Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli

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

John Cass,
Institute of Ophthalmology, University College London, London, UK

School of Psychology, University of Western Sydney, Sydney, Australia

John A. Greenwood,
Institute of Ophthalmology, University College London, London, UK

Peter J. Bex
Institute of Ophthalmology, University College London, London, UK

Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA

Abstract

We investigated how crowding—a breakdown in object recognition that occurs in the presence of nearby distracting clutter—works for complex letter-like stimuli. Subjects reported the orientation (up/down/left/right) of a T target, abutted by a single flanker composed of randomly positioned horizontal and vertical bars. In addition to familiar retinotopic anisotropies (e.g., more crowding from more eccentric flankers), we report three *object-centered* anisotropies. First, inversions of the target element were rare: errors included twice as many ±90° as 180° target rotations. Second, flankers were twice as intrusive when they lay above or below (*end-flanking*) compared to left or right (*side-flanking*) of an upright T target (an effect that holds under global rotation of the target–flanker pair). Third, end flankers induce subjects to make erroneous reports that resemble the flanker (producing a structured pattern of errors), but errors induced by side flankers do not (instead producing random errors). A model based on probabilistic weighted averaging of the feature positions within contours can account for these effects. Thus, we demonstrate a set of seemingly "high-level" object-centered crowding effects that can arise from "low-level" interactions between the features of letter-like elements.

Keywords

spatial vision; object recognition; perceptual organization; shape and contour; visual acuity; crowding

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

Introduction

Our ability to identify objects in the peripheral visual field, compared to the fovea, is disproportionately vulnerable to the presence of nearby visual clutter, a phenomenon known as *visual crowding* (for a recent review see Levi, 2008). Several principles determine the strength of crowding. The first is the similarity of the target and flanking object(s) (Figure 1a) with dissimilar objects crowding less than those that are similar. This finding holds across "low-level" stimulus dimensions including contrast polarity, color, orientation, spatial frequency, direction, and speed (Bex & Dakin, 2005; Chung, Levi, & Legge, 2001; Kooi, Toet, Tripathy, & Levi, 1994; van den Berg, Roerdink, & Cornelissen, 2007; Wilkinson, Wilson, & Ellemberg, 1997). For "higher level" stimulus properties such as facial identity (Farzin, Rivera, & Whitney, 2009; Louie, Bressler, & Whitney, 2007), one finds that the "objectness" of the flankers determines the level of crowding: upright flanking faces crowd more than inverted flanking faces irrespective of the target orientation (Figure 1b). This is the principal evidence for "object-centered" crowding. Such effects do not seem to hold for words that are equally well crowded by inverted words (Yu, Akau, & Chung, 2010), as well as letters once featural similarities are taken into account (Chastain, 1982). This suggests that the limits imposed by crowding on the recognition of letters and words likely originates from interactions between constituent features, whereas more "holistic" interactions may take place for complex objects such as faces.

In addition to target–flanker similarity, a major determinant of crowding is the location of the target and flanker elements, both in relation to one another and in the visual field. Flankers affect target identification only if they fall in a spatial region surrounding the target known as the *interference zone* (Figure 1c). Interference zones grow with the target's eccentricity (Bouma, 1970; Toet & Levi, 1992) but are independent of target size (Bouma, 1970; Pelli, 2008; Tripathy & Cavanagh, 2002). The shape of these zones is roughly elliptical, with the principal axis falling on a line extending from the zone center to the fovea (Toet & Levi, 1992). Along this axis, there is a further "centrifugal anisotropy" such that flankers that are nearer to fixation can get closer to the target without interfering with identification than more eccentric flankers (the so-called "in–out" anisotropy; Bex, Dakin, & Simmers, 2003; Chastain, 1982; Petrov & Popple, 2007; Toet & Levi, 1992).

The strength of crowding is also determined by the configuration of the target and flanker elements. For example, crowding can be relieved by arranging flankers so that they group into contours that exclude the target (Livne & Sagi, 2007). A similar effect that employs Gestalt grouping principles, observed by Saarela and Herzog (2009), is illustrated in Figure 1d. Notice that the proximity between the target and nearest flankers is identical in the second and third configurations, yet the target orientation is much clearer when the flankers group together in the rightmost example. These effects indicate that crowding is not just determined by local feature interactions. In general terms, this is consistent with the above evidence for object-centered crowding. Specifically, here we see a release from crowding when flankers are arranged to promote grouping into objects, while the works of Farzin et al. (2009) and Louie et al. (2007) demonstrate an increase in crowding when the holistic identity of flankers is more clearly evident. These global properties can modulate crowding even when featural similarity and proximity remains constant.

Thus, crowding can be determined by a combination of the position, similarity, and global configuration of both target and flankers and may occur among mechanisms tuned for low- and high-level stimulus properties. Beyond this, we understand little about the specific mechanisms of crowding or how they interact. Given the strong disruption to both letter and object recognition (Flom, Heath, & Takahashi, 1963; Pelli & Tillman, 2008) and discrimination of the orientation and spatial frequency of Gabor stimuli (Solomon, Felisberti, & Morgan, 2004; Wilkinson et al., 1997), some authors have attributed crowding to a failure of attentional resolution such that the identity of cluttered elements is lost from awareness (He, Cavanagh, & Intriligator, 1996). This view is however inconsistent with the observation that the errors in crowding tasks correlate highly with the identity of flanking elements (Huckauf & Heller, 2002; Nandy & Tjan, 2007; Strasburger, 2005; Strasburger, Harvey, & Rentschler, 1991) as well as on repeated presentations of the same stimuli (Dakin, Cass, Bex, & Watt, 2009). This has most commonly been attributed to the *substitution* of either flankers (Chastain, 1982; Strasburger et al., 1991) or their independent features (Chastain, 1982; Wolford, 1975) into the target location. However, a recent study demonstrates that the perceived position of features within letter-like stimuli can be attributed to a noisy weighted averaging of position information (Greenwood, Bex, & Dakin, 2009), which can account for both the systematic flanker-driven errors and the threshold elevation that occurs under crowding. This fits with prior work demonstrating that averaging can account for judgments of orientation made with an array of crowded Gabor stimuli (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). While averaging may seem inconsistent with measured distributions of the perceived orientation of crowded targets, which tend toward bimodality rather than being normally distributed (Gheri & Baldassi, 2008), this pattern is consistent with both a "Signed Max" rule—in which the observer's response is based on the maximum output from a set of noisy filters (Baldassi, Megna, & Burr, 2006)—as well as a later decisional strategy (Greenwood et al., 2009) known as "reference repulsion" (Jazayeri & Movshon, 2007; Rauber & Treue, 1998). Target–flanker averaging is likely to precede these later decisional processes.

Although we can formulate simple rules for predicting the strength of crowding between simple features, such as lines, based on their similarity and proximity, Figure 1d illustrates that this is not enough to predict the percept. One still needs to understand the layout of other lines in the scene, i.e., the global organization of the scene, to predict the magnitude of crowding. This finding, others like it (Livne & Sagi, 2007), and "holistic crowding" (Farzin et al., 2009; Louie et al., 2007) suggest that higher order relationships between the target and features that are not necessarily immediate neighbors are critical. In this paper, we use stimuli that are defined by conjunctions of features to examine these relationships, extending previous studies that have focused on judgments of a single feature within simple elements (such as the orientation of Gabors), to try and move toward the richer class of complex object stimuli.

We utilize "letter-like" stimuli that are composed of two lines, one horizontal and one vertical (Greenwood et al., 2009). These stimuli can be configured as tumbling "Ts" when one feature is placed at the extreme of the stimulus space and more ambiguous cross-like stimuli with intermediate positions (e.g., left flanker in the second panel of Figure 2). By constraining the target elements to be Ts at one of the four cardinal orientations (up/down/

left/right), and allowing the position of flanker features to vary continuously, we can examine complex multi-dimensional interactions in crowding with a simple four-alternative forced-choice response. Furthermore, rather than focus on overall error rates, as in many prior studies of crowding, the approach we take is to examine, on a trial-by-trial basis, what combinations of flanker and targets lead subjects to report that the target is in a particular state. This is a richer data set that we will show places more rigorous constraints on models of crowding. In this way, we can examine interactions between the standard retinotopic (e.g., in–out) anisotropies, determined by absolute position of the target and flanker in the visual field, and *object-centered* anisotropies, i.e., variations in the pattern of errors that arise for a given combination of target and flanker (irrespective of the global, retinotopically defined rotation of the pair). These results can be modeled using a variant of weighted averaging that incorporates both the probabilistic nature of interference zones around the target (essentially switching the system between crowded and uncrowded states), demonstrating that seemingly high-level effects in crowding (at least for our letter-like stimuli) can emerge from a low-level model.

Crowding of complex stimuli likely reflects the cumulative effect of various levels of the object recognition process dealing with insufficient information (e.g., due to under-sampling in the periphery). The utility of our simple "letter-like" stimuli is that they are "minimal objects" that can be characterized fully by the coordinate of their feature crossing, but that still require feature integration and the simultaneous representation of the position and orientation of multiple features. Such stimuli can help us build on our current understanding of how individual features are corrupted under crowing, to begin to understand how the integration processes operating in the recognition of more complex stimuli works. To be clear, we do not believe that even a full account of crowding of such stimuli will fully explain crowding at the level of whole complex objects (such as faces) especially when the experiments devised to probe such processes have carefully matched local feature interactions (Farzin et al., 2009; Louie et al., 2007). Instead, we aim to set limits on how much models based entirely on low-level feature interactions can predict whole-object crowding effects in increasingly complex stimuli. The principles we uncover may have application for understanding crowding in higher order visual information, which plays a part in crowding of complex stimuli such as faces.

General methods

Subjects

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

Apparatus

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 maximum luminance of 50 and 100 cd/m^2, respectively.

Stimuli

Targets were white (100 cd/m^2) "Ts" with Sloan proportions (stroke width = 1/5 letter height) presented at one of the four cardinal orientations (0°, 90°, 180°, or 270°). Stimuli were presented monocularly (to the dominant eye) 12.5° in the upper visual field. Targets were always abutted by a single flanker composed of one horizontal and one vertical stroke, randomly positioned (within ±0.5 letter width) around the flanker center. The size of the target and flank was under continuous control (see below), but the center-to-center distance between them was fixed at 1× the target size. Stimuli were presented for 250 ms and were pre- and post-masked using a 50% contrast, phase-scrambled version of the stimulus (Figure 2).

We ran four interleaved conditions (256 trials each) with flankers either positioned above, below, to the left, or to the right of the target. Before each trial, a pre-cue was presented at fixation to highlight the target location (consisting of a small letter: U, D, L, or R, respectively). An adaptive staircase (QUEST; Watson & Pelli, 1983) maintained stimulus size (target and flanker were scaled together) to elicit 82% correct identification (4AFC) of target orientation (N, S, E, or W). This procedure was repeated 6–7 times for each subject. Early termination of some runs led to SCD, JRC, and LA finally running 6355, 5926, and 7457 trials, respectively. Our use of independent QUESTs to minimize the influence of absolute target–flanker configuration meant that stimuli were presented at different sizes for different configurations. However, note that our subsequent analysis of object-centered effects always pooled data across all four conditions, minimizing any confound that might have arisen from this difference.

Results

Figure 3 plots data from the experiment, starting with the minimum target size that supports 82% correct performance in the four interleaved QUEST runs for each flanker location (each repeated 6–7 times). These thresholds (Figure 3a) resemble the familiar centrifugal anisotropy illustrated in Figure 1c: threshold size was 80% greater when the flanker was located in the more eccentric location (i.e., above the target) compared to when the flanker was less eccentric (below the target) or iso-eccentric (left or right of the target). Recall that a QUEST procedure rescaled stimuli to maintain performance at 82% threshold within any one run, which completely factors out the influence of this centrifugal anisotropy in any subsequent analysis of the data. For data pooled across these conditions, Figure 3b shows that subjects' error rates are broadly similar across the four target identities. Figure 3c plots the frequency of the three possible incorrect reports pooled across the different letter-orientation conditions and normalized to be relative to an upright "T". Note that only 20% of errors arise from the observer reporting a 180° rotation of the target with the remaining 80% coming from their reporting a ±90° rotation of the target. This is an object-centered effect in the sense that it depends on the target identity rather than the retinotopic target orientation.

Figure 4 plots all 2817 errors made by the three observers in the experiment. The color of each plotted symbol indicates what target orientation subjects (erroneously) reported on a particular trial while the location of the symbol within the panel indicates the shape of the flanker (see Key). Thus, dots near the center indicate that the flanker was "+-like", dots near the lower left indicate that the flanker was "L-like," etc. In the main black panels, each column of four sub-panels is derived from a given interleaved run, with each sub-panel derived from a particular combination of target identity and flanker location. Cyan labels indicate *side-flanked* conditions—that is, the entire flanker–target combination could be rotated into registration with an upright T flanked on either the left or right. Magenta labels indicate *end-flanked* conditions where the flanker–target combination could be registered with an upright T, flanked either above or below. Note that there are substantially fewer symbols in cyan-labeled than in magenta-labeled conditions; end flankers lead to more errors. We also observe somewhat higher error rates within end-flanking conditions when the flanker fell below the target (again, relative to an upright T) than when it fell above. This was particularly evident for subject LA whose error rates were 23% and 6.4% (for flankers below and above) but was less so for the other two subjects: SCD, 19% and 20% and JRC, 27% and 20% (flankers below and above). Data are thus broadly consistent with the end- versus side-flanking division, although there seems to be some variation among observers as to the importance of element placement within end-flanking conditions.

As well as leading to more errors, end-flanked conditions led observers to make target reports that typically resembled the flanker. Consider the top left panel of Figure 4a, which plots errors for subjects reporting the orientation of a rightward-pointing T in the presence of a flanker to the right (note that this is classed as end-flanking because in this configuration the flanker position is above the target, relative to an upright T target). Many of the data are labeled yellow or blue, indicating that observers mostly reported upright or inverted Ts (respectively) when they made errors. Note also that the yellow dots cluster at the top and the blue at the bottom of the panel (respectively) indicating that, e.g., flankers that resembled upright Ts led subjects to report that the target was an upright T. This is consistent with subjects reporting the target orientation that most closely resembled the flanker on error trials, a strategy we term *substitution* (Greenwood et al., 2009). As discussed in the Introduction section, this trend could reflect flankers as a whole replacing the target, or the constituent features of flankers combining with those of the target. Consistent with the trend shown in Figure 3c, note that the scarcity of green symbols (leftward T reports, i.e., a 180° rotation of the target) indicate that a simple random wholesale substitution of the flanker cannot account for all errors. We return to this point below.

Errors are not only less frequent in side-flanked conditions, but they also lead to reports that bear a weaker resemblance to the flanker that induced them. Pooling across subjects, 23/24 end-flanked configurations but only 9/24 side-flanked configurations achieved significance at the 1% or less level in a χ^2 test on the distribution of errors that were consistent or inconsistent with a substitution report. To explore this lack of structure further, we pooled subjects' data across conditions to generate a single map of errors arising from either side- or end-flanked conditions for each observer. Starting with the 16 maps above, we first exploited the similarity of errors that is evident for similar target identities within a given row and pooled these pairings (maintaining the end-versus side-flanker distinction) to reduce

the number of maps to 8 for each subject. We next took those maps and rotated each into registration with the map arising from performance with an upright T (which involves physical rotation and recoding error reports) and then pooled within all end- and side-flanked conditions to generate a pair of maps for each subject. These are plotted in Figure 5. Note that all errors are now encoded relative to an upright T. The left part of the figure shows schematically the conditions that have been averaged to generate each map. Bar graphs plot the total number of errors contributing to each plot for each subject. A χ^2 test on the frequency of errors in each response category showed that the data for end-flank and side-flank conditions for all subjects were significantly different from chance ($p < 0.01$ for all comparisons: χ^2 end: SCD = 116.48; JRC = 100.84; LA = 107.77: side: SCD = 12.02; JRC = 29.49; LA = 78.93). These data reveal that subjects make twice as many errors in end- as side-flanked conditions and twice as many 90° rotation as 180° rotation errors irrespective of flanker location relative to the target. The pooled data also make the point that errors are much more structured in the end- than in the side-flanked conditions, and that the structure in the errors indicate that subjects' erroneous reports are frequently similar to the structure of the flankers.

To summarize, the psychophysical experiment, as well as measuring and then factoring out well-known centrifugal anisotropies in the spatial extent of crowding, has revealed three novel object-centered crowding effects for this class of stimulus:

1.   Only 20% of errors arise from subjects reporting a 180° inversion of the target. The other 80% arise from them reporting ±90° target rotations.

2.   When an upright T is end-flanked (i.e., the flanker falls above or below it), it is twice as likely to lead to an error than when it is side-flanked. This result holds under retinotopic rotation of the whole flanker–target combination—i.e., it is an *object-centered* effect.

3.   Errors arising from end-flanking are more predictable than errors arising from side-flanking, in that they tend to resemble the flanker.

Modeling

We now consider how such findings constrain several quantitative models of crowding illustrated in Figure 6. Consistent with the high levels of positional uncertainty in the peripheral field (Hess & Hayes, 1994; Levi & Klein, 1986; Rentschler & Treutwein, 1985), all models include a stage of early positional noise. The first model we consider is the *substitution* account, whereby either the whole flanker (Chastain, 1982; Strasburger et al., 1991) or its features (Chastain, 1982; Wolford, 1975) are probabilistically substituted for the target or its features. As shown in Figure 6, the model fails to predict any of the three effects reported: it predicts identical performance for side and end flankers (both in terms of error rates and structure of errors) and a number of inversion errors that is either equal to (whole flanker) or greater than (feature) the 90° rotation errors. The second model is *weighted positional averaging* (WPA) of feature positions, which has been shown to account for the attraction of target feature positions to those of flankers under crowding (Greenwood et al., 2009). Here, the final percept is determined using a weighted combination of the noisy

estimates of feature positions from both the target and flankers. This model can account for inversion effects because it captures the notion that distance between the flanker and target features matters: perceived inversion of an upright T is rare when target (T) and flanker (F) positions are equally weighted (T50:F50) because this requires that the average position of the horizontal stroke in the target (i.e., at the top) and in the flanker (anywhere from top to bottom) must fall below the center. This outcome is only possible for an equally weighted average when the values are sufficiently perturbed by positional noise. As the weights are increasingly biased toward flankers, the number of inversion errors increases (e.g., T25:F75; Figure 6), and, with a complete bias toward the flankers (T0:F100), the model output resembles substitution. Intermediate weights can reproduce the 20% rate of inversion errors. However, this model cannot predict the difference between end- and side-flanked conditions.

We propose an extension of this model—*probabilistic anisotropic positional averaging* (PAPA; Figure 6)—that is similar in structure to the WPA model except that an extra stage is added between the early encoding of positional noise and weighted averaging of positions (for details, see Supplementary materials). This intermediate stage uses an interference zone around each feature to determine both the probability that crowding will arise and the strength of influence of a flanker on the average; the closer a flanker feature is to the target feature, the higher this probability and the more the flanker will contribute to the average. Under this view, crowding is "all or nothing". If crowding arises, then the prediction is generated the same way as in WPA. If crowding does not arise, the target feature is returned veridically (corrupted only by early noise).[1] Given recent suggestions that crowding may relate to processes of contour integration (Livne & Sagi, 2007; May & Hess, 2007), interference zones were allowed to vary independently in size along each axis, relative to the contour orientation. In this way, we can mimic integration along contours containing near-collinear elements ("snakes") and/or near-parallel elements ("ladders"; Dakin & Hess, 1999; Field, Hayes, & Hess, 1993) as in Figure 6. We generated predictions from the PAPA model using a Monte Carlo simulation of our experiment and optimized its five free parameters to produce the best match to our data.[2] The free parameters are (specified relative to target width, W): standard deviation of early positional noise ($\sigma_{\text{pos}} = W/4$), dimensions of the interference zone ($2.5\,W \times 1.0\,W$), weight of the interference zone to determine contribution of the flanker to the average ($w_{\text{average}} = 1.00$), and weight of the interference zone to determine probability that crowding occurs ($w_{\text{prob}} = 0.66$). The predictions are shown in the error maps and bar graphs at the bottom of Figure 6. This model gives an excellent account of the rather complex pattern of psychophysical findings: half as many inversion errors as rotation errors, fewer errors in side- than end-flanked conditions, and errors that reflect flanker structure only in the end-flanked conditions. Below, we outline the reasons for this agreement with our psychophysical data. Here we also note a limitation of the model in that it fails to capture the somewhat lower density of errors in the center of the maps (i.e., when flankers were cross-like). We return to this point in the Discussion section.

---

[1]Note that we ran simulations using many variants of this model without the probabilistic component; no deterministic model came close to capturing the pattern of psychophysical results we have described.
[2]MatLab code for the model is available on request from SCD.

Discussion

Object-centered crowding

Results from the present study provide an insight into both the instances in which crowding is likely to occur and the mechanisms that underlie crowding when it does occur. We first replicate the standard retinotopic centrifugal anisotropy—i.e., that there is greater crowding from more eccentric than more foveal flankers (Bex, Simmers, & Dakin, 2003; Chastain, 1982; Petrov & Popple, 2007; Toet & Levi, 1992). In addition, when the influence of these retinotopic anisotropies are factored out, we report three novel *object-centered* anisotropies. First, when crowding did occur, the resultant errors included twice as many misperceptions of ±90° rotations as 180° rotations of the target. Second, relative to an upright T target, flankers were much less intrusive when they lay to the left or right (*side flankers*), compared to above or below (*end flankers*) the target. Third, while errors that were induced by end flankers tended to resemble the flanker, errors induced by side flankers did not.

A general principle to draw from this pattern of errors is that in addition to being retinotopically dependent, the appearance of crowded targets is largely determined by an interaction between the identity of the target and the flanker. That is, rather than simply adding noise or suppressing the identity of target elements, crowding involves specific interactions between the target and flanker identities. This is consistent with the correlation between flanker identities and subjects' errors in letter-recognition tasks (Huckauf & Heller, 2002; Nandy & Tjan, 2007; Strasburger et al., 1991), the biases induced by simpler stimuli such as oriented Gabors (Solomon et al., 2004), and shifts in the position of lines within similar letter-like stimuli (Greenwood et al., 2009). These types of biases have previously been shown to correlate highly across repeated presentations of the same stimuli (Dakin et al., 2009) consistent with their basis in low-level pre-attentive mechanisms. The end result of these interactions is a change in target appearance that tends to increase the similarity between target and flanker elements (Greenwood, Bex, & Dakin, 2010).

That we observe crowding effects determined more by the identities of targets and flankers, than by the physical arrangement of lines on the screen, classifies this squarely as an object-centered phenomenon. Interestingly, although these object-centered anisotropies might appear likely to arise from higher level mechanisms (i.e., operating after the extraction of target identity), we can account for this complex pattern of errors using a relatively simple, feed-forward model in which estimates of the positions of object features are corrupted by position uncertainty (the only source of noise in the model) followed by an averaging stage. Consistent with prior work (Greenwood et al., 2009), the substitution of either flanker elements, or the features of flankers, gives an inadequate characterization of the errors induced by crowding. As described in the Introduction section, Signed Max models also predict substitution-like behavior, because observers report the peak tuning of the noisy filter giving the largest absolute response (Baldassi et al., 2006; Gheri & Baldassi, 2008), which is always close to the orientation of the target or one of the flanking elements. Furthermore, these models predict an equal rate of rotation and inversion errors that is quite unlike the actual patterns produced. Weighted averaging, on the other hand, accounts for this ratio of errors quite well, by modulating the effect of the flankers on the target. The simplest version

of this scheme fails to encapsulate our data completely, however, by promoting structured errors (i.e., incorrect target reports that resemble the flankers) regardless of the flanker location. The relatively unstructured side flanker errors require that we incorporate elongated interference zones into the model, which determine weighting of both the strength of crowding (the amount flankers contribute to the average) and the probability that crowding happens, depending on the location of features within the elements.

The PAPA model

Figure 7 illustrates how the PAPA model produces this complex pattern of behavior for the example of an upright T target. First, Figure 7a shows how the model predicts the low incidence of target inversions (i.e., "downward-facing T" responses to an upright T target) among errors. This happens because horizontal strokes, which drive inversions in models averaging the locations of target and flanker features, are more likely to fall outside the anisotropic interference zones than vertical strokes, thus "switching off" crowding. Second, Figure 7b shows that the PAPA model successfully predicts higher error rates with end flankers because vertical features drive the most frequent errors (±90° rotation) for this configuration, and it is these vertical features that are more likely to fall inside elongated interference zones when they are contained in end compared to side flankers. Third, as a consequence of these differences in the response of the model to end- and side-flanked configurations, there is a closer relationship between errors and the identity of end flankers than there is between errors and side flankers (compare colors of flankers and reports in Figures 7a and 7b). Figure 7c illustrates how random noise can lead to errors that are inconsistent with the flanker. Essentially, this arises when early positional noise switches the orientation of the flanker from (say) left- to right-facing; if the result of averaging this noisy flanker with the target resembles the *noise-corrupted* flanker, this will lead to a target report that is inconsistent with the *original* flanker configuration. The presence of substantial positional noise means that such errors are frequent in all (side- and end-flanked) conditions, but the lack of structured errors with side flankers means that these are all we observe in our maps (Figures 4 and 5).

The success of this model depends on three key elements: First, *positional averaging*. Building on early work with crowded orientation judgements (Parkes et al., 2001), it is clear that averaging has a central role in crowding, whether in the domain of stimulus attributes that can be read-off directly from visual filters (e.g., orientation) or from feature attributes such as spatial position (Greenwood et al., 2009). By building on our earlier modeling of interactions of feature positions in letter-like stimuli, the PAPA model is able to account for aspects of the data that are both consistent (e.g., target reports resembling the flankers) and are inconsistent with flanker substitution (e.g., infrequency of 180° target rotations among errors). What differs from this earlier work (Greenwood et al., 2009) is the second key element of the PAPA model: *elongated interference zones*. The shape of these zones would promote extended contours among both collinear ("snake") and parallel ("ladder") elements, thus regularizing oriented structure across the visual field. This notion is consistent with a growing consensus (Livne & Sagi, 2007; May & Hess, 2007) that the processes that produce crowding may be the same as those that limit contour integration in the periphery (Hess & Dakin, 1997).

Probabilistic crowding

The third novel and likely most controversial aspect of the model is that it is, at its core, *probabilistic*. Interference zones determine both the strength of crowding and the probability that crowding will occur. That is, close proximity between a target and flanker both increases the likelihood of the flanker interfering with the target and leads to stronger interference when it arises. Consequently, crowding appears to have an "all-or-nothing" element that, to our knowledge, has yet to be directly tested psychophysically. There are, however, two existing lines of evidence bearing on this proposition. The first comes from Greenwood et al. (2009) who examined the effect of cross-like flankers on the perception of features within a cross-like target. Figure 8 summarizes findings from the first experiment of this study. The black symbols plot the probability that a subject (JAG) reported that the horizontal bar of a target was displaced upward, as a function of the bar's physical offset. The resulting psychometric function is centered on a bar offset of zero, indicating the subject was unbiased, i.e., equally likely to report that a centered bar was offset upward or downward. Crowding with a pair of cross-like flankers (gray symbols) made the psychometric function shallower (i.e., threshold increased) but left performance unbiased. Adding flankers with the bar offset upward (red symbols) or downward (blue symbols) shifted psychometric functions left or right, respectively, inducing bias that was consistent with observers perceiving a target-bar position that was shifted in the direction of the flanker bar. Predictions of a weighted averaging model (Greenwood et al., 2009) and the PAPA model (lines) do a good job of capturing these results. If one plots JAG's bias under a variety of offsets of the flanker horizontal bar, we again see good agreement between data and models. It is only when one considers the dependence of *threshold* on flanker offset that one sees a discrepancy. Because standard weighted averaging rigidly averages flanker and target positions on every trial, it predicts that there should be no change in threshold with increasing magnitude of flanker-bar offset or, by extension, with increasing observer bias. PAPA, on the other hand, predicts a shallow U-shaped dependence of threshold on the size of the flanker-bar offset that we also observe in all subjects tested in Greenwood et al. (2009; Figure 1c). These results were obtained with essentially similar parameters of the model as used above except that the early positional noise (σ_{pos}) was raised from 0.25 to 0.32 (reflecting the greater target eccentricity in Greenwood et al., 2009) and w_{inter} (the strength of crowding) was reduced from 1.0 to 0.8.

The small number of data points in Figure 8c limits how strong a conclusion we should draw from this analysis, but it allows us to make a strong and rather general prediction from the PAPA model: that threshold will rise with increasing bias under crowding.

The second line of evidence bearing on this issue comes from Solomon et al. who have examined the perception of tilted Gabor elements under crowding and report that not only do tilted flankers cause horizontal targets to appear tilted but that this is accompanied by an elevation of orientation discrimination thresholds (Solomon et al., 2004; Solomon & Morgan, 2006, 2009). They note that this increase is inconsistent with simple averaging of target and flanker orientations and instead propose that "acuity losses can be explained if the extent of recalibration were to fluctuate from trial to trial, leading in effect to a fluctuating bias, which would be indistinguishable from an increase in sensory noise" (Solomon &

Morgan, 2006). Our results accord closely with this suggestion. The stochastic element of a PAPA model applied to orientation would result in tilted flankers crowding on some proportion of trials, leading to exactly such a "fluctuating bias" and corresponding elevation of threshold (Solomon & Morgan, 2009). Solomon and Morgan specifically show that the magnitude of threshold elevation should rise in proportion to the magnitude of the bias; this agrees with the behavior of the model in Figure 8c.

Low-level factors?

It is well known that decreased sampling density (e.g., of retinal ganglion cells) leads to greater intrinsic blur in the periphery (Levi & Klein, 1990). Since the phenomenon we are modeling is positional averaging, we wondered what role increased blur in the periphery might play in such averaging. To assess this, we ran a control experiment with one observer (SCD) who performed the original experiment (4AFC target-orientation identification) with inward (i.e., less eccentric) or outward (more eccentric) flankers in randomly interleaved conditions in the presence of variable amount of Gaussian blur applied to the stimulus.

Results are shown in Figure 9, circles are for less eccentric "in-flankers" and squares are for more eccentric out-flankers, and (a) confirm that "out-flankers" are more disruptive at low levels of blur (leading to higher size thresholds) than "in-flankers" and (b) show that blur starts having an effect on thresholds for in-flankers at ~10 arcmin and for out-flankers at ~23 arcmin (estimates are from the fits of a standard equivalent blur model). These results are essentially consistent with a scale-invariant mechanism; if blur is scaled by stimulus size, then both conditions start going up at the same blur level. Inset are depictions of the stimuli, at SCD's size threshold, with the estimated intrinsic blur applied to both. Note that these data set an *upper limit* on the amount of intrinsic blur; we have shown only that blur beyond this level would disrupt performance further, not that this is necessarily the intrinsic blur present at this eccentricity. Even so, this estimate of maximum possible intrinsic blur clearly cannot entirely account for the types of positional averaging we observe. Subjectively, individual strokes remain delineated from one another. That said, it is possible that intrinsic blur promotes some within-contour averaging (e.g., note that segregation of the vertical bars in the upper left inset may be compromised under blur). Importantly, however, intrinsic blur could not predict the infrequency of target inversions, nor the more disruptive influence of end versus side flankers (both of which are predicted by within-letter positional averaging).

Limitations of the approach

The current version of the model is limited in several respects. First, as alluded to in the Results section, the model predicts a higher density of errors with cross-like flankers (i.e., those plotted in the central region of our element space) than observed in human data. This could be accommodated by several means. For example, we have assumed that position is encoded linearly, so that shifting a feature by 0.1 deg produces the same change in its internal representation irrespective of its position within the target. A non-linear transformation that took into account the visibility of shifts (i.e., where small shifts from the center might not be encoded at all) could produce the desired reduction in the influence of cross-like flankers. Decision processes such as a Signed Max rule (Baldassi et al., 2006) could potentially produce similar non-linearities. This suggestion may be related to a second

limitation of the PAPA model: that it operates on noisy but essentially idealized estimates of spatial position, without specifying how these might be measured from images. There are several ways this could be achieved, including use of the centroid of the zero-bounded thresholded response of a spatial filter (Morgan & Glennerster, 1991; Watt, 1984) and incorporating some form of hierarchical representation from coarse to fine spatial scales (Watt, 1987, 1991). Alternatively, it may be possible to encode fine position using an estimate of instantaneous phase (Bennett & Banks, 1991) by taking the arctangent of the response ratio from a quadrature pair of filters. The particular problem for a position code based on local phase estimates is what origin phase estimates are made relative to. Although the use of our cross-like stimuli bypasses the issue of how these feature positions are encoded—since position is necessarily represented relative to the stimulus dimensions—it may be that the way in which the visual system solves this contributes to the pattern of errors observed in crowding. For example, an early positional code based on spatial phase might emphasize local position estimates that are consistent with cardinal phases (i.e., 0°, 90°, 180°, or 270°) leading to a bias toward seeing bar positions that are either centered (i.e., in even phase) or at extreme right, left, up, or down positions within the letter (i.e., odd phases).

A third limit of the model as proposed is that two-dimensional position is encoded strictly as a pair of independent one-dimensional positional estimates, each derived from features at a particular orientation. Although this could emerge naturally from a positional code operating on the output of oriented receptive fields in V1 (e.g., the phase-encoding scheme described above), we are unaware of direct psychophysical evidence showing that our ability to make two-dimensional judgements of position is limited by one-dimensional positional coding. We are currently undertaking such experiments. This also bears on the greater question of the fundamental unit upon which crowding operates—i.e., whether these interactions take place between the line-stroke features of letters or the elements as a whole. Though Martelli, Majaj, and Pelli (2005) convincingly demonstrate that crowding operates between the letters of a word, their assertion that crowding cannot operate between letter features relies on the observation that crowded letter recognition can be restored to uncrowded levels with sufficient letter spacing. This assumes, however, that uncrowded performance does not already have some degree of self-crowding between the letter features. The success of the PAPA model at modeling observers' apparently complex patterns of errors on a two-dimensional judgement leads us to be cautiously optimistic that crowding is operating at the level of one-dimensional positional encoding (i.e., between letter strokes) and that such interactions are occurring within the output of independent orientation channels.

It should be clear from our description above and in preceding sections that the proposed model is specifically designed to deal with interactions of feature position operating within the output of single orientation bands. We are currently working to extend the model to deal with simultaneous estimates of orientation and position (as a further step to a more complete model of crowding within letters). That said, this does not pretend to be a general model of object-centered crowding. We note that other object-centered effects, such as "holistic crowding" between Mooney faces, are specifically designed to minimize the role of low-level feature interactions by using stimuli that are challenging to recognize from individual features even in isolation. While we make no specific claims that the approach adopted here

can be generalized to, e.g., face crowding, we do note that one way in which faces are special is that they seem to convey a disproportional amount of relevant information through horizontal orientation structure (Dakin & Watt, 2009). It would be interesting to know how much this structure (compared to other orientation information) is being degraded by crowding. The proposal would then be that although crowding of faces does not operate on the level of simple interactions between features it may operate on representations of primitive facial information that form the input to face recognition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Baldassi S, Megna N, Burr DC. 2006; Visual clutter causes high-magnitude errors. PLoS Biology. 4:e56. [PubMed: 16494527]

Bennett PJ, Banks MS. 1991; The effects of contrast, spatial scale, and orientation on foveal and peripheral phase discrimination. Vision Research. 31:1759–1786. [PubMed: 1767496]

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

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

Bex PJ, Simmers AJ, Dakin SC. 2003; Grouping local directional signals into moving contours. Vision Research. 43:2141–2153. [PubMed: 12855250]

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

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

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

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

Dakin SC, Cass JR, Bex PJ, Watt RJ. 2009; Dissociable affects of crowding and attention on orientation averaging. Journal of Vision. 9(11):28, 1–16. DOI: 10.1167/9.11.28

Dakin SC, Hess RF. 1999; Contour integration and scale combination processes in visual edge detection. Spatial Vision. 12:309–327. [PubMed: 10442516]

Dakin SC, Watt RJ. 2009; Biological "bar codes" in human faces. Journal of Vision. 9(4):2, 1–10. DOI: 10.1167/9.4.2

Farzin F, Rivera SM, Whitney D. 2009; Holistic crowding of Mooney faces. Journal of Vision. 9(6):18, 11–15. DOI: 10.1167/9.6.18

Field DJ, Hayes A, Hess RF. 1993; Contour integration by the human visual system: Evidence for a local "association field". Vision Research. 33:173–193. [PubMed: 8447091]

Flom MC, Heath GG, Takahashi E. 1963; Contour interaction and visual resolution: Contralateral effects. Science. 142:979–980. [PubMed: 14069233]

Gheri C, Baldassi S. 2008; Non-linear integration of crowded orientation signals. Vision Research. 48:2352–2358. [PubMed: 18723044]

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

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

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

Hess RF, Dakin SC. 1997; Absence of contour linking in peripheral vision. Nature. 390:602–604. [PubMed: 9403687]

Hess RF, Hayes A. 1994; The coding of spatial position by the human visual system: Effects of spatial scale and retinal eccentricity. Vision Research. 34:625–643. [PubMed: 8160382]

Huckauf A, Heller D. 2002; What various kinds of errors tell us about lateral masking effects. Visual Cognition. 9:889–910.

Jazayeri M, Movshon JA. 2007; A new perceptual illusion reveals mechanisms of sensory decoding. Nature. 446:912–915. [PubMed: 17410125]

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

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

Levi DM, Klein SA. 1986; Sampling in spatial vision. Nature. 320:360–362. [PubMed: 3960118]

Levi DM, Klein SA. 1990; Equivalent intrinsic blur in spatial vision. Vision Research. 30:1971–1993. [PubMed: 2288101]

Livne T, Sagi D. 2007; Configuration influence on crowding. Journal of Vision. 7(2):4, 1–12. DOI: 10.1167/7.2.4

Louie EG, Bressler DW, Whitney D. 2007; Holistic crowding: Selective interference between configural representations of faces in crowded scenes. Journal of Vision. 7(2):24, 21–11. DOI: 10.1167/7.2.24 [PubMed: 18217839]

Martelli M, Majaj NJ, Pelli DG. 2005; Are faces processed like words? A diagnostic test for recognition by parts. Journal of Vision. 5(1):6, 58–70. DOI: 10.1167/5.1.6

May KA, Hess RF. 2007; Ladder contours are undetectable in the periphery: A crowding effect? Journal of Vision. 7(13):9, 1–15. DOI: 10.1167/7.13.9

Morgan MJ, Glennerster A. 1991; Efficiency of locating centres of dot-clusters by human observers. Vision Research. 31:2075–2083. [PubMed: 1771793]

Nandy AS, Tjan BS. 2007; The nature of letter crowding as revealed by first- and second-order classification images. Journal of Vision. 7(2):5, 1–26. DOI: 10.1167/7.2.5

Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. 2001; Compulsory averaging of crowded orientation signals in human vision. Nature Neuroscience. 4:739–744. [PubMed: 11426231]

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

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

Pelli DG, Tillman KA. 2008; The uncrowded window of object recognition. Nature Neuroscience. 11:1129–1135. [PubMed: 18828191]

Petrov Y, Popple AV. 2007; Crowding is directed to the fovea and preserves only feature contrast. Journal of Vision. 7(2):8, 1–9. DOI: 10.1167/7.2.8

Rauber HJ, Treue S. 1998; Reference repulsion when judging the direction of visual motion. Perception. 27:393–402. [PubMed: 9797918]

Rentschler I, Treutwein B. 1985; Loss of spatial phase relationships in extrafoveal vision. Nature. 313:308–310. [PubMed: 3969141]

Saarela T, Herzog M. 2009; Crowding in multi-element arrays: Regularity of spacing [Abstract]. Journal of Vision. 9(8):1017, 1017a.doi: 10.1167/9.8.1017

Solomon JA, Felisberti FM, Morgan MJ. 2004; Crowding and the tilt illusion: Toward a unified account. Journal of Vision. 4(6):9, 500–508. DOI: 10.1167/4.6.9

Solomon JA, Morgan MJ. 2006; Stochastic recalibration: Contextual effects on perceived tilt. Proceedings of the Royal Society B: Biological Sciences. 273:2681–2686.

Solomon JA, Morgan MJ. 2009; Strong tilt illusions always reduce orientation acuity. Vision Research. 49:819–824. [PubMed: 19268684]

Strasburger H. 2005; Unfocused spatial attention underlies the crowding effect in indirect form vision. Journal of Vision. 5(11):8, 1024–1037. DOI: 10.1167/5.11.8

Strasburger H, Harvey LO Jr, Rentschler I. 1991; Contrast thresholds for identification of numeric characters in direct and eccentric view. Perceptions & Psychophysics. 49:495–508.

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

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

van den Berg R, Roerdink JB, Cornelissen FW. 2007; On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. Journal of Vision. 7(2):14, 11–11. DOI: 10.1167/7.2.14

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

Watt RJ. 1984; Towards a general theory of the visual acuities for shape and spatial arrangement. Vision Research. 24:1377–1386. [PubMed: 6523758]

Watt RJ. 1987; Scanning from coarse to fine spatial-scales in the human visual system after the onset of the stimulus. Journal of the Optical Society of America. 4:2006–2021. [PubMed: 3430211]

Watt, RJ. Understanding vision. London: Academic Press; 1991.

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

Wolford G. 1975; Perturbation model for letter identification. Psychology Reviews. 82:184–199.

Yu, D; Akau, MU; Chung, STL. The mechanism of word crowding. Poster was presented at VSS 2010: The Vision Science Society 2010 Annual Meeting; Naples, Florida. 2010.
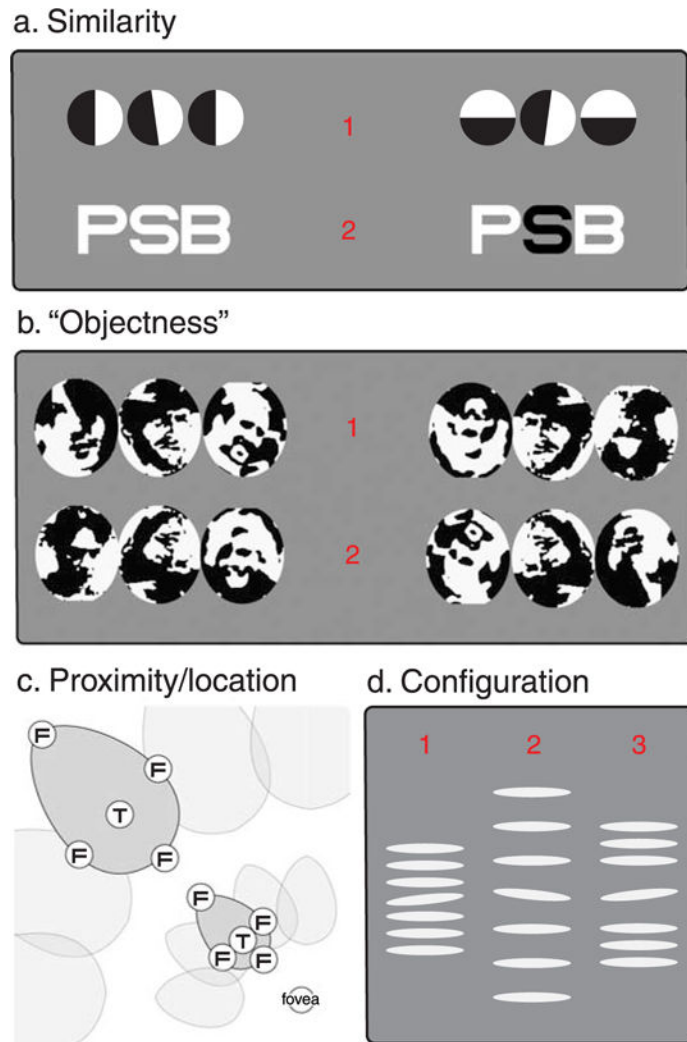
Figure 1.
Determinants of crowding. (a) The more similar objects are, in terms of their low-level feature properties, the more they crowd one another. Fixating the red numbers in turn, it is harder to recognize the central element of the triplets on the left than on the right. (b) For objects, where local similarity has been factored out, upright flanker faces crowd more than inverted flankers, irrespective of target orientation (Mooney faces after Farzin et al., 2009). Again, targets on the right should be easier to identify then targets on the left. (c) "Interference zones" for crowding. The region within which flankers ("F") interfere with identification of a target ("T") depends on target and flanker locations in the visual field. (d) Configuration effects (after Saarela & Herzog, 2009). Fixating "1" the identity of the central element is lost (objects are similar, and close). Increasing separation helps a little (fixate "2") but fixating "3" the central identity is clear, even though the target and its immediate neighbors have not changed. That is, global configuration can also affect crowding.
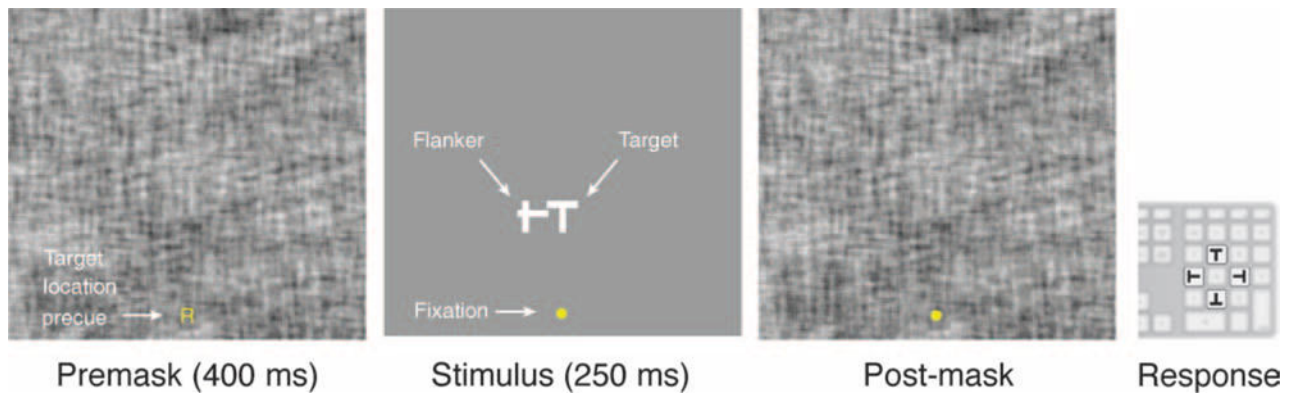
Figure 2.
Time course of a trial. Subjects fixated on the yellow marker, which for the first 400 ms was a pre-cue indicating whether the target would be right, left, above, or below the flanker element. The stimulus—a T in one of four (N, S, E, W) orientations abutted by a flanker— was displayed for 250 ms, followed by a post-mask. Subjects indicated the perceived orientation of the target using the keyboard.
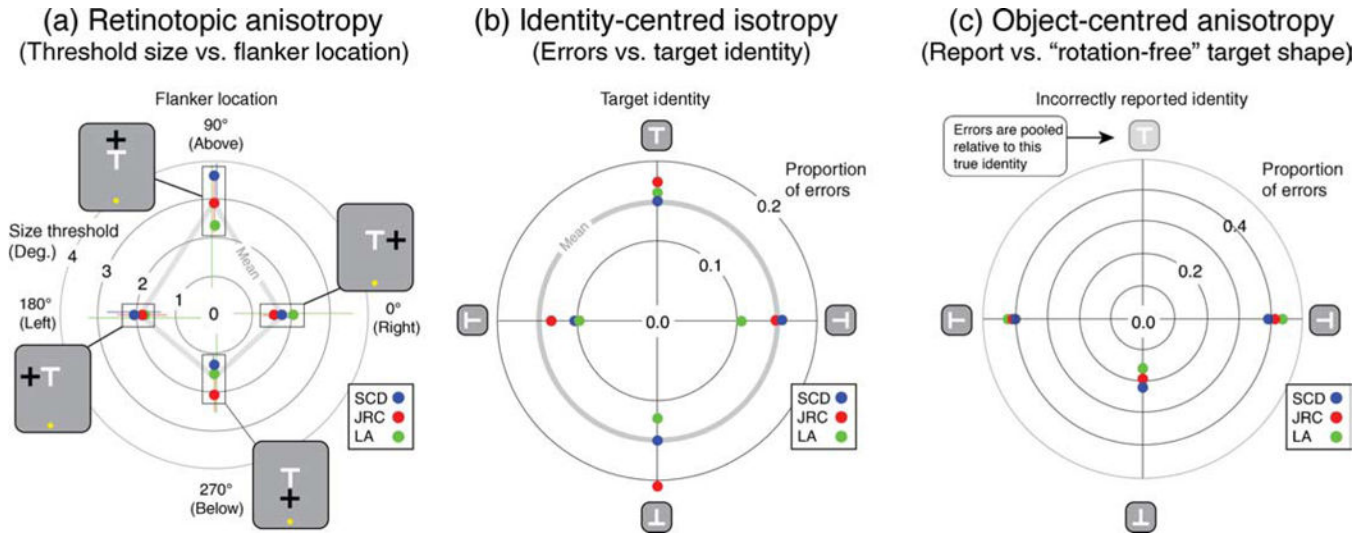
Figure 3.
(a) Threshold stimulus size as a function of the location of the flanker with respect to the target. Note that the more eccentric flanker interferes at a greater target separation (cf. Figure 1b). Subjects (b) do not make more errors with any given target but (c) are twice as likely to report a 90° compared to a 180° rotation of the target, when they do make mistakes.
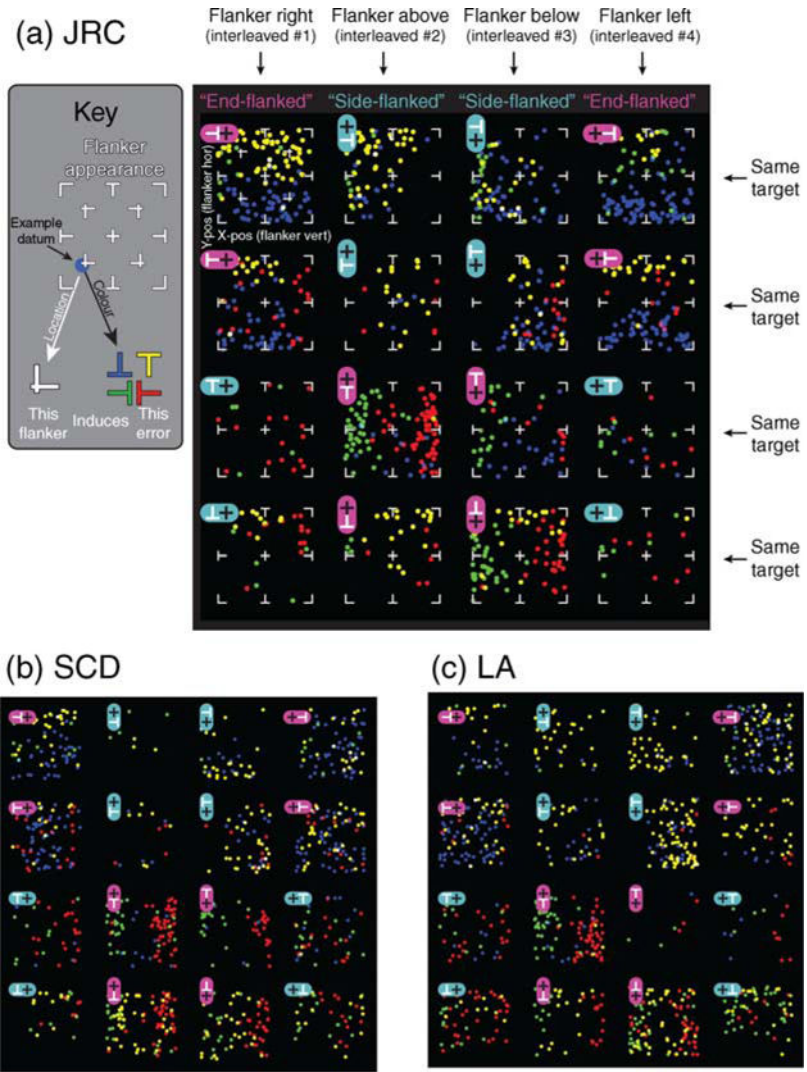
Figure 4.
(a) Each of the 4 × 4 panels plots the errors arising for a given flanker–target configuration (a schematic stimulus is inset in each panel). Dot color codes the target orientation reported by the subject, and the position of the dot within the plot codes the flanker configuration on that trial (see "Key"). Note lower error rates in cyan- compared to magenta-labeled panels, corresponding to *side-flanked* and *end-flanked* conditions, respectively. This refers to the flanker location relative to the target after the whole stimulus was rotated so the target was an upright T. Note also that errors arising from end flankers (magenta-labeled) correlate with flanker appearance, whereas errors from side flankers (cyan-labeled) are less structured.
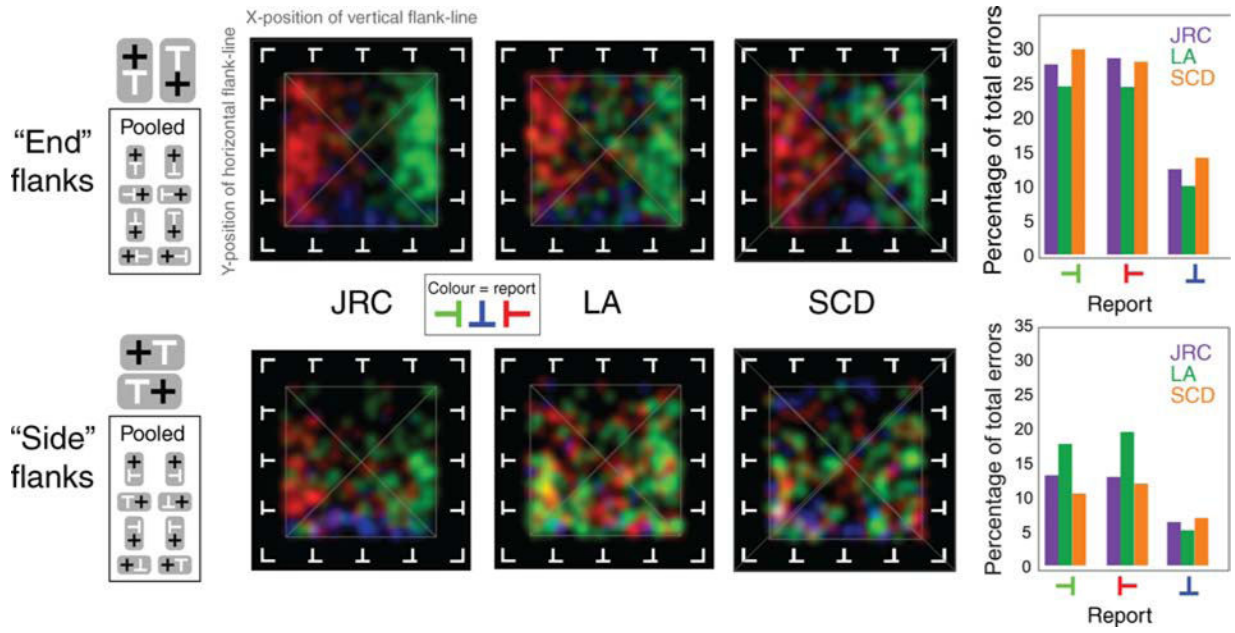
Figure 5.

Errors from the three subjects have been rotated so that they are relative to an upright "T" and pooled within end- and side-flanked conditions. Flanker configuration strongly determines subjects' report in end-flanked conditions but less so in side-flanked conditions. Bar graphs indicate that (a) error rates are approximately twice as high in end- compared to side-flanked conditions and (b) errors arising from 180° target rotations are only half as frequent as reports of 90° rotation. In subsequent discussion of these findings, it is important to recall that these are *object-centered* results; end-flanking can arise with any absolute target orientation, or with any absolute configuration of target and flanker.
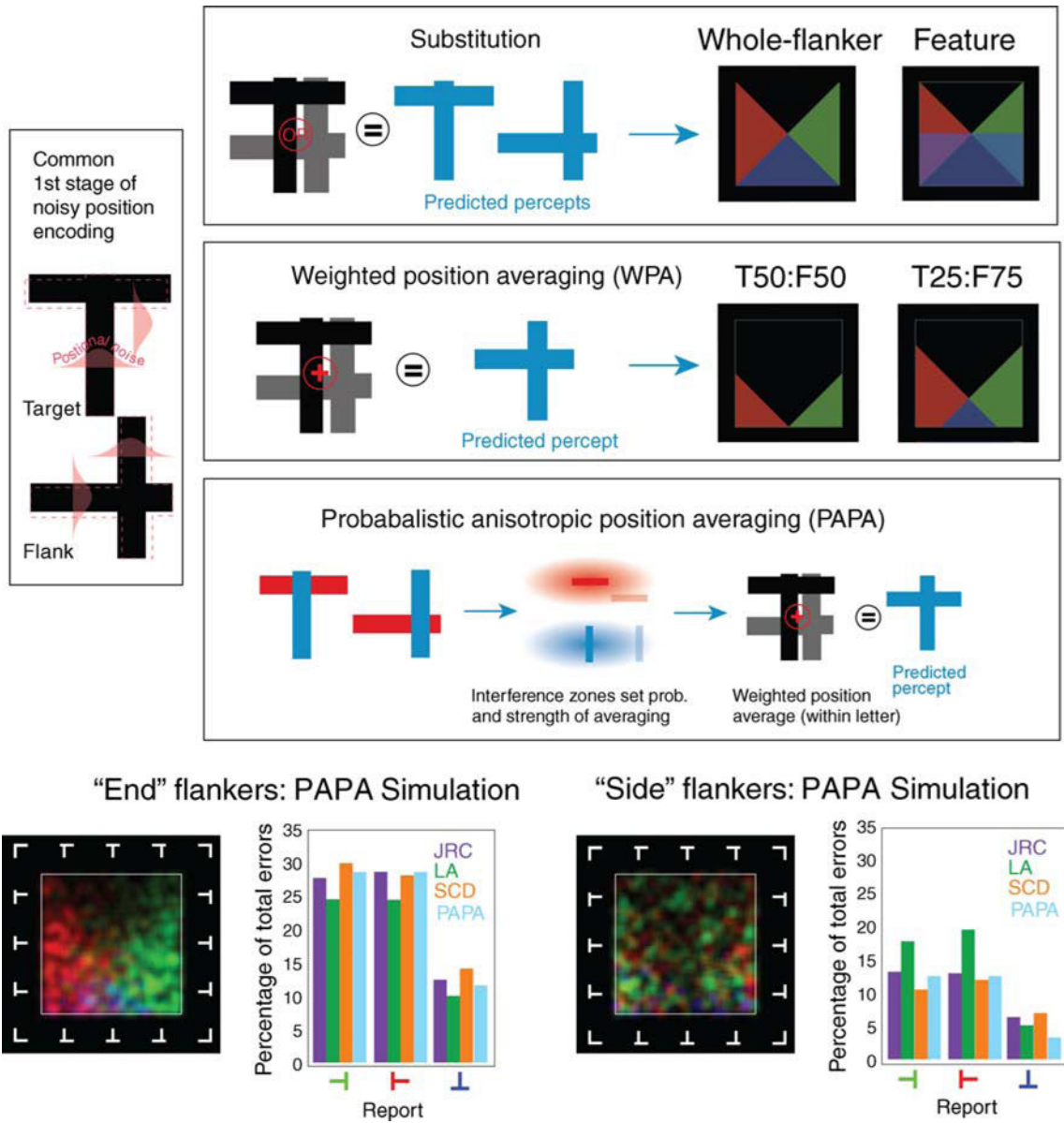
Figure 6.
Three models for predicting the errors induced by crowding. All share a first stage where the location of the features of target and flankers are perturbed with Gaussian position noise. Substitution models then propose that either the whole flanker or its features randomly replace the target or its features. Our data rule out this model since it can predict neither the infrequency of 180° target rotations among errors nor differences between end- and side-flanking. The second model proposes that the locations of crowded target features are the weighted positional average (WPA) of the target and flanker feature positions. This model can predict the infrequency of 180° target inversions but not end- versus side-flanker differences. The third, probabilistic anisotropic position averaging (PAPA), is similar but incorporates an additional stage, between positional encoding and averaging, that uses

elongated interference zones to set the *probability* and weighting of averaging. Its behavior (bottom panel) closely matches that of the subjects.
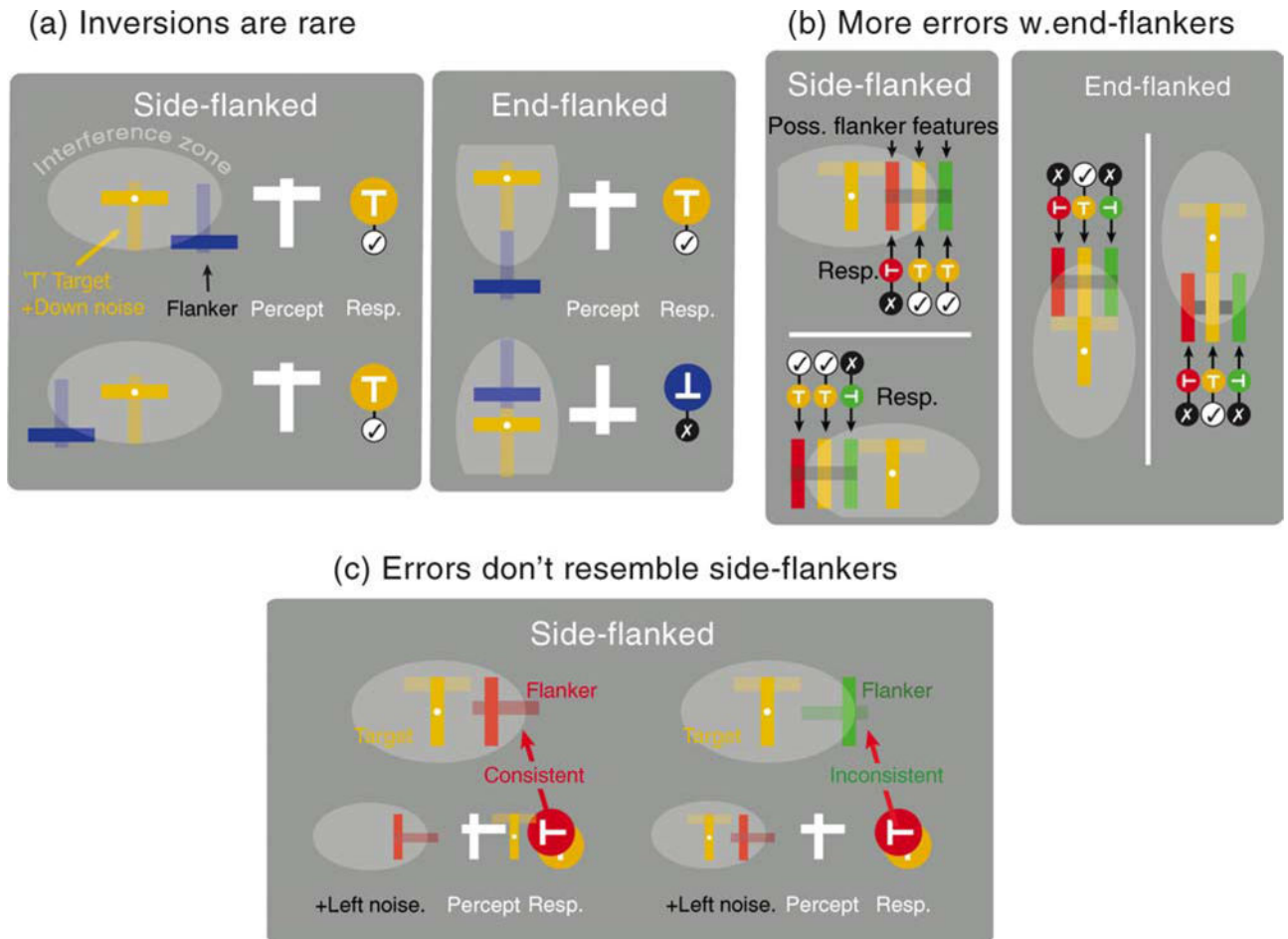
Figure 7.
Why does the PAPA model work? (a) Under averaging of target and flanker feature positions, inversions of an upright T target can only arise when the horizontal stroke of the flanker falls well below its center. Such features are unlikely to fall within the elongated interference zones surrounding the target, making target inversions rare. (b) For an upright T target, most erroneous reports are ±90° rotations. Elongation of the interference zones means that bigger offsets of the vertical flanker feature (which drives ±90° target rotations) always interfere with end flankers but can fall outside the interference zone with side flankers (leading to fewer errors). In both (a) and (b) note the stronger correlation of flanker structure with response category in end- but not side-flanked conditions. This means that proportionally more of the erroneous reports in side-flanked conditions are due to random positional noise. (c) Early positional noise can lead to errors that are inconsistent with the flanker. In the left condition, a leftward-pointing flank has left noise added: the resultant average of target and flanker is consistent with the flanker. In the right example, a rightward-pointing flanker is corrupted by noise to become left-facing. Under averaging this leads to a leftward report that is inconsistent with the original flanker orientation.
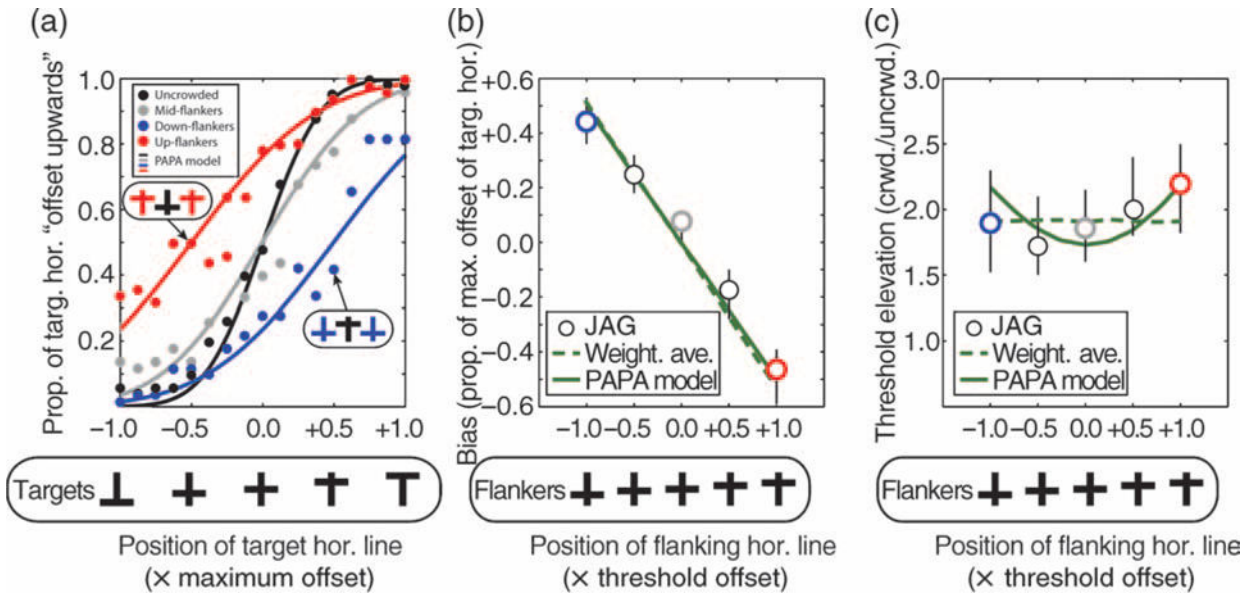
Figure 8.

(a) Proportion of times observer JAG reported that the horizontal target bar was offset upward, as a function of the offset of the bar (black symbols; data from Greenwood et al., 2009). Crowding with a mid-flanker (gray symbols) disrupts performance, but observers remain unbiased; flankers with upward-displaced horizontal bars (red symbols) both disrupt and bias performance, making observers more likely to report that the target horizontal was offset in the same direction as the flanker bar. Predictions from the PAPA model (lines) accord well with these results. (b) Bias and (c) threshold elevation (crowded/uncrowded) derived from similar psychometric functions for various offsets of the horizontal bar of the flanker (red and blue symbols are from corresponding psychometric functions in (a)). The pattern of JAG's bias (b) is consistent with both weighted positional averaging (WPA, Greenwood et al., 2009) and the PAPA model. (c) However, WPA predicts strict independence of threshold and offset whereas PAPA correctly predicts a modest U-shaped dependence of threshold on offset.
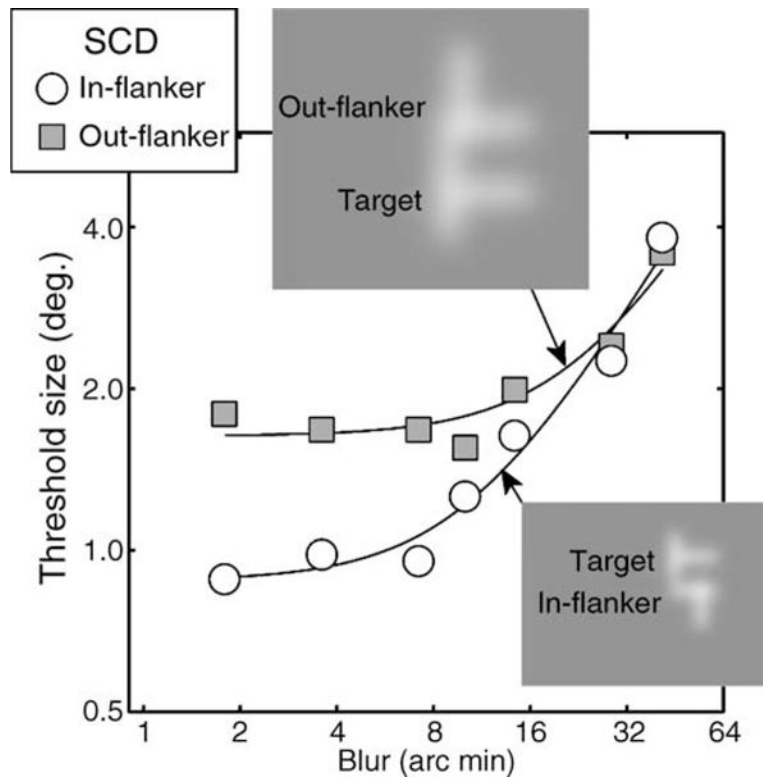
Figure 9.
The effect of Gaussian blur on T-orientation identification in the presence of out-flankers (circles) and in-flankers (squares). Note poorer performance with out-flankers (as Figure 1c) at low levels of blur. Note also that more blur is required to impact on performance with out-flankers than with in-flankers. This is unsurprising given that better performance led to smaller stimuli in the in-flanker condition so that the same blur would be expected to exert more influence in a scale-invariant mechanism. Insets show stimuli, at threshold, with levels of intrinsic blur indicated by the arrows.