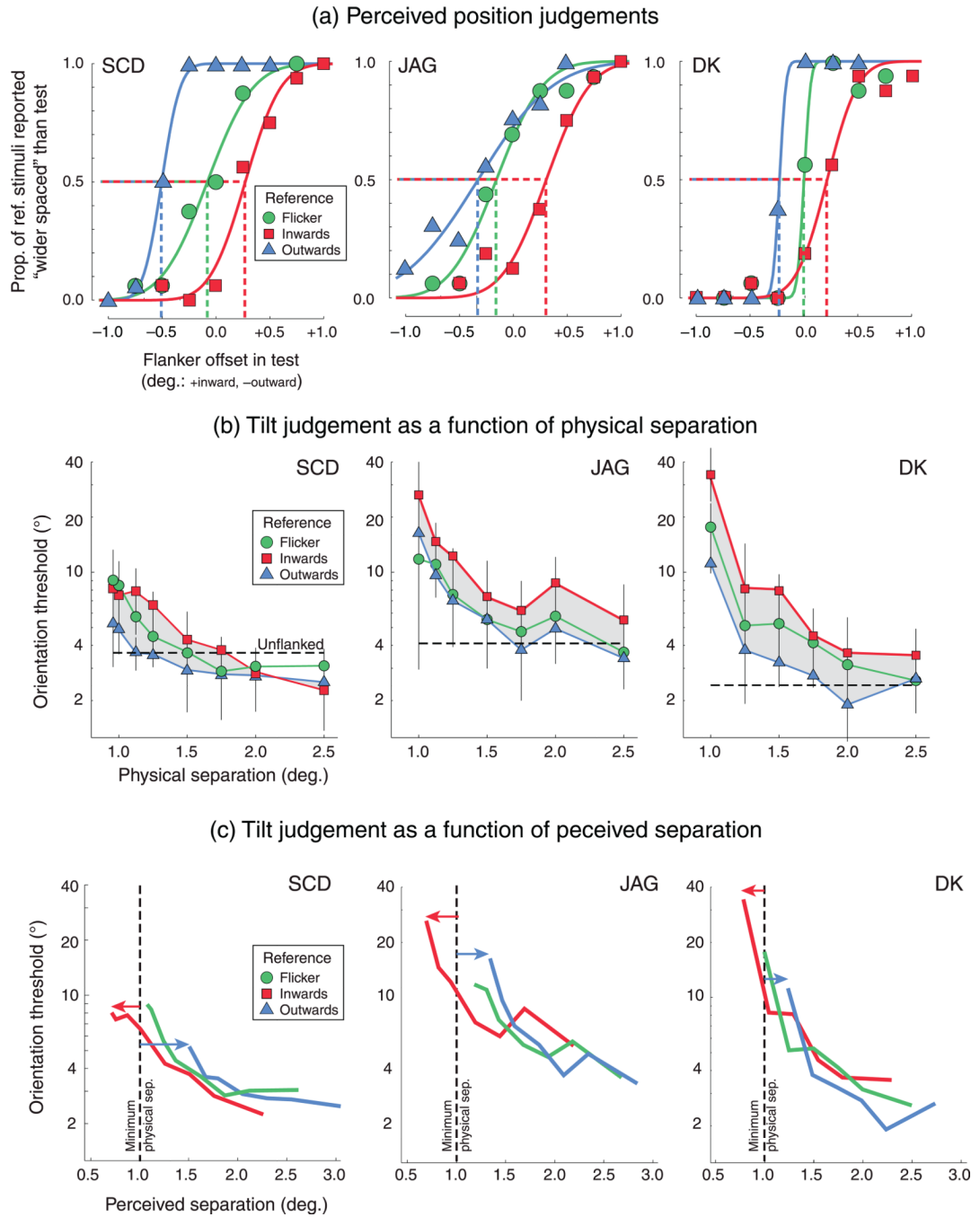
(a) Perceived position judgements

(b) Tilt judgement as a function of physical separation

(c) Tilt judgement as a function of peceived separation

Figure 3.
Results from the (a) perceived position and (b, c) crowded tilt discrimination experiments.
(a) Compared to performance with flickering flankers (green circles), outward-drifting
flankers (blue triangles) expand perceived separation, while inward-drifting flankers (red
squares) compress separation. Dashed lines indicate the point of subjective equality, at
which separation the test stimuli in each condition were seen as equivalently spaced to the
flickering reference stimuli. (b) Lowest tilt thresholds occur with flankers that drift outward
and highest thresholds with inward-drifting flankers. Error bars are T1 standard deviation of
the bootstrapped threshold estimates. (c) We can now plot thresholds from (b) as a function
of perceived separation by simply horizontally shifting functions (arrows) by the PSE values

from the perceived separation experiment. Although there is variation between subjects, thresholds are now interleaved and (broadly) fall on top of each other.
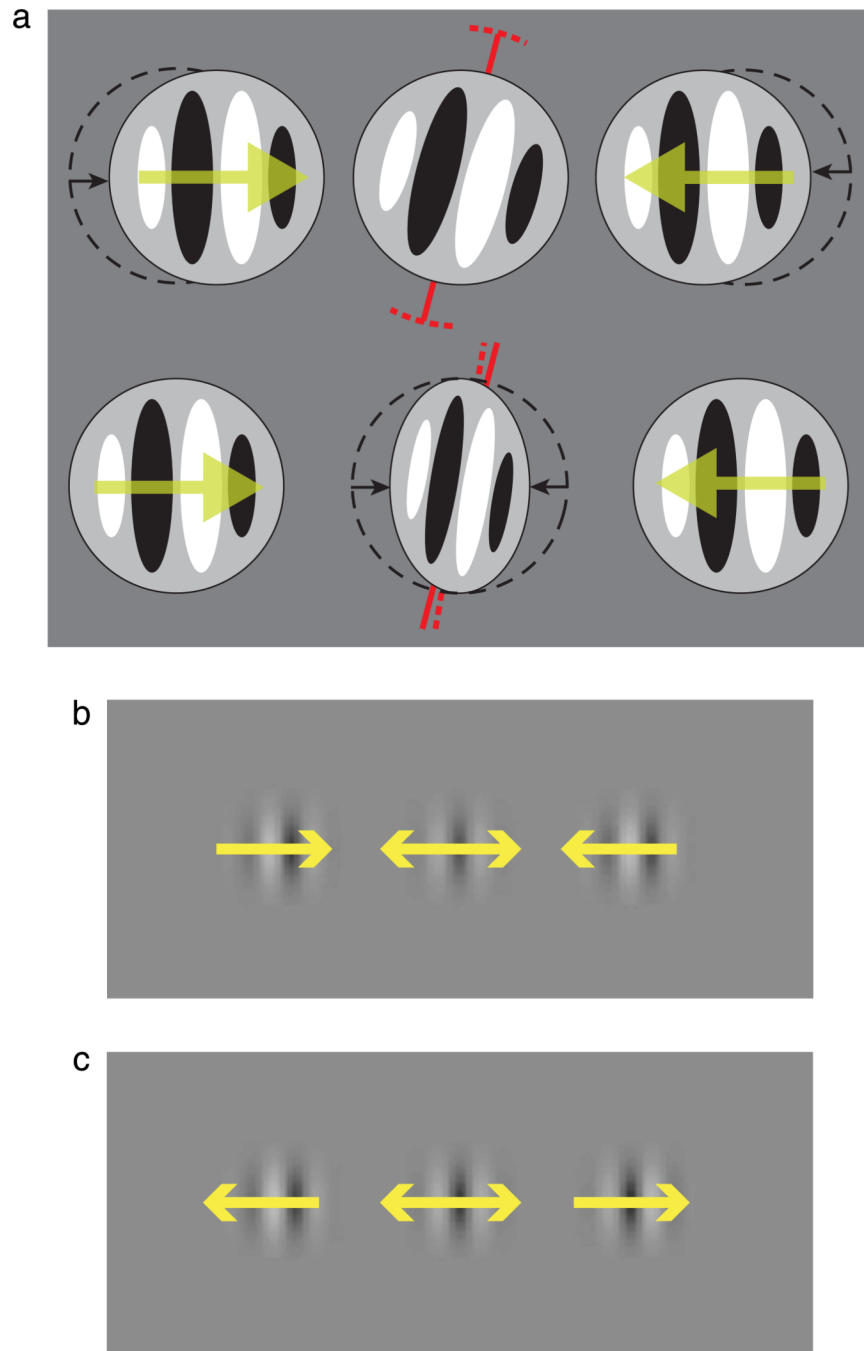
Figure 4.
(a) Two possible mechanisms by which inward-drifting flankers could elevate orientation discrimination thresholds: (upper) position offset or (lower) contraction of space at the target location. The former predicts an increase in spatial frequency discrimination thresholds due to the closer perceived position of the flankers, without an effect on the mean perceived spatial frequency. The latter predicts an increase in perceived spatial frequency at the target location without a necessary increase in threshold elevation. (b, c) Similar to Figures 1c and 1d except now the stimulus changes not its orientation but its spatial frequency.
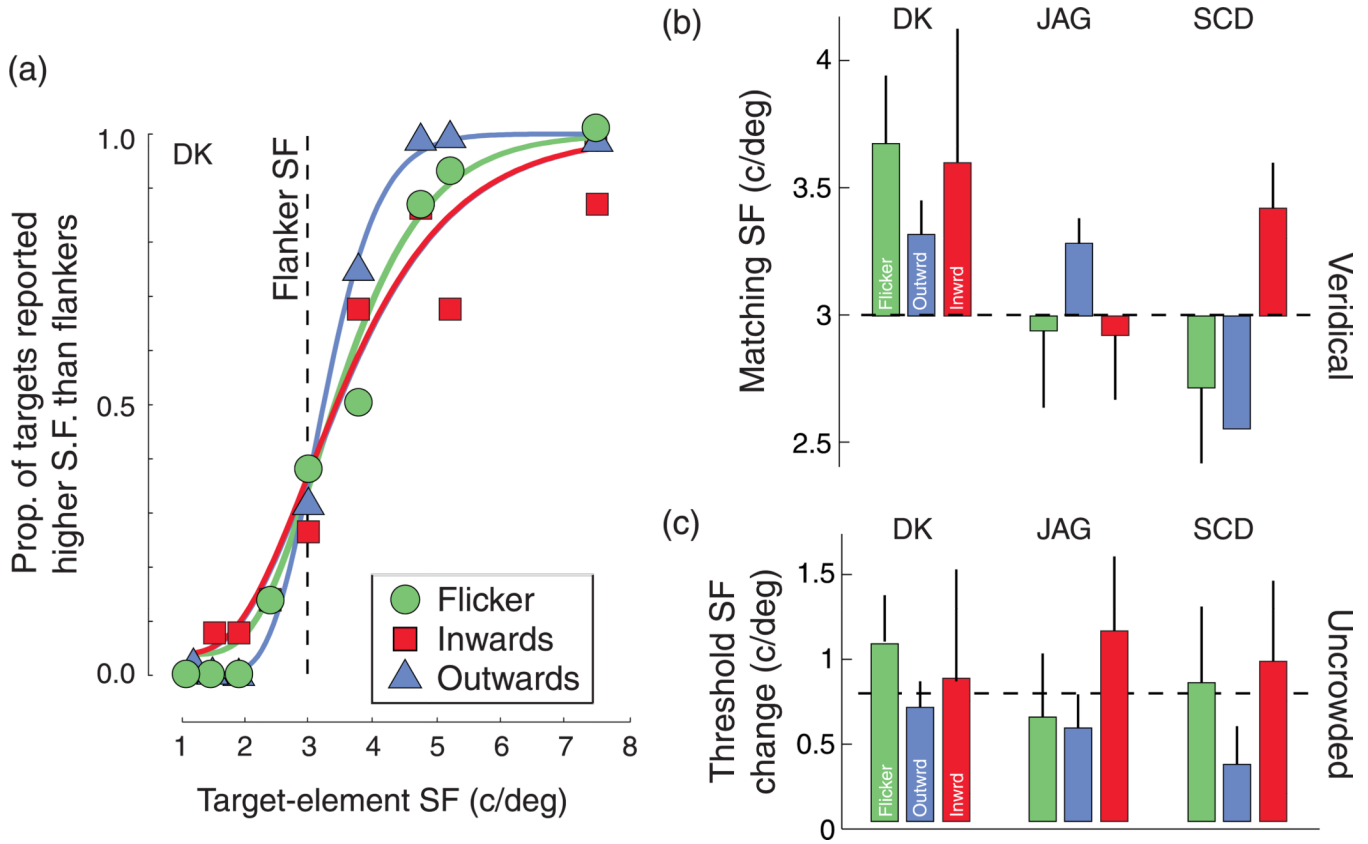
Figure 5.
Results from the spatial frequency (SF) discrimination experiment. (a) Typical psychometric functions for discrimination of the SF of a target Gabor from the SF of the flanking elements surrounding it, as a function of flanker SF. (b, c) Bar graphs plot (b) matching SF of the flanker (i.e., bias) and (c) threshold SF change for three observers with the three classes of flanker used. Units are c/deg and error bars show T1 standard deviation of the corresponding estimates (based on a bootstrap). The black dashed line indicates the (b) veridical match and (c) average uncrowded threshold. Note that flanker condition has little consistent effect on matching SF but that thresholds are lower with outward (blue bars) compared to inward (red bars) flankers, replicating the effects on orientation discrimination described in Experiment 1, for SF discrimination.
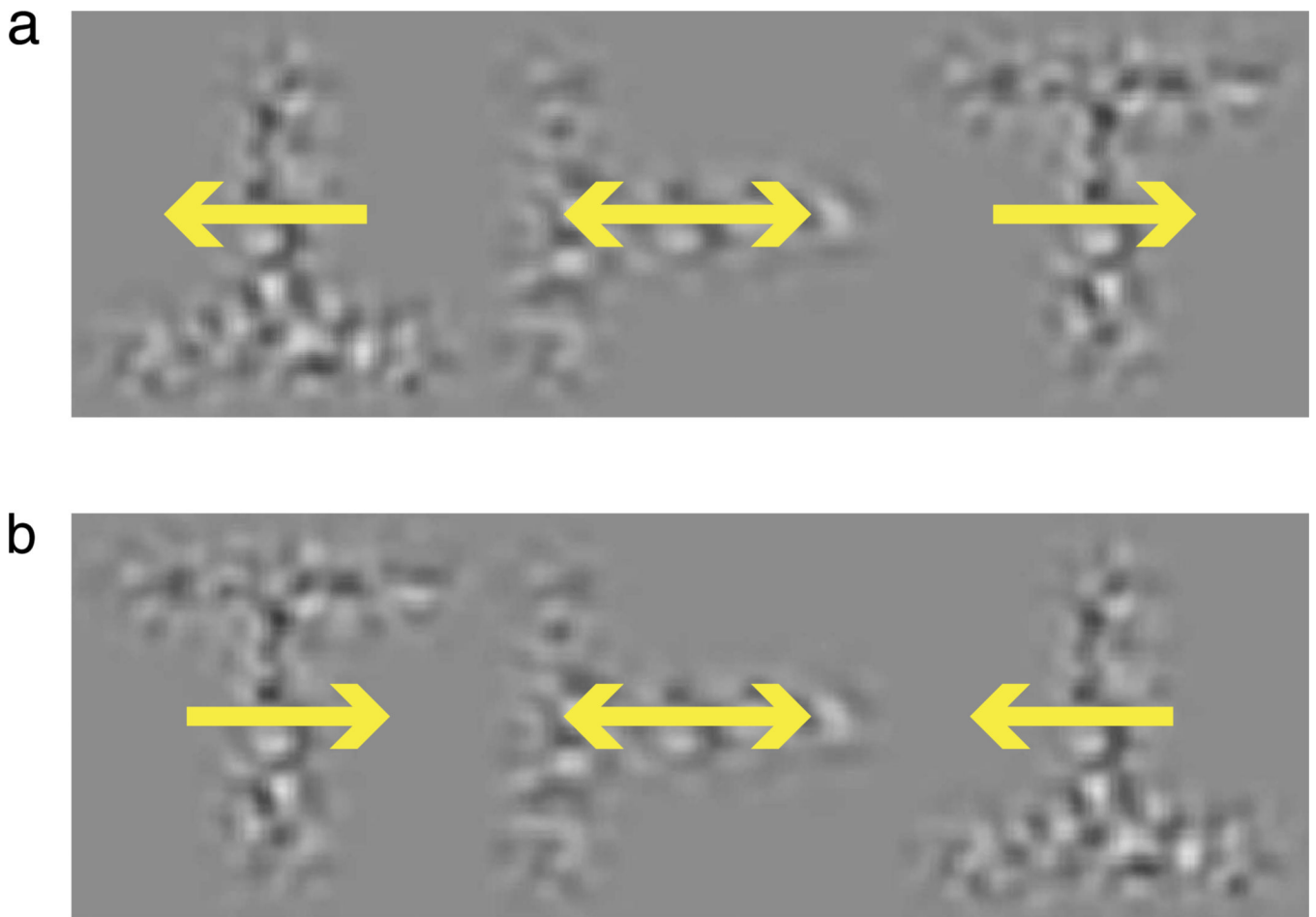
Figure 6.
A demonstration of the effects of perceived position on the crowding of T-orientation discrimination (a task similar to letter recognition). Each element is constructed by moving band-pass filtered noise within T-shaped apertures. Viewing these movies in one's peripheral vision, it should become apparent that the inward-moving flankers present in (b) make the recognition of the central "T" element considerably more difficult than the outward-moving flankers present in (a).