

Crowding with conjunctions of simple features

Endel Pöder

Laboratory of Experimental Psychology,
University of Leuven, Leuven, Belgium, &
Department of Psychology, University of Tartu,
Tartu, Estonia



Johan Wagemans

Laboratory of Experimental Psychology,
University of Leuven, Leuven, Belgium



Several recent studies have related crowding with the feature integration stage in visual processing. In order to understand the mechanisms involved in this stage, it is important to use stimuli that have several features to integrate, and these features should be clearly defined and measurable. In this study, Gabor patches were used as target and distractor stimuli. The stimuli differed in three dimensions: spatial frequency, orientation, and color. A group of 3, 5, or 7 objects was presented briefly at 4 deg eccentricity of the visual field. The observers' task was to identify the object located in the center of the group. A strong effect of the number of distractors was observed, consistent with various spatial pooling models. The analysis of incorrect responses revealed that these were a mix of feature errors and mislocalizations of the target object. Feature errors were not purely random, but biased by the features of distractors. We propose a simple feature integration model that predicts most of the observed regularities.

Keywords: crowding, features, conjunctions, localization, feature integration

Citation: Pöder, E., & Wagemans, J. (2007). Crowding with conjunctions of simple features. *Journal of Vision*, 7(2):23, 1–12, <http://journalofvision.org/7/2/23/>, doi:10.1167/7.2.23.

Introduction

Perception of a visual object can be impaired when there are other objects nearby in the visual field (e.g., Bouma, 1970; Strasburger, Harvey, & Rentschler, 1991). This is known as the crowding effect. In the fovea, only a small crowding effect is usually observed. In the periphery, the spatial extent of crowding is approximately proportional to the eccentricity of the target object and reaches $0.5E$ (E —eccentricity of target) (Bouma, 1970; Toet & Levi, 1992). It seems that detection of simple visual features is very little affected by crowding (Andriessen & Bouma, 1976; Levi, Hariharan, & Klein, 2002; Pelli, Palomares, & Majaj, 2004), and the main mechanism of adverse interaction must therefore be located at some level after feature detection.

The exact nature of crowding, in a computational sense, is not clear. Theoretical models of crowding are based on an interaction between feature detectors (e.g., Bjork & Murray, 1977; Wolford & Chambers, 1984), or on pooling (integration) the signals of feature detectors over some larger area (e.g., Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli et al., 2004). Intriligator and Cavanagh (2001) have argued that crowding is an effect of insufficient spatial resolution of attention which is limited by the large size of the receptive fields at some higher level of visual processing.

Several authors have suggested that some sort of positional noise may account for the crowding effects

(Neri & Levi, 2006; Wolford, 1975). According to this account, visual features are incorrectly localized, or can migrate to incorrect locations, thereby producing errors of object perception. Simple pooling models assume a full loss of positional information within a certain field of integration. Parkes et al. (2001) found that orientation discrimination thresholds in crowding conditions could be accounted for by a simple averaging of feature (orientation) values over a certain region. With similar stimuli, Baldassi, Megna, and Burr (2006) obtained results that were not consistent with averaging and suggested that some other (non-linear) combination rule had to be used to model their data. Anyway, it is not clear how to extend these simple pooling rules to more complex stimuli with a number of features.

Pooling (integration) models should predict a strong effect of the number of distractors within the integration area, because with a larger number of distractors, the target signal is diluted in a larger amount of irrelevant activity. The results obtained with simple stimuli (Parkes et al., 2001) are consistent with this prediction. With letters, the results are ambiguous: Strasburger et al. (1991) found relatively strong effect of the number of flankers; Pelli et al. (2004) found no difference between 2 and 4 flankers.

Many studies with alphanumeric characters have found that flanking objects are reported frequently instead of the target (e.g., Eriksen & Rohrbaugh, 1970; Huckauf & Heller, 2002; Strasburger, 2005). These studies suggest

that positions of integrated objects rather than these of features are perceived incorrectly. Strasburger (2005) explains this by imprecise focusing of attention. Some studies have suggested that object identification and localization errors are differentially affected by target–flanker distance (Butler & Currie, 1986) and exposure duration of stimuli (Styles & Allport, 1986). With alphanumeric characters, it is difficult to separate the roles of features and their combinations, because the relevant features are largely unknown. Wolford and Shum (1980) used specially designed stimuli and reported some support for the idea that object localization errors are mediated by different (higher level) mechanisms as compared with feature localization errors.

There are different ideas about the effects of target–flanker similarity. Some early studies, using alphanumeric characters, found more degradation of performance with more similar flankers and argued for an interaction between mechanisms sensitive to similar features (e.g., Bjork & Murray, 1977). Estes (1982), however, demonstrated that these effects could be explained by the criterion shifts induced by flankers. He found a strong bias towards the target similar or identical to the flankers, but discriminability of targets was little if at all affected by the target–flanker similarity. Also, he found that position errors were more frequent when target–flanker similarity was high.

Several recent studies have demonstrated that pop-out of the target with a unique visual feature can reduce the crowding effect (e.g., Felisberti, Solomon, & Morgan, 2005; Kooi, Toet, Tripathy, & Levi, 1994; Pöder, 2006). The role of salience in more complex conditions with multidimensional and/or heterogeneous stimuli is not clear.

Illusory conjunctions are perceptual errors when visual features from different objects are incorrectly combined. Pelli et al. (2004) drew attention to the similarity between conditions of crowding and these of illusory conjunctions. Both appear predominantly in the visual periphery, and when irrelevant objects are located not far from the target. While some early studies suggested that features can be incorrectly conjoined regardless of the distance between them (e.g., Treisman & Schmidt, 1982), the more recent studies have shown a strong effect of spatial proximity (e.g., Ashby, Prinzmetal, Ivry, & Maddox, 1996; Prinzmetal, Ivry, Beck, & Shimizu, 2002). The effect of target–distractor similarity is more controversial. Treisman and Schmidt (1982) found that feature migrations causing illusory conjunctions were independent of similarity of the target and distractor objects on the other feature dimensions. Ivry and Prinzmetal (1991) studied the migration of features when the similarity of objects on the same feature dimension was varied and they found more frequent illusory conjunctions with more similar features. Donk (1999),

however, reported an effect of similarity along other dimensions and no effect of similarity along the same dimension. In the studies of illusory conjunctions, different measures (dual tasks, backward masking) have been used in order to avoid focused attention to the target, and it is hard to say whether these findings are valid for more simple crowding displays.

The purpose of this study is to explore the role of features and their combinations in a simple crowding experiment. We use multidimensional stimuli with simple and well-defined visual features that can be combined to create a set of objects comparable to letters or numerals. We try to reveal the ways by which distractor objects and their features affect perception of the target, and whether this could be predicted by any simple model.

Methods

Examples of stimuli are depicted in Figure 1. The stimuli consisted of Gabor patches—cosine profile luminance gratings windowed by two-dimensional Gaussian (sigma 5 pixels or 0.15 deg, from a 60-cm viewing distance). The maximum luminance contrast was about 85%. The Gabors could vary independently in three feature dimensions—orientation, spatial frequency, and color. They could be either vertical or horizontal, either of low (3 cpd) or high (6 cpd) spatial frequency, and either red or green (for the red stimuli, the voltage of the red gun was increased by 25% and that of the green gun was reduced by the same amount; for green stimuli, the changes were opposite). The color and spatial frequency differences were chosen in order to approximately equate the probabilities of errors across the feature dimensions. All 8 feature combinations could occur in the stimuli.

The Gabor positioned in the center of a group was the target. It was selected randomly from the set of eight possible objects. The target was surrounded by 2, 4, or 6 flankers that were selected independently and with replacement from the remaining 7 objects. The flankers were located at the same distance from the target (0.8 deg, from center to center), in equal steps around it (the angular position of the first flanker was selected randomly). The combination of presentation parameters (target–flanker distance, eccentricity, exposure duration) was chosen to induce a nearly perfect performance without flankers and a strong crowding effect with six flankers, while avoiding a spatial overlap of the target and flankers.

The stimuli were presented on a gray background (with the luminance about 40 cd/m²). The luminance function of the monitor was approximately linearized using the gamma correction option of the video driver.

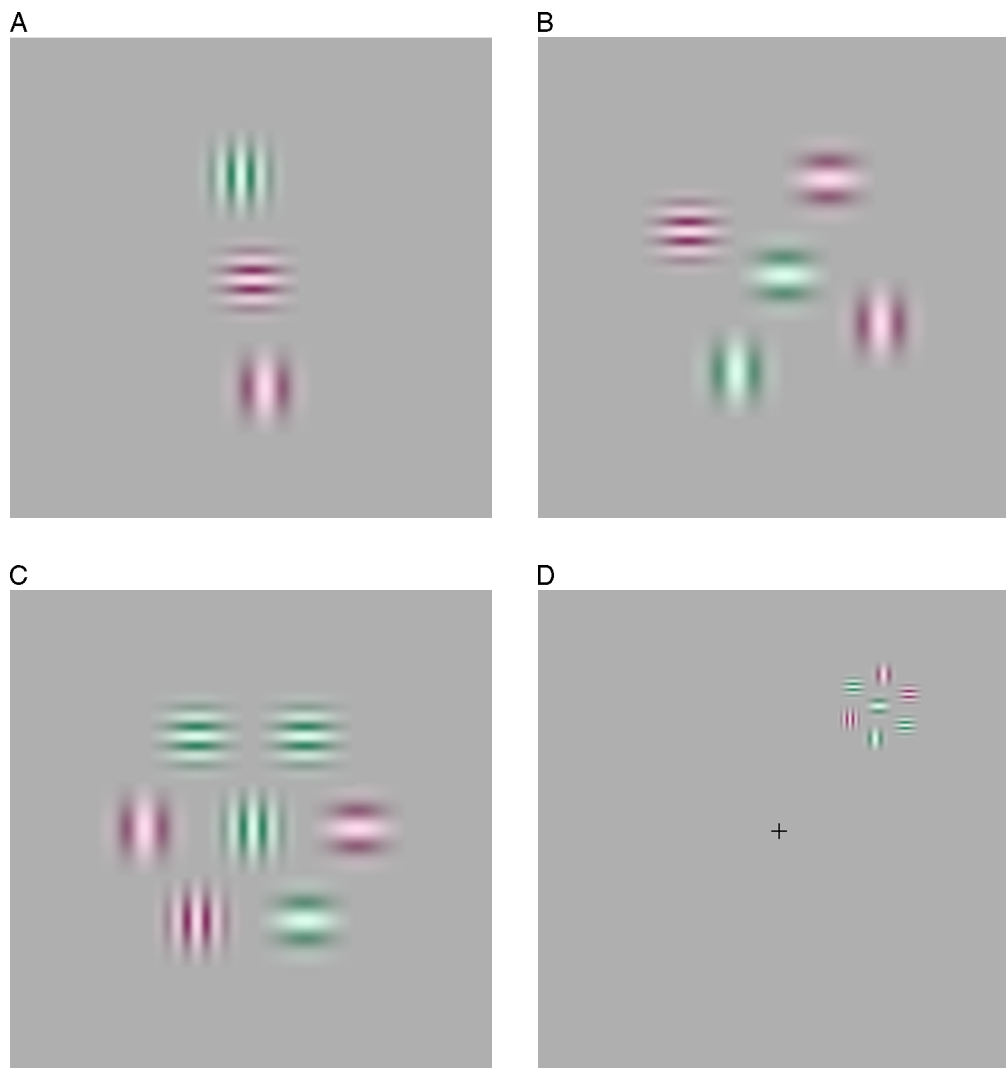


Figure 1. Examples of stimuli used in this study. A target Gabor with (A) two, (B) four, and (C) six flanking Gabors. (D) Stimuli were presented in random positions around the fixation point (eccentricity 4 deg).

On each trial, the target and its flankers were presented for 60 ms (for observer EP, 120 ms) at a constant radius of 4 deg, in random direction from the fixation point. The observer's task was to identify the central object, the target, and to indicate it by clicking an icon in the response panel (containing 8 alternatives), located in the monitor screen, below the stimulus presentation area. A feedback message informed whether the response was right or wrong.

Three observers took part in the experiment (one of the authors among them). They had normal or corrected to normal vision. The observers had no or very little experience with these particular stimuli, but had participated in similar experiments with a brief presentation of stimuli. Each of them ran 1000–1600 trials (EP 600, 500, and 500; both LP and SE 400, 200, and 400 trials with 2, 4, and 6 flankers, respectively). The number of flankers was held constant within a block of trials.

Results

Effect of the number of flankers

All three observers exhibited a strong effect of the number of flankers (see Figure 2). This result is obviously more or less consistent with various pooling models but appears to contradict Pelli et al. (2004), who found no difference between 2 and 4 flankers in a letter identification task.

Part of the effect of the number of flankers can be explained by the difference of crowding in the radial vs. tangential direction of the visual field (e.g., Toet & Levi, 1992). With two flankers, the performance with a radial configuration was, on average, 18% worse than a tangential configuration of flankers. But even in the most difficult radial condition the performance with two

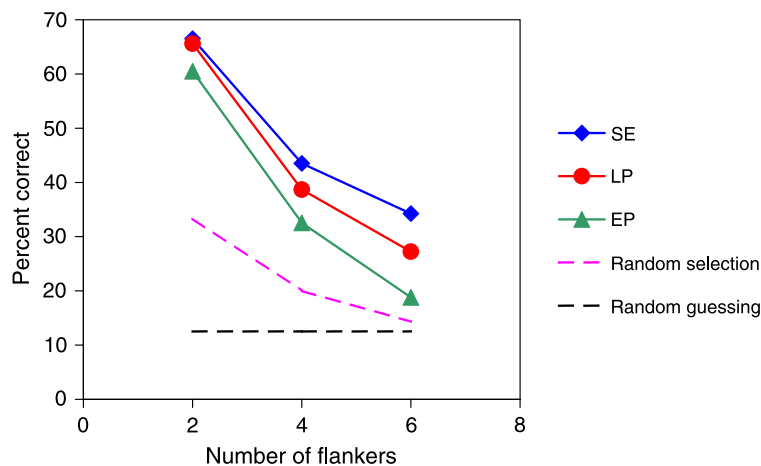


Figure 2. Performance as dependent on number of flankers for three observers. The level of random guessing and the prediction of simple random selection (full loss of spatial information) are indicated by dashed lines.

flankers (55% correct) was considerably better as compared with the 4-flanker condition (38% correct).

In Figure 2, together with random guessing, the prediction of a simple random selection model is indicated. This model assumes that we are unable to use positional information within the group of objects and select a “target” object randomly from all the objects displayed in a given trial. It is clear that the data cannot be predicted by this model, and that the assumption of a full loss of spatial information must be wrong.

Saliency

We tried several measures of saliency based on Euclidean and city-block distances in 3D feature space between the target and distractors and between the distractors themselves, including also Mahalanobis distance. However, we found that a very simple measure—the number of feature dimensions on which the target was unique among the distractors, captured most of the saliency effects on performance in our experiment. The effect was statistically significant ($p < 0.01$) for all three observers for the 2-flanker condition. For the 4-flanker condition, the correlations were significant ($p < 0.05$) for two observers, and nearly significant ($p = 0.08$) for the third. However, there was virtually no saliency effect with 6 flankers. These results (averaged across observers) are given in Figure 3.

An obvious reason for not finding a saliency effect with 6 flankers is the near absence (very low probability) of configurations with unique target features among our randomly generated stimuli. Even the effects with 2 and 4 flankers are based on relatively small numbers of trials with highly salient targets. Thus, the saliency cannot explain much of the total variance of our present data.

Actually, there is another process that may counterbalance the effect of saliency. We found that the effect of

the number of distractors with features that are identical to the target features tends to be non-monotonic: Performance can improve also when most of the distractors are similar to the target. A more detailed analysis revealed that percentage correct on any single feature dimension tends to increase with the number of distractors that are identical to the target on that dimension, and to decrease with the number of distractors that are identical to the target on the other dimensions. The first effect is relatively stronger and presumably caused by the feature pooling (or bias) or incorrect object selection (that we will discuss in the next part); the other may be related with saliency.

Analysis of errors

Usual crowding studies with letter recognition allow discriminating two types of errors only: misidentification

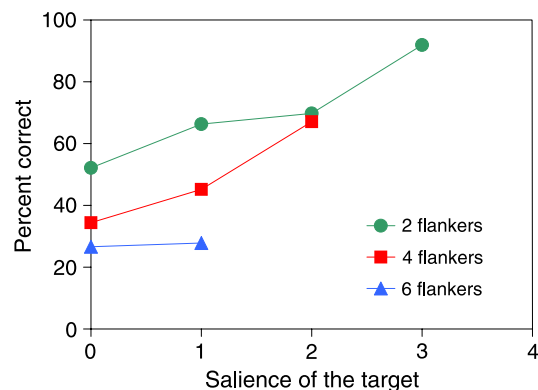


Figure 3. The effect of saliency of the target on performance. The number of feature dimensions on which the target was unique among the distractors is used as the measure of saliency. The results are averaged across the observers.

and mislocalization. Present multidimensional stimuli offer much more possibilities because each feature may be either right or wrong, and either presented or not among the distractors.

The distributions of responses across different combinations of feature errors are given in Figure 4. (We present the results averaged across observers, the individual

results were qualitatively similar.) It is clear that a majority of incorrect responses differ from the target by one feature only, and errors in all three feature dimensions are very rare. This pattern seems to show that incorrect responses are, to a large extent, generated by (independent) feature errors. Also, there are no large differences across different feature dimensions.

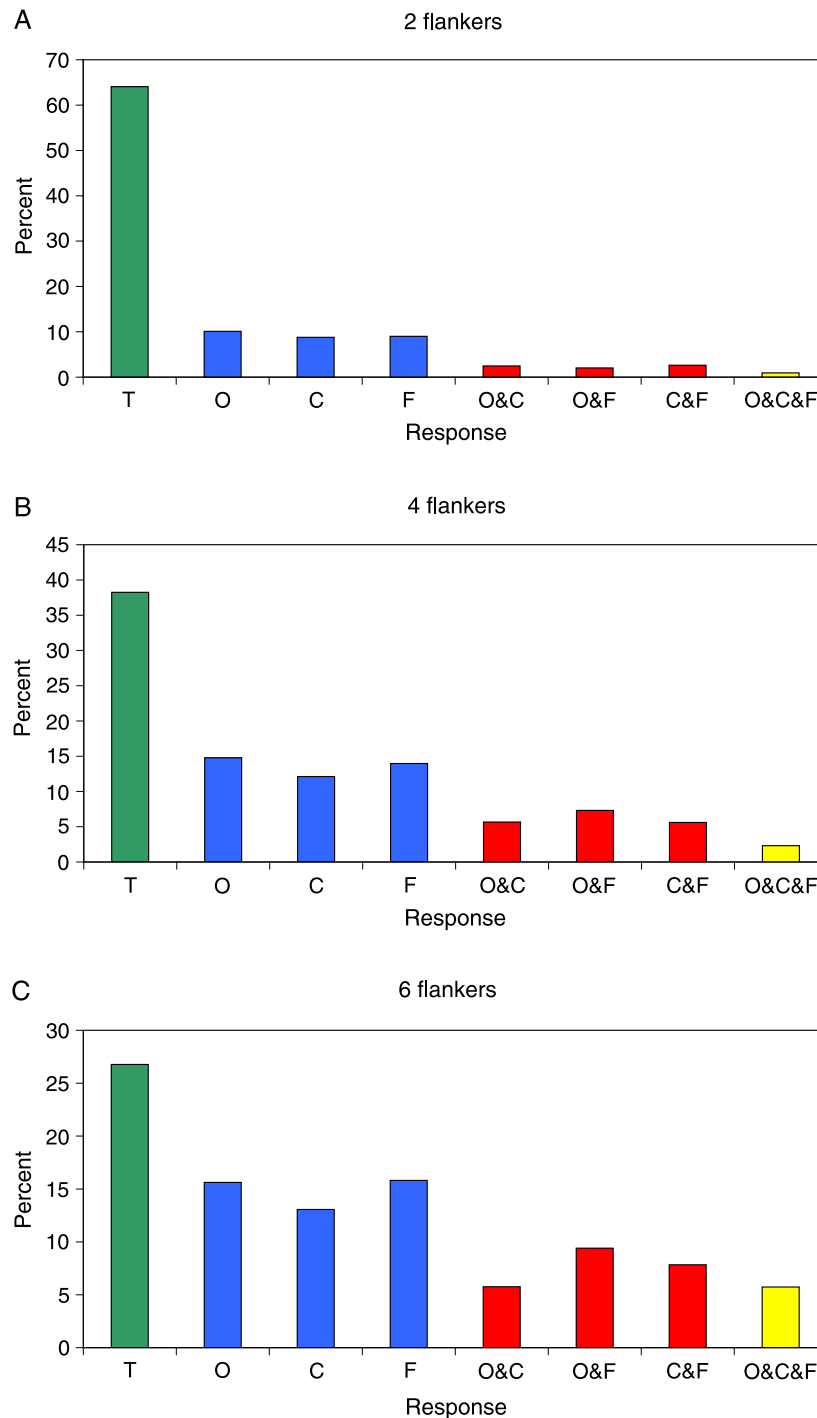


Figure 4. Distributions of responses across correct answer and different feature errors (T—correct answer; O, C, and F—error on one dimension only: orientation, color, or spatial frequency, respectively; O&C, O&F, C&F—errors on two dimensions; O&C&F—errors on all three dimensions). The average data of 3 observers are presented.

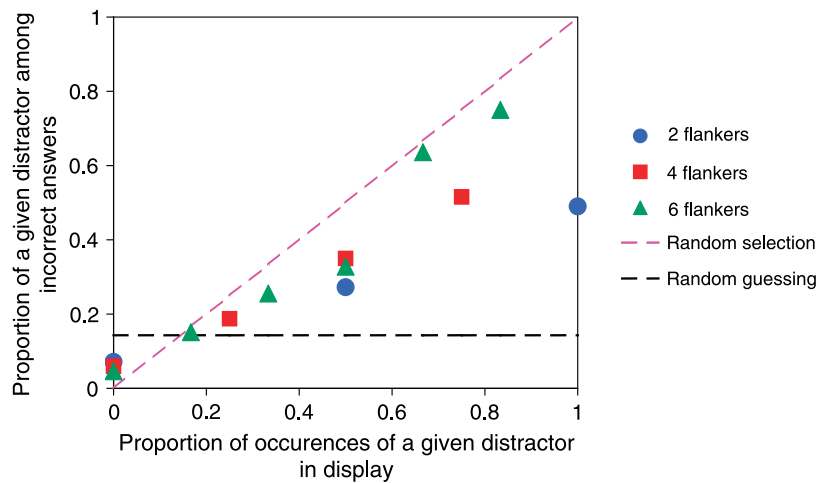


Figure 5. Probability of reporting a particular distractor instead of the target as predicted by the random object selection model and corresponding empirical data for the 2-, 4-, and 6-flanker conditions (pooled over three observers).

However, there is another side of the results as well. If incorrect responses were independent of flankers presented in a given trial, then the proportion of responses corresponding to the flankers relative to the total incorrect responses should be a simple function of the number of flankers (NF): $1 - (1 - 1/7)^{NF}$. According to this equation, the proportion of responses corresponding to the flankers among incorrect responses should be 0.26, 0.46, and 0.60 for 2, 4, and 6 flankers, respectively. Actual proportions were significantly larger for all numbers of flankers and for all three observers (averages 0.54, 0.70, and 0.82 for 2, 4, and 6 flankers). Consequently, there is a tendency to report flankers as incorrect responses (this is just an empirical regularity that may be caused by different mechanisms).

This aspect of the results can be demonstrated even more directly by using the fact that, in the present experiment, each particular object could occur more than

once among the flankers. Figure 5 plots the average proportion of any object reported among incorrect responses, as a function of the proportion of this object among the flankers in a given trial. There is a direct proportionality between the number of particular objects in a display and the probability of reporting this object instead of the target. Especially the incorrect answers for the 6-flanker condition followed this probabilistic model quite closely.

A report of a particular flanker instead of the target suggests that these objects (as conjunctions of the three features) are perceived correctly. However, an occurrence of a given object in a display is strongly correlated with the occurrence of its component features. Indeed, we found similar effects of flankers when analyzing each feature dimension separately. For example, with a larger number of green objects in a display, a green object was chosen as the response more frequently (Figure 6). A

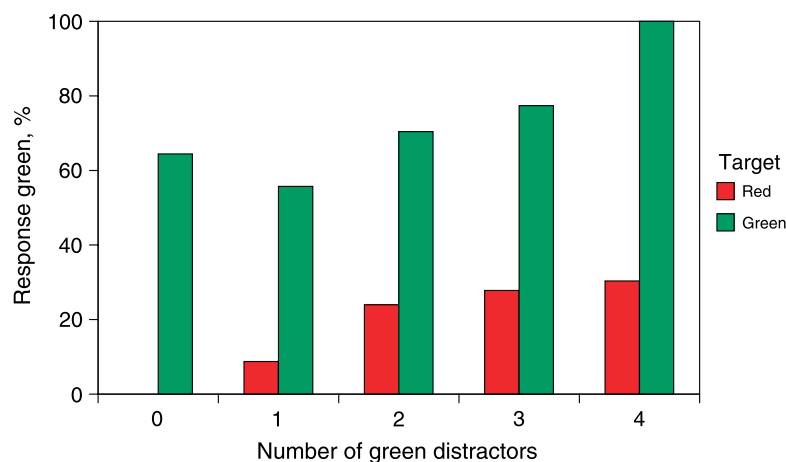


Figure 6. Probability of selection a green object for response as dependent on the target color and the number of flankers with green color. This example depicts the data for the 4-flanker condition, averaged across observers. Qualitatively similar results were observed for other dimensions, other numbers of flankers, and for individual observers.

Number of flankers	Observer	Object model	Feature model	Difference
2	EP	0.37	0.34	0.03
	LP	0.21	0.19	0.02
	SE	0.25	0.18	0.07
4	EP	0.30	0.24	0.06
	LP	0.24	0.25	-0.01
	SE	0.28	0.25	0.02
6	EP	0.31	0.22	0.09**
	LP	0.23	0.16	0.07**
	SE	0.22	0.19	0.03

Table 1. Correlations of occurrence of particular objects as incorrect response with number of occurrences of this object in a display (object model), and with prediction based on proportions of respective features in a display (feature model). In the last column, the difference between correlations of two models is given (**difference significant with $p < 0.01$).

combination of three simultaneous feature-based effects could mimic an object-based effect quite well.

We made an attempt to tell apart these accounts using a correlation analysis. For a feature-based model, we calculated the predicted proportions of a particular object among incorrect answers as the multiplication of the proportions of the respective features in a display. The analysis revealed that the number of particular objects in a display was a significantly better predictor of the response probabilities than the feature-based model for the 6-flanker condition only (Table 1). For 2 and 4 flankers, the differences were not significant, although mostly in the same direction.

These results are consistent with incorrect answers being a mix of object mislocations and random binding of the features from a given display. (By simulations, we verified that when responses are generated by one of these processes, then the correlations with that model should be systematically higher as compared with the alternative one.)

Modeling

The analysis of incorrect answers suggested that there may be several processes behind the observed response distributions. We attempted to study these mechanisms more quantitatively by fitting several simple models to our data. Basically, we assumed that there are two main sources of errors: report of a flanker instead of the target, and making feature errors when reporting the correctly located target object.

Thus, there is a probability P_L of selecting an object at the central (target) location, and probabilities P_{E1} , P_{E2} , and P_{E3} , for making an error in identifying each of the three target features, conditional on the correct spatial selection. Consequently, the probability of the correct

answer is $P_C = P_L \cdot (1 - P_{E1}) \cdot (1 - P_{E2}) \cdot (1 - P_{E3})$. With probability $1 - P_L$, one of the presented distractor objects is selected. For simplicity, we assume that a distractor object, if selected, will be reported with probability 1. (There is a lot of evidence that surrounding objects can be identified more accurately than the central one; e.g., Estes, 1982; Huckauf & Heller, 2002; Styles & Allport, 1986.)

In the simplest model (Model 1), feature errors were completely random with a constant probability, independent across feature dimensions, and independent of flanking objects in a display ($P_{E1} = P_{E2} = P_{E3} = P_E$).

In Model 2, the probabilities of feature errors (P_{E1} , P_{E2} , and P_{E3}) were proportional to the number of occurrences of respective non-target features in a display, $P_{Ek} = P_{EM} \cdot N_{NTFk}/N$, where P_{Ek} is probability of error on dimension k , P_{EM} is maximum probability of the feature errors, N_{NTFk} is number of objects with non-target features on dimension k , and N is the number of objects in display (P_{EM} was assumed to be the same for all three feature dimensions).

In Model 3, an additional assumption was made that the effect of flankers on probability of feature errors is modified by the similarity between the target and flankers in other feature dimensions (following the idea that feature migrations may be more frequent between more similar objects). For this model, $P_{Ek} = P_{EM} \cdot \sum(D_{TDik}/D_{TDi})/N$, where D_{TDik} is the difference between the target and distractor i on feature dimension k , and D_{TDi} is the difference (dissimilarity) between the target and distractor i , summed over all three feature dimensions (Σ is summation over all distractors i).

We chose to fit the distribution of responses across 7 categories based on two supposedly important variables: (1) difference of the response from the target (number of feature differences) and (2) correspondence vs. non-correspondence of the response to any of the flankers. These distributions (averaged across observers) for 2, 4, and 6 flankers are shown in Figure 7. However, we fitted the data of each observer separately. We used Microsoft Excel Solver to minimize the log-likelihood ratio statistic $G^2 = 2\sum N_j \ln(N_j/N_{pj})$, where N_j is the observed and N_{pj} is the predicted number of cases (trials) in response category j .

The fits are shown in Table 2. Each of these models has 2 free parameters and they are directly comparable. (We tried also more simple models with one free parameter, assuming either mislocation of integrated objects, or feature errors only. The fit was much worse.)

It seems that Models 2 and 3 are better than Model 1, implying that feature errors are not independent, but biased by the features of flankers. The more complex effect of modifying feature bias with object similarity (Model 3) improved the fit for one observer only. While either Model 2 or Model 3 fit the 4- and 6-flanker data well, none of these models is very good for the 2-flanker condition. We could fit all these data with a 3-parameter

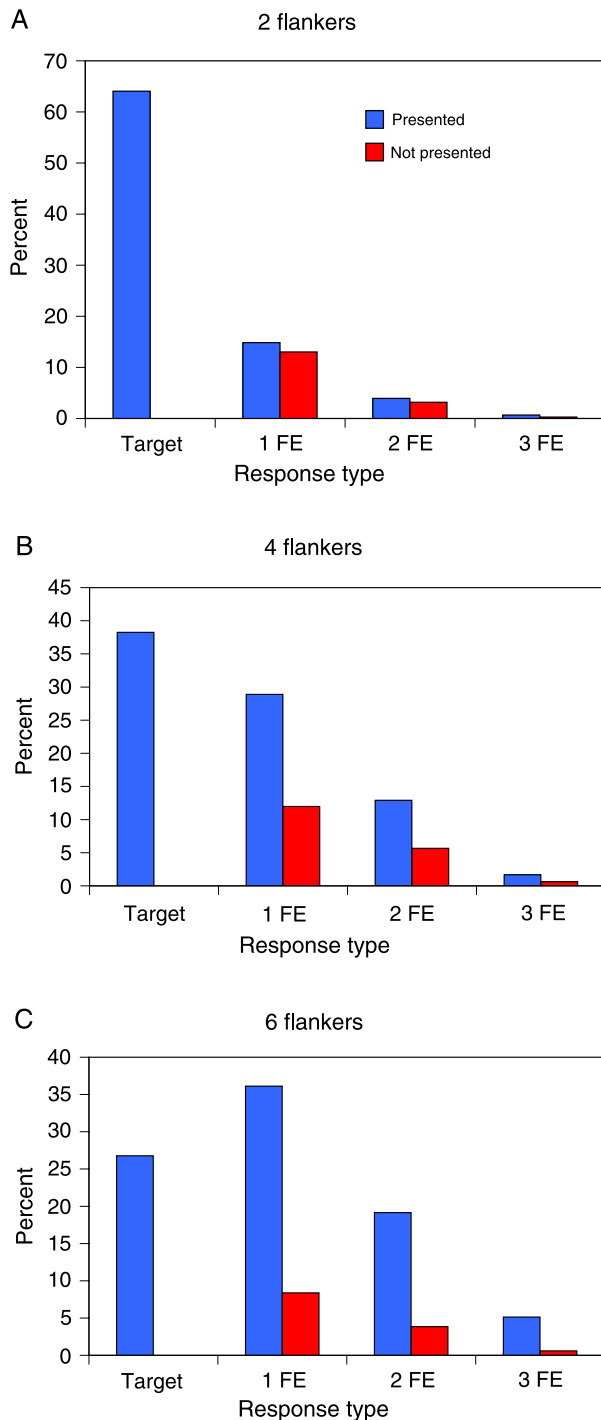


Figure 7. Distributions of responses used for modeling (averages across 3 observers). 1FE, 2FE, and 3FE signify 1, 2, and 3 feature errors, respectively.

model, allowing a proportion of pure guesses combined with Model 3 (referred as Model 3G in the tables).

The optimal values of the parameters for all these models are given in Table 3. The guessing parameter of Model 3G seems to vary a lot both between and within subjects, especially for 4 and 6 flankers conditions. However, this parameter is not critical for these conditions.

Acceptable fits for 4 and 6 flankers were obtained also with the optimal 2-flanker parameters and with mean parameter values across conditions.

Also, we used a combination of object localization errors and flanker-dependent feature errors (Model 2) for the simulation of observer’s responses and were able to reproduce the pattern of correlations found earlier (Table 1) between the probabilities of a particular object among incorrect responses and proportions of this object among flankers, and proportions of the respective features in a display.

Our modeling results seem to support the idea that crowding is a mix of (at least) two processes: flanker-dependent feature errors (mostly “illusory conjunctions”) and object localization errors. The object localization errors and feature errors as estimated by Model 3 are shown in Figure 8 (we use these transformed parameters instead of the original ones for a more intuitive comparability). There are several, potentially interesting, differences between the behaviors of these parameters. The object localization errors tend to increase more rapidly with larger number flankers, while feature errors seem to level off after 4 flankers. Also, there seems to be more variability across the observers for object localization errors, especially for 6 flankers.

However, we cannot be fully satisfied with these models. We noticed that nominally different processes can produce quite similar results, and consequently, there is a trade-off between parameters. For example, the flanker-dependent feature errors resemble the selection of wrong objects to some extent. Furthermore, there is a possibility of similarity-dependent object selection errors (Estes, 1982) that makes the separation of two types of errors even more problematic. Also, the probabilities of both object localization and feature errors increase with the number of flankers and also seem to correlate across observers. Is it possible that there is a common mechanism behind them?

Interestingly, there is a simple model based on the Feature Integration Theory (Treisman & Gelade, 1980)

Number of flankers	Observer	Model 1	Model 2	Model 3	Model 3G
2	EP	35.8**	21.9**	13.0*	7.0
	LP	14.9**	9.8*	15.3**	7.0
	SE	19.8**	17.7**	22.1**	1.0
4	EP	40.1**	20.8**	3.2	3.2
	LP	6.3	4.8	9.0	4.0
	SE	6.8	2.8	1.8	0.9
6	EP	14.9**	9.6*	4.6	4.5
	LP	6.0	3.2	5.3	5.3
	SE	8.0	5.1	5.9	0.4

Table 2. Fits of the models (values of G^2). Significant differences between observed and predicted distributions of responses. Note: ** $p < 0.01$, * $p < 0.05$.

Number of flankers	Observer	Model 1		Model 2		Model 3		Model 3G		
		P_L	P_E	P_L	P_{EM}	P_L	P_{EM}	P_L	P_{EM}	P_G
2	EP	0.83	0.10	0.89	0.22	0.91	0.39	0.90	0.35	0.03
	LP	0.94	0.12	0.98	0.23	0.98	0.38	0.98	0.31	0.07
	SE	0.93	0.12	0.97	0.23	0.97	0.38	0.97	0.28	0.09
4	EP	0.71	0.23	0.87	0.48	0.88	0.82	0.88	0.82	0
	LP	0.81	0.23	0.95	0.49	0.92	0.78	0.86	0.58	0.20
	SE	0.81	0.19	0.91	0.39	0.91	0.65	0.89	0.59	0.06
6	EP	0.48	0.27	0.59	0.54	0.60	0.91	0.59	0.87	0.02
	LP	0.64	0.24	0.75	0.49	0.75	0.82	0.75	0.82	0
	SE	0.77	0.25	0.90	0.50	0.87	0.82	0.83	0.66	0.16

Table 3. The fitted parameters of the models (P_L —probability of correct localization of the target, P_E —probability of feature errors, P_{EM} —maximum probability of feature errors, P_G —proportion of trials with guessing).

that could generate a mix of object and feature mislocalizations. Assume that within a spotlight of attention (or a corresponding receptive field) the “binding” of features is completely random (probabilistic). It means that if there are more red objects and more vertical objects within the receptive field then the perception of “red and vertical” becomes more probable, regardless of the actual conjunctions. With a small spotlight, centered on the target, only features of the target are sampled, and the correct conjoining is warranted. With a larger spotlight, features

from distractors are probabilistically combined with these from the target, and between themselves, and illusory conjunctions are perceived. Further, the exact position of the spotlight (or receptive field) relative to the target can be varied. Sometimes, it may be close to one of the distractors. Then, predominantly the features of this distractor will be selected, and with high probability, their conjunction corresponding to this distractor will be reported.

We wondered whether it is possible to find a combination of the size and the positional variability of attended receptive field that could accommodate our data. We assumed that receptive fields are circular, with Gaussian spatial profile, and that profile determines the probability of selection of features from any spatial position within the receptive field. We also assumed that the position of the attended receptive field relative to the center of the target varies according to the 2D normal distribution. We used a computer simulation to generate the probabilities of different types of responses as dependent on the size (σ_s) and positional variability (σ_p) of hypothetical receptive fields. We searched for an appropriate pair of parameters that could reproduce the empirical data. The best compromise we found corresponds to the parameters σ_s and σ_p both equal to about 0.4, measured in the units of the target–flanker distance (or 0.32 deg of visual angle). The predicted response distributions are given in Figure 9. Although there are some discrepancies, the overall similarity to the empirical data (Figure 7) is rather impressive (especially because there were only two adjustable parameters relative to the 18 degrees of freedom of the modeled data set). The unexpected finding that parameters of size and position variability must be equal looks interesting, but the meaning of this is not clear.

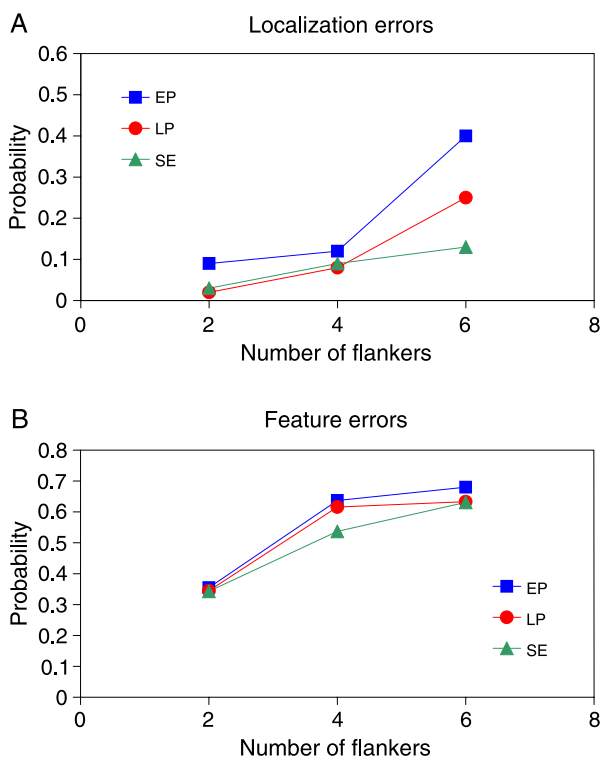


Figure 8. Probabilities of object localization errors and feature errors as estimated by Model 3. For feature errors, the probability of occurrence of at least one feature error, conditional on correct target localization, is given.

Discussion

Several recent studies (e.g., Levi et al., 2002; Pelli et al., 2004) have related crowding with a feature integration

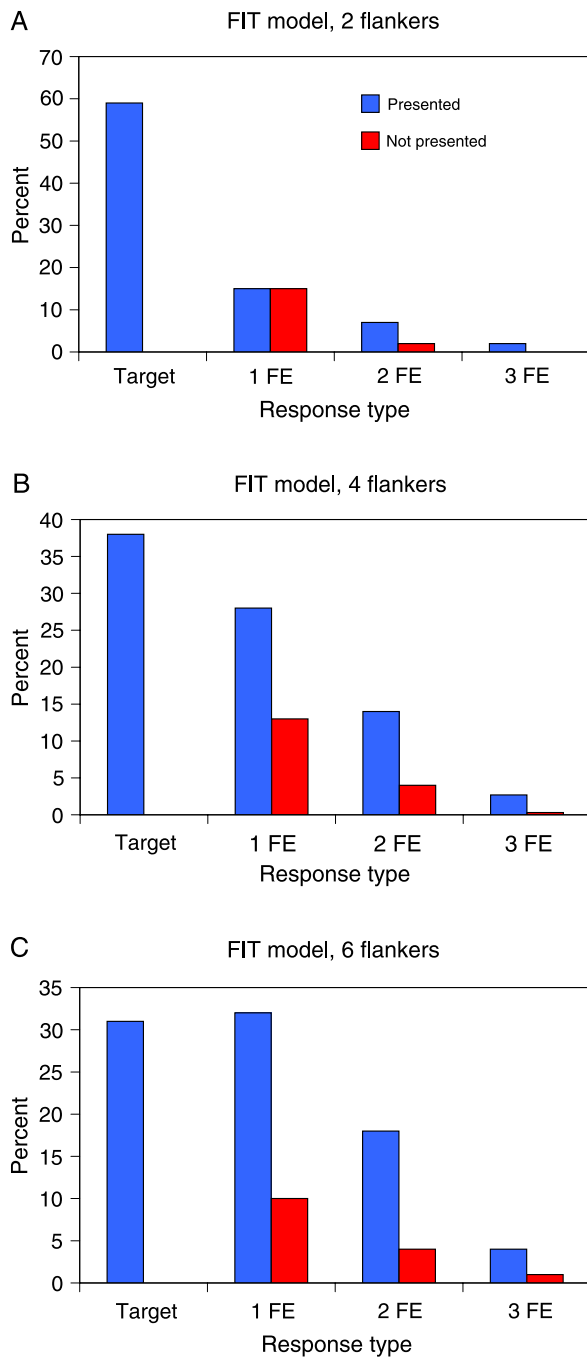


Figure 9. Distributions of responses predicted by the feature integration model. 1FE, 2FE, and 3FE signify 1, 2, and 3 feature errors, respectively (compare with the empirical distributions in Figure 7).

stage in visual processing. We used simple multidimensional stimuli in order to study the properties of this integration mechanism.

Our experiment showed that adjacent irrelevant objects affect the perception of the target in several ways. Sometimes, a flanking object is perceived and reported instead of the target; sometimes, the target seems to be located correctly, but observers make errors in reporting

one or two features. These feature errors are not completely random, but strongly biased by the features of distractors.

Our results are well in accord with Strasburger (2005), Huckauf and Heller (2002), and others who have suggested that the important mechanism of crowding is the selection of incorrect object. With their stimuli, these authors could not study the role of the features. The results are also consistent with a recent study by Nandy and Tjan (2007), who found the mislocalization of features being the main component of crowding. Their methods did not allow observing mislocalization of whole objects.

In each feature dimension separately (see Figure 6), the effect of flankers is largely consistent with a pooling account if we assume that the target has a larger weight relative to the distractors, and the proportions of the features rather than averages are calculated. Qualitatively, the same effect could be considered as a bias or criterion shift (in the sense of Signal Detection Theory), induced by flankers. However, different from Estes (1982), our data seem to imply that the both criterion and sensitivity are affected by the flankers. Thus, the pooling model looks more attractive.

We found that very simple feature integration model could qualitatively predict both the distribution of responses across the different categories and the effect of number of flankers found in our experiment. This model follows the original idea of the Feature Integration Theory (Treisman & Gelade, 1980) that features are conjoined simply by the fact of falling simultaneously within the spotlight of spatial attention. The theory does not specify exactly what happens when the spotlight is too large and includes several objects. We used the simple assumption that “binding” could then be completely random. However, because of the feature sampling according to the Gaussian profile, the features of the target (located near the center of the spotlight) are usually preferred. The additional assumption of imprecise location of the spotlight relative to the target makes it possible to predict responses that correspond to the integrated distractor objects. It is interesting that both feature and object localization errors can be generated by essentially the same probabilistic mechanism, and there is no clear border between them.

We found that the optimal receptive field size (sigma of Gaussian profile) for our model was 0.32 deg at 4 deg eccentricity. A reasonable estimate of full receptive field radius ($2 \times$ sigma) is therefore 0.64 deg (or 0.16E), yielding a diameter of 1.28 deg. This size is somewhat smaller than the size of integration fields suggested by some other models (e.g., Pelli et al., 2004) because our model explains a considerable fraction of the crowding effect by the position uncertainty of the attended receptive field.

Looking at the available neurobiological data (Gattass, Gross, & Sandell, 1981; Gattass, Sousa, & Gross, 1988; Smith, Singh, Williams, & Greenlee, 2001; Van Essen,

Newsome, & Maunsell, 1984), we could suggest that our receptive field size (diameter 1.28 deg at 4 deg eccentricity) resembles best the receptive field size of the visual brain area V2. However, the measurements of the receptive field sizes in a given brain area and eccentricity vary a lot, and also, it is not clear whether crowding is determined by the average or the smallest available receptive fields. Consequently, we cannot make any strong assertions about the possible neural site.

Despite of its attractiveness, this mechanistic model has clear limitations. It cannot explain the usual observation that we have no problems to perceive several objects (e.g., the target and one of the flankers) simultaneously (e.g., Huckauf & Heller, 2002; Popple & Levi, 2005). To explain this, maybe several simultaneous spotlights are needed. Also, this model will have problems with configuration effects in crowding displays as reported by Livne and Sagi (2007). At present, it is hard to imagine how it could work with “within-object conjunctions” (relative positions of features within object). Anyway, this simple model may be a useful building block for more complex ones.

In conclusion, the present study shows that simple multidimensional stimuli with explicit feature dimensions make it possible to address many interesting questions about the mechanisms of crowding and feature integration in vision.

Acknowledgments

We would like to thank the reviewers, who provided useful comments to improve our paper, and Hans Op de Beeck for a discussion of neurobiological receptive field data. EP was supported by visiting fellowships from the Scientific Research Fund-Flanders (GP.018.06N) and from the Research Fund of the University of Leuven (F/07/007), and by the Estonian Science Foundation (grant #6796). The research was partly funded by a large-scale grant from the Research Fund of the University of Leuven (GOA 2005/03) to JW.

Commercial relationships: none.

Corresponding author: Endel Pöder.

Email: endel.poder@psy.kuleuven.be.

Address: Laboratory of Experimental Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium.

References

- Andriessen, J. J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research*, *16*, 71–78. [PubMed]
- Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review*, *103*, 165–192. [PubMed]
- Baldassi, S., Megna, N., & Burr, D. C. (2006). Visual clutter causes high magnitude errors. *PLoS Biology*, *4*, 387–394. [PubMed] [Article]
- Bjork, E. L., & Murray, J. T. (1977). On the nature of input channels in visual processing. *Psychological Review*, *84*, 472–484. [PubMed]
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178. [PubMed]
- Butler, B. E., & Currie, A. (1986). On the nature of perceptual limits in vision: A new look at lateral masking. *Psychological Research*, *48*, 201–209. [PubMed]
- Donk, M. (1999). Illusory conjunctions are an illusion: The effect of target–nontarget similarity on conjunction and feature errors. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1207–1233.
- Eriksen, C. W., & Rohrbaugh, J. W. (1970). Some factors determining efficiency of selective attention. *American Journal of Psychology*, *83*, 330–342.
- Estes, W. K. (1982). Similarity-related channel interactions in visual processing. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 353–382. [PubMed]
- Felisberti, F. M., Solomon, J. A., & Morgan, M. J. (2005). The role of target salience in crowding. *Perception*, *34*, 823–833.
- Gattass, R., Gross, C. G., & Sandell, J. H. (1981). Visual topography of V2 in the macaque. *Journal of Comparative Neurology*, *201*, 519–539. [PubMed]
- Gattass, R., Sousa, A. P., & Gross, C. G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *Journal of Neuroscience*, *8*, 1831–1845. [PubMed] [Article]
- Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effects. *Visual Cognition*, *9*, 889–910.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, *43*, 171–216. [PubMed]
- Ivry, R. B., & Prinzmetal, W. (1991). Effect of feature similarity on illusory conjunctions. *Perception & Psychophysics*, *49*, 105–116. [PubMed]
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*, 255–279. [PubMed]
- Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in

- peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision*, 2(2):3, 167–177, <http://journalofvision.org/2/2/3/>, doi:10.1167/2.2.3. [PubMed] [Article]
- Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision*, 7(2):4, 1–12, <http://journalofvision.org/7/2/4/>, doi:10.1167/7.2.4. [PubMed] [Article]
- Nandy, A. S., & Tjan, B. S. (2007). The nature of letter crowding as revealed by first- and second-order classification images. *Journal of Vision*, 7(2):5, 1–26, <http://journalofvision.org/7/2/5/>, doi:10.1167/7.2.5. [PubMed] [Article]
- Neri, P., & Levi, D. M. (2006). Spatial resolution for feature binding is impaired in peripheral and amblyopic vision. *Journal of Neurophysiology*, 96, 142–153. [PubMed] [Article]
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4, 739–744. [PubMed] [Article]
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, 4(12):12, 1136–1169, <http://journalofvision.org/4/12/12/>, doi:10.1167/4.12.12. [PubMed] [Article]
- Pöder, E. (2006). Crowding, feature integration, and two kinds of “attention.” *Journal of Vision*, 6(2):7, 163–169, <http://journalofvision.org/6/2/7/>, doi:10.1167/6.2.7. [PubMed] [Article]
- Popple, A. V., & Levi, D. M. (2005). The perception of spatial order at a glance. *Vision Research*, 45, 1085–1090. [PubMed]
- Prinzmetal, W., Ivry, R. B., Beck, D., & Shimizu, N. (2002). A measurement theory of illusory conjunctions. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 251–269. [PubMed]
- Smith, A. T., Singh, K. D., Williams, A. L., & Greenlee, M. W. (2001). Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cerebral Cortex*, 11, 1182–1190. [PubMed] [Article]
- Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, 5(11):8, 1024–1037, <http://journalofvision.org/5/11/8/>, doi:10.1167/5.11.8. [PubMed] [Article]
- Strasburger, H., Harvey, L. O., Jr., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*, 49, 495–508. [PubMed]
- Styles, E. A., & Allport, D. A. (1986). Perceptual integration of identity, location and colour. *Psychological Research*, 48, 189–200. [PubMed]
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32, 1349–1357. [PubMed]
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107–141. [PubMed]
- Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97–136. [PubMed]
- Van Essen, D. C., Newsome, W. T., & Maunsell, J. H. (1984). The visual field representation in striate cortex of the macaque monkey: Asymmetries, anisotropies, and individual variability. *Vision Research*, 24, 429–448. [PubMed]
- Wolford, G. (1975). Perturbation model for letter identification. *Psychological Review*, 82, 184–199. [PubMed]
- Wolford, G., & Chambers, L. (1984). Contour interaction as a function of retinal eccentricity. *Perception & Psychophysics*, 36, 457–460. [PubMed]
- Wolford, G., & Shum, K. H. (1980). Evidence for feature perturbations. *Perception & Psychophysics*, 27, 409–420. [PubMed]