Journal of Vision (2015) 15(2):21, 1–14 http://www.journalofvision.org/content/15/2/21

Eyes on crowding: Crowding is preserved when responding by eye and similarly affects identity and position accuracy

Funda Yildirim
Laboratory of Experimental Ophthalmology,
University Medical Center Groningen,
University of Groningen, Groningen, The Netherlands

Vincent Meyer
Laboratory of Experimental Ophthalmology,
University Medical Center Groningen,
University of Groningen, Groningen, The Netherlands

Frans W. Cornelissen
Laboratory of Experimental Ophthalmology,
University Medical Center Groningen,
University of Groningen, Groningen, The Netherlands

Peripheral vision guides recognition and selection of targets for eye movements. Crowding—a decline in recognition performance that occurs when a potential target is surrounded by other, similar, objects—influences peripheral object recognition. A recent model study suggests that crowding may be due to increased uncertainty about both the identity and the location of peripheral target objects, but very few studies have assessed these properties in tandem. Eye tracking can integrally provide information on both the perceived identity and the position of a target and therefore could become an important approach in crowding studies. However, recent reports suggest that around the moment of saccade preparation crowding may be significantly modified. If these effects were to generalize to regular crowding tasks, it would complicate the interpretation of results obtained with eye tracking and the comparison to results obtained using manual responses. For this reason, we first assessed whether the manner by which participants responded—manually or by eye—affected their performance. We found that neither recognition performance nor response time was affected by the response type. Hence, we conclude that crowding magnitude was preserved when observers responded by eye. In our main experiment, observers made eye movements to the location of a tilted Gabor target while we varied flanker tilt to manipulate target–flanker similarity. The results indicate that this similarly affected the accuracy of peripheral recognition and saccadic target localization. Our results inform about the importance of both location and identity uncertainty in crowding.

Introduction

In crowding, recognition of an object is impaired when it is surrounded by other—similar—objects. The current main theories explain crowding either in terms of excessive feature pooling (e.g., Greenwood, Bex, & Dakin, 2010; Pelli & Tillman, 2008; van den Berg, Roerdink, & Cornelissen, 2010) or as due to a loss of positional information (source confusion) resulting in reporting a flanking object as the target (e.g., Dakin, Cass, Greenwood, & Bex, 2010; Greenwood, Bex, & Dakin, 2009; Strasburger, Harvey, & Rentschler, 1991; Strasburger & Malania, 2013). A recent model integrates both of these accounts by assuming that uncertainty (i.e., the width of the internal noise distribution) about both stimulus positions and identities depends on flanker proximity (van den Berg, Johnson, Martinez Anton, Schepers, & Cornelissen, 2012). To further test this idea, crowding studies would ideally assess target position and identity in tandem, but to our knowledge, hardly any studies have done so thus far (a study by Greenwood, Bex, & Dakin, 2012, forms a notable exception, and we will discuss it later).

A very natural way to simultaneously assess what an observer saw and where s/he saw it is through using eye tracking. Perceived identity can be inferred from the selected object, and the saccadic localization data can be used to evaluate perceived location. This is the approach we will take in our main experiment that will be aimed at simultaneously assessing perceived position

Citation: Yildirim, F., Meyer, V., & Cornelissen, F. W. (2015). Eyes on crowding: Crowding is preserved when responding by eye and similarly affects identity and position accuracy. *Journal of Vision*, *15*(2):21, 1–14, http://www.journalofvision.org/content/15/2/21, doi:10.1167/15.2.21.

doi: 10.1167/15.2.21 Received April 29, 2014; published February 16, 2015 ISSN 1534-7362

and identity in crowding. However, before we can proceed with this approach, we first need to verify whether eye movements themselves influence crowding in our type of task. This is assessed in Experiment 1.

Experiment 1: Comparing crowding when responding manually or by eye

The possibility that eye movements influence crowding is suggested by several recent studies that assessed the relationship between crowding and eye movements (Harrison, Mattingley, & Remington, 2013a; Harrison, Retell, Remington, & Mattingley, 2013). These studies suggest that just prior to the observer making an eye movement, crowding may substantially change. If these effects were to generalize to other crowding tasks (e.g., tasks that require identifying and locating targets with one's eyes), this would complicate the interpretation of results obtained with eye tracking. In fact, it would call into question the relevance of the majority of previous crowding studies for understanding natural vision (as these mostly used manual responses and required participants to fixate).

Harrison, Mattingley et al. (2013a) reported that about 50 ms before a saccade is initiated toward a crowded object, the magnitude of crowding was reduced. Moreover, the spatial area within which crowding occurred was approximately halved. Harrison, Mattingley et al. (2013a, p. 2927) concluded that "eye movement preparation effectively enhances object discrimination in peripheral vision at the goal of the intended saccade. These presaccadic changes may enable enhanced recognition of visual objects in the periphery during active search of visually cluttered environments." Various mechanisms may underlie the reduced crowding observed by Harrison, Mattingley et al. (2013a). As an explanation, they proposed that extraretinal signals during saccade preparation stop the obligatory averaging of flanker and distractor features, thereby partially releasing the target from crowding. Two alternative explanations for the phenomenon were suggested by van Koningsbruggen and Buonocore (2013). Preparing an eye movement might lead to a shift of covert attention to the saccade target (Deubel & Schneider, 1996), thereby enhancing discrimination performance at the location of the saccade goal (Deubel, 2008). To the extent that this influence is sufficiently selective, perception of the target might be enhanced more than that of the distractors, effectively reducing crowding magnitude. However, this explanation has been disputed by Harrison, Mattingley, and Remington (2013b; p. 1) because "observers knew the target's position and approximate timing in the no-eye movement and eye movement conditions so that attention could be allocated in the same manner on every trial." Second, performance might have been enhanced through saccadic unmasking (De Pisapia, Kaunitz, & Melcher, 2010; Hunt & Cavanagh, 2011). In this presaccadic effect, the target and its distractors—although presented at the same physical location—are being perceived at different spatial locations. As perceived—and not physical—position determines crowding (Dakin, Carlson, & Greenwood, 2011) this would effectively release a target from the crowding influence of its distractors. In another study, Harrison, Retell et al. (2013) reported a different perisaccadic phenomenon, which they referred to as "remapped crowding." Harrison, Retell, et al. found that flankers flashed at the postsaccadic location of a target but prior to the actual saccade (and thus rather distant from the target), nevertheless affecting the magnitude of crowding. Authors explained these results on the basis of a predictive remapping of receptive fields prior to saccades. Effectively, this phenomenon could increase the amount of crowding around saccadic eye movements.

There is one important reason to question whether presaccadic phenomena would be able to improve saccadic goal selection is the existence of a *saccadic dead time*. This is the brief period just prior to saccadic execution during which neither the execution of the saccade can be cancelled nor its goal changed (Hooge & Erkelens, 1996; Hooge, Beintema, & van den Berg, 1999). Hooge and Erkelens (1996) estimated that this saccadic dead time starts approximately 70 ms prior to a saccade. Thus, there appears to be a substantial overlap between this estimate of saccadic dead time and the period in which the reduced crowding has been observed. De Vries, Hooge, Wiering, and Verstraten (2011) reported the presence of crowding in an eye movement–based visual search paradigm, thus suggesting that crowding is still present in the presence of eye movements. However, they did not compare crowding magnitude with and without eye movements, so reduced crowding might still have been present. Note that this question has not been answered by the studies of Harrison, Mattingley et al. (2013a) or Harrison, Retell et al. (2013) as, in order to measure crowding magnitude, observers made manual responses after having moved their eyes. Hence, the question of whether presaccadic phenomena can influence perception during natural tasks remains open.

Therefore, in this experiment preceding the main one of the present study, we set out to assess the influence of responding by eye on crowding magnitude. To do so, we adapted the most common crowding paradigm—recognition of a single isolated or flanked object in peripheral vision—for use with eye movements. In our paradigm, observers are simultaneously shown an (isolated or flanked) object to the left and to the right

Figure 1. An example stimulus. In different blocks of trials, observers were instructed to indicate the most rightward-tilted target (central object) by either making a saccade or by pressing a button while maintaining fixation.

sides of a fixation mark and respond by making an eye movement to the one perceived to be the target. Note that, in essence, this paradigm mimics what the visual system has to do in most natural viewing behavior, e.g., visual search: select a potential target—out of the usually many available in peripheral vision—for further scrutiny and plan a saccade to its location. During manual responding, observers fixate and indicate the target location by pressing one of two buttons. Our hypothesis is that recognition performance in this crowding task does not change between eye and manual responses.

Methods

Overview

In the experiment, we measured the errors in discriminating between a target and a reference presented left and right of fixation. Target and reference could either be presented in isolation or be surrounded by four flankers. The target could be discerned from the reference on the basis of tilt: The observer's task was to choose the most right-tilted object. In different blocks of trials, observers indicated their responses by (a) making a saccade to the perceived target or (b) pressing a left or right arrow key. Below the experiment is described in detail.

Observers

Eight observers (age range: 20–49; three women) participated in the experiment. Authors FY and FWC were among the observers. The remaining observers were naïve as to the purpose of the experiment. All observers had normal or corrected-to-normal vision.

Materials

Observers viewed stimuli on a 22-in. CRT RGB monitor with a frame rate of 75 Hz (LaCie) from a distance of 59 cm. Stimulus presentation, eye movement recording, and response collection were programmed in Matlab (MathWorks) using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and EyeLink Toolbox extensions (Cornelissen, Peters, & Palmer, 2002). Eye movements were recorded at 250 Hz with an EyeLink 1000 (SR Research, Kanata, Ontario, Canada) infrared eye tracker. We used the EyeLink's built-in nine-point calibration procedure. Background luminance during the experiment was 35 cd/m^2. We used a chin rest and a forehead rest to stabilize the observer's head position.

Stimuli and procedure

Prior to the start of a trial, a white horizontal line (a minus sign; 0.2°) was presented at the center of the display. The observer pressed a key to commence the trial. Following the key press, the horizontal line changed into a fixation cross. Two hundred milliseconds thereafter, the stimulus was presented for 200 ms. An example stimulus is shown in Figure 1. The stimulus consisted of a target and reference that were presented left and right of fixation, at either 8° or 10° of eccentricity and either in isolation or surrounded by flankers. The side at which the target appeared was determined randomly. Target, reference, and flankers were Gabor patches (width = 1.0°, spatial frequency 3.0 c/°). Base target and reference tilt was set to 45°. To distinguish the target from the reference, it was tilted clockwise from base tilt by 5°, and the reference was tilted counterclockwise by 5°. In 20% of the trials, target and reference were presented in isolation. In the remaining 80% of the trials, four flankers surround the target and reference. Flankers were positioned at the four corners of an invisible square with the target or reference at the center. Flanker tilt was either the same as the base tilt (45°) to create a high (difference 0°) or differed (by 45° or –45°) to create a low target–flanker similarity condition (see Figure 2 for an example). The 0° flanker tilt condition occurred twice as often as each of the 45° or –45° flanker tilt conditions. Flankers were always presented at 25% contrast. Target and reference were presented at individually determined contrast and tilt thresholds (see below).

The observer's task was to choose the most right-tilted target (the right one in the example of Figure 1). In

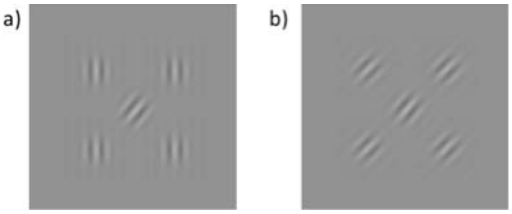

Figure 2. Example stimuli. (a) Low target–flanker similarity and (b) high target–flanker similarity.

alternating blocks of trials, observers either indicated their response by making a saccade to the target or by pressing one of two keys. After the response, the fixation point turned either red (error) or green (correct) to provide feedback to the observer. During manual responding, observers were required to maintain steady fixation throughout the trials, and their gaze was monitored. A single block of trails consisted of 260 trials. In their first session, observers first completed a 100-trial training block for each condition. Following these, observers completed two blocks of experimental trials in a row. They commenced with a block in which they responded manually, followed by a block in which they responded by eye. In each of their next three sessions, observers completed two more blocks of trials in which the order of eye and manual responding was alternated each time. All eight observers thus completed eight blocks of trials for a total of 1,040 trials in each response mode.

Individual contrast and tilt threshold determination

In a session preceding the experimental ones, we individually determined the tilt and contrast levels that enabled observers to achieve 80% correct recognition performance (manual responses) at each eccentricity for isolated targets and references. Observers were asked to report the more rightward tilted target. Target and reference tilt were set to 45° ± 5°, respectively, and were presented at either 8° or 10° of eccentricity. In different trials, target and reference were presented at 13 different levels of contrast (1.0%, 1.3%, 1.7%, 2.2%, 2.9%, 3.8%, 5.0%, 6.5%, 8.5%, 11.2%, 14.6%, 19.1%, or 25.0% contrast). For the first two experimental sessions, we determined an individual contrast level for each observer that enabled them to achieve 80% correct performance. We did so by first fitting a cumulative normal distribution to a participant's performance data as a function of contrast level. Next, using this distribution, we determined the threshold level of contrast. For the final two sessions, thresholds were determined slightly differently. The contrast was fixed, and the required tilt difference was determined. Contrast was set to the level to achieve 70% correct performance. Next, we determined the tilt level that enabled 80% correct performance. Target and references were presented at 13 different tilt values (difference from base tilt: 1.0°, 1.3°, 1.7°, 2.2°, 2.9°, 3.8°, 5.0°, 6.3°, 7.9°, 10.0°, 12.6°, 15.9°, or 20.0°). Again, a cumulative normal distribution was fitted to the performance data as a function of tilt, and this function was used to calculate the threshold tilt level. In the Results section, the results for all sessions will be integrally presented (after verifying that the type of threshold setting had no significant influence on the results).

Eye movement analysis

Saccades were determined using the EyeLink's built-in analyses routines. Prior to entering the statistical analysis, eye movement responses were filtered based on saccadic amplitude, saccade latency, and saccadic direction. Trials were removed in which saccades were (a) made within 150 ms or after 3000 ms following the start of the stimulus presentation or (b) in which saccadic direction differed more than 15° from horizontal (the direction of the target or reference) or (c) in which saccadic amplitude was less than two thirds of the target or reference eccentricity. On average, this excluded about 8% of the eye-response trials. For the manual trials, responses that were accompanied by an eye movement were excluded from the analysis. On average, this excluded about 3% of the manual-response trials.

Statistical analysis

We used repeated-measures ANOVA to determine whether differences in performance were statistically significant. Within-observer factors were response type (eye, manual), flanker mode (isolated target, low target–flanker similarity, high target–flanker similarity), and threshold (contrast, tilt). To test for differences in response time between manual and eye responses, first, a Gaussian was fit to the individual distributions, and the mean and standard deviations were determined. Paired t tests were used to test for differences in these fitted parameters.

Results

As we used two slightly different ways to set the individual thresholds, we first determined whether the factor threshold resulted in significant performance differences. Repeated-measures ANOVA indicated that this was not the case, $F(1, 7) = 1.1, p > 0.05$. Interactions of this factor with the other factors were not significant

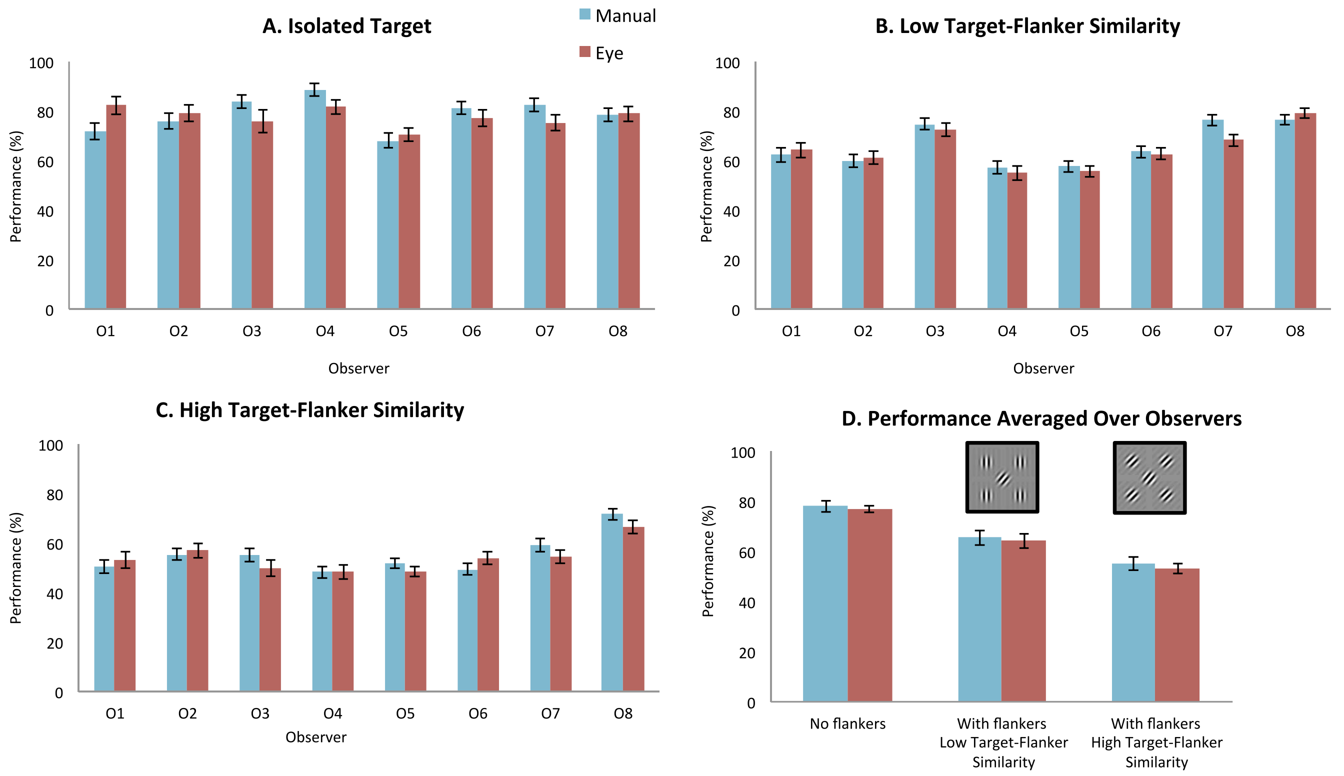

Figure 3. Discrimination performance of individual observers for manual and eye responses. (a) Results for the isolated target condition, (b) for the low target–flanker similarity condition, and (c) for the high target–flanker similarity condition. (d) Results averaged over observers. Bars in (d) indicate the standard error of the mean over observers.

either: response type, $F(1, 7) = 2.7, p > 0.05$; flanker mode, $F(2, 14) = .14, p > 0.05$. For this reason, in the remainder of the analysis, this factor is not further considered.

Figure 3a through c shows the results for the individual observers with different panels for the isolated target and the low and high target–flanker similarity conditions. Figure 3d shows the results averaged over observers. In each panel, results are split by response mode (eye or manual). Not surprisingly, observers performed best when the target and reference were presented in isolation (Figure 3a and d). As expected, adding flankers decreased performance (Figure 3b through d). This decrease in performance was largest when target–flanker similarity was highest (Figure 3c, d). Overall, the small difference between the two different response modes was not significant, $F(1, 7) = 0.7, p > 0.05$, and it showed no significant interaction with target-flanker similarity either, $F(2, 14) = 0.5, p > 0.05$.

To examine the crowding effect in more detail, we calculated crowding magnitude by subtracting performance in the two flanker conditions from performance in the isolated target condition. This is a standard procedure to calculate the difference in performances for isolated and flanked conditions. The results are shown in Figure 4. It is obvious that the results for the manual and the eye responses are nearly identical.

Finally, we compared the time that it took observers to respond when making manual or eye responses. Figure 5 shows the response time histograms for both manual and eye responses. Response times and saccadic latencies were binned in 50-ms bins. As Figure 5 shows, the histograms of response times for the manual and eye responses are very similar. Neither the means nor the standard deviations of individually fitted

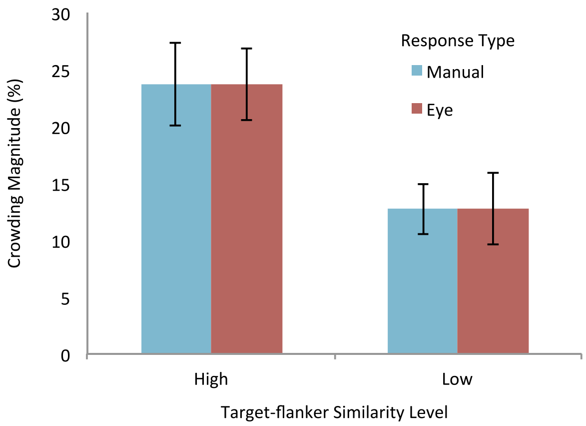

Figure 4. Crowding magnitude. Results are shown for the conditions with low and high target–flanker similarity. Average results of eight observers. Bars indicate standard error of the mean over observers.

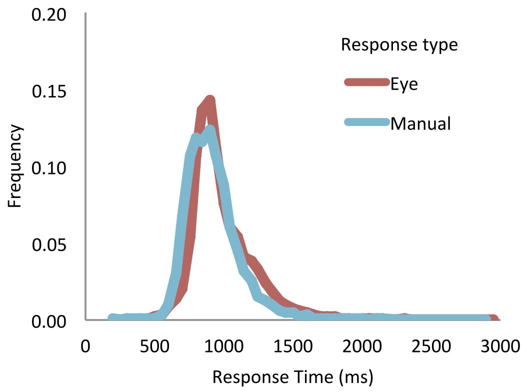

Figure 5. Frequency distribution of saccadic and manual response times.

Gaussians differed between eye and manual responses; means: manual: 845 ms and eye: 892 ms, $t(7) = -1.153$, $p > 0.05$; standard deviations: manual: 68 ms and eye: 63 ms, $t(7) = 2.203$, $p > 0.05$.

Discussion

In normal visual behavior, information from peripheral vision is used in saccade goal selection. In the task at hand, the visual system uses saliency and other metrics of relevance to determine the priority with which peripheral locations need to be scrutinized by foveal vision (Fecteau & Munoz, 2006; Hayhoe & Ballard, 2005; Itti & Koch, 2001; Tatler, Hayhoe, Land, & Ballard, 2011; Yanulevskaya, Marsman, Cornelissen, & Geusebroek, 2011). Peripheral vision is also the part of the visual field that is most strongly affected by crowding (Bouma, 1970; Korte, 1923; Stuart & Burian, 1962; Toet & Levi, 1992).

If eye movement preparation would effectively change peripheral discrimination—and thereby saccade goal selection—this would profoundly affect our understanding of the relationship between peripheral vision and crowding in natural viewing. In our view, it thus seemed worthwhile to compare crowding magnitude when observers responded either manually or used their eyes to respond. The main finding of our first experiment is that crowding magnitude is preserved when observers use their eyes to respond in a crowding task, thus confirming our hypothesis. This implies that previously observed changes in crowding as a result of eye movement preparation (Harrison, Mattingley et al., 2013a; Harrison, Retell et al., 2013) do not appear to change crowded object discrimination as measured in our—relatively standard—crowding paradigm. Therefore, we also conclude that eye tracking is a valid tool to assess crowding.

An important reason for not finding a difference was already suggested in the Introduction. The temporal window of the observed enhancement and integration effects is such that they show a substantial overlap with *saccadic dead time*. During this brief period just before saccade initiation, no changes to the saccade execution are possible anymore. The period of the saccadic dead time (approximately 70–100 ms) is very similar in size to the period during which changes in crowding were observed (approximately 50 ms prior to saccade initiation). It is important to note that our results do not necessarily demonstrate that eye movement preparation does not cause changes in crowding. However, our results do indicate that eye movement preparation does not affect discrimination performance in our paradigm.

We modeled our paradigm to mimic what we believe the visual system has to do in most natural viewing behavior, e.g., visual search: select a potential target—out of the usually many available in peripheral vision—for further scrutiny and plan a saccade to its location. One might object that an (forced) orientation-discrimination task preceding the eye movement is not natural. Yet planning an eye movement requires, at the least, a preattentive target-selection process (e.g., based on saliency). Therefore, in our view, our paradigm with a forced selection based on orientation followed by an eye movement comes close to the natural situation.

Our present design—with both a target and a reference symmetrically presented left and right of fixation—guaranteed that there was no net advantage on discrimination performance of moving one's eyes early. As in natural viewing behavior, our paradigm requires observers to decide on the location of the target before programming the saccade. Therefore, one could argue that discrimination performance could not have been affected by saccade preparation. However, Allik, Toom, and Luuk (2003) showed that saccade direction and amplitude can be planned separately. In our paradigm, observers could plan saccade direction only after having made the decision about which side contained the target. However, they might have started planning the saccadic amplitude as soon as the stimulus was shown, giving this process a substantial head start and also leaving sufficient time for any putative enhancement and integration effects to occur.

Experiment 2: Measuring perceived target identity and location using eye movements

Having shown that crowding is preserved for eye responses, we now can use eye tracking to simultaneously quantify errors in identity and perceived

Figure 6. An example stimulus for the second experiment. In different blocks of trials, observers were instructed to indicate either the most rightward-tilted target (central object) or the most leftward-tilted target by making a saccade.

location caused by crowding. It is not a priori certain that both aspects would be affected—let alone similarly affected—by crowding. "What" and "where" information is considered to be processed in different streams in the visual cortex (Goodale & Milner, 1992). A number of studies locate crowding effects beyond the early visual areas (Dakin et al., 2011; Liu, Jiang, Sun, & He, 2009), increasing the likelihood that position and identity information are differentially affected by crowding. Psychophysical support for separate coding of what and where aspects in crowded conditions has also been reported (Strasburger, 2005).

Another reason that might result in differences in identification and localization performance is the following: A number of studies have indicated the importance of perceived target position in crowding (Dakin et al., 2011; Maus, Fischer, & Whitney, 2011). However, as recently demonstrated using a motion illusion, (saccade) localization need not necessary follow perceived position (Lisi & Cavanagh, 2014). Greenwood et al. (2009, p. 13130) suggested that "crowding is a preattentive process that uses averaging to regularize the noisy representation of position in the periphery." Based on this—if anything—one would expect that crowding would tend to minimize localization errors, and consequently, identity and position errors would perhaps even be anticorrelated.

On the other hand, a recent crowding model proposed that flanker proximity affects uncertainty about both stimulus positions and identities (van den Berg et al., 2012). This model therefore predicts that both aspects would be affected by crowding and, most likely, to a similar extent. Moreover, a study by Greenwood et al. (2012, p. 1) investigated the binding of feature and relative position information in crowding. In one of their experiments, observers were asked to indicate both the target identity and perceived location. The authors concluded that "…crowding is a singular process that affects bound position and orientation values in an all-or-none fashion." Based on this finding, recognition and localization errors would also be expected to be largely correlated.

To assess how identity and position uncertainty are affected by crowding, we used eye tracking to measure the influence of target–flanker similarity on both recognition and saccadic localization performance. Based primarily on our model (van den Berg et al., 2012), our hypothesis is that increasing target–flanker similarity will lead to both increased recognition and localization errors.

Methods

Overview

We quantified the errors in recognition performance and saccadic landing position of a crowded target. The target was defined by tilt; the observer's task was to choose the most right- or left-tilted target, respectively, by making a saccade to the target.

Observers

Six different observers (age range: 23–27; all males) participated in the experiment. All of the observers were naïve as to the purpose of the experiment. All observers had normal or corrected-to-normal vision. One observer was excluded because of poor attention, resulting in poor performance, leaving five observers for the results reported.

Materials

The materials used in this experiment were identical to those used in the first experiment.

Stimuli and procedure

Prior to the start of a trial, a white fixation cross (0.2°) was presented at the center of the display. The observer initiated a trial by fixating on this fixation cross. If the fixation was stable for at least 250 ms, the trial commenced by presentation of the stimulus. It was presented until the observer made an eye response. If there was no response given within 3000 ms, the trial was marked as invalid. An example stimulus is shown in Figure 6. The stimulus consisted of a target and reference that were presented left and right of fixation, either at 8° or 10° of eccentricity and either in isolation or surrounded by flankers. The side at which the target appeared was determined randomly. Target, reference, and flankers were Gabor patches (width = 1.0°, spatial frequency 3.0 c/°).

In 10% of the trials, target and reference were presented in isolation. In the remaining 90% of the trials, flankers were present, and these were positioned at the four corners of an invisible square with the target or reference at the center. Center-to-center distance of the target and flankers was 0.15 times the target eccentricity. All four flankers always had the same tilt, which was randomly chosen to be one of 12 different values in the interval −75° to 90° (step size 15°). Flankers were always presented at 25% contrast. Target and reference were presented at 5% contrast.

In alternating blocks of trials, the base target and reference tilt was set to either 45° or −45° from vertical. To distinguish the target from the reference, targets varied either 5° or 10° from this base tilt, always mirrored for the reference Gabor. So the target and the reference differed in orientation by either 10° or 20°. The observer's task varied depending on the base tilt. In blocks in which base tilt was 45° (−45°), the observer's task was to choose the target by making a saccade to the most rightward (leftward) tilted Gabor.[1] After the response, the fixation point turned either red (error) or green (correct) to provide feedback to the observer. During manual responding, observers were required to maintain steady fixation throughout the trials, and their gaze was monitored. A single block of trails consisted of 300 trials.

In each of six sessions, observers completed two blocks of 300 trials each. For each session, the base tilt for the first of the two blocks was determined randomly. The other base tilt was used in the second block. Observers were offered a short break in between the two blocks of trials to minimize fatigue. All six observers thus completed 12 blocks of trials for a total of 3,600 trials per observer. Prior to the actual experiment, observers completed two training blocks of trials, one for each base tilt and task combination. These trials were used to verify that the observer's performance for isolated targets was in the range 70%–90%. The training trials were not used in the further analysis.

Eye movement analysis

Saccades were determined using the EyeLink's built-in analyses routines. Prior to the statistical analysis, eye movement responses were filtered based on saccadic amplitude, saccade latency, and saccadic direction. Trials were removed in which saccades were (a) made within 150 ms following the start of the stimulus presentation or (b) in which saccadic direction differed more than ±15° from horizontal (the direction of the target or reference) or (c) in which saccadic amplitude was less than two thirds of the target or reference eccentricity. On average, this excluded 13% of the eye response trials.

Analysis of recognition performance (crowding magnitude)

The first part of the analysis was done separately for each observer. First, we determined the average recognition performance in the "no-flanker" condition separately for each eccentricity and base tilt. To obtain crowding magnitude, this value was subtracted from performance in each of the flanker conditions, again separately for each eccentricity and base tilt. For the final analysis, the results were averaged over observers.

Analysis of localization errors

To determine the occurrence of changes in localization performance, we included the saccades made during both erroneous and correct recognition responses. Our reasoning behind also using the recognition errors is that, in those cases, the observer apparently considered the reference to be the target. Therefore, for determining localization performance, an error is as informative as a correct response.

First, saccadic landing positions to the left of fixation were mirrored in the origin. Next, we determined the average saccadic landing position for isolated target/reference trials (i.e., no flanker presented). This was done separately for each eccentricity and base tilt. Next, for each trial, we calculated the absolute distance between the saccadic landing position in that particular trial and the average saccadic landing position in the no flanker condition for the corresponding eccentricity and base tilt. Note that this removes any bias (e.g., due to saccadic undershoot) a participant may have had. We refer to this value as the localization error.

Next, average localization errors were calculated for each flanker orientation. Subsequently, the average localization error in the no-flanker condition was subtracted from the average localization error in each of the flanker conditions separately for each eccentricity and base tilt. Note that this measure is therefore analogous to the calculation of crowding magnitude in that it indicates the extent to which the localization error changed in magnitude as a result of the presence of flankers.

Finally, to compare the changes in recognition performance and localization error induced by the presence of flankers, each measure was also converted into a "z-score" by subtracting the average value and dividing by the standard deviation (over flanker

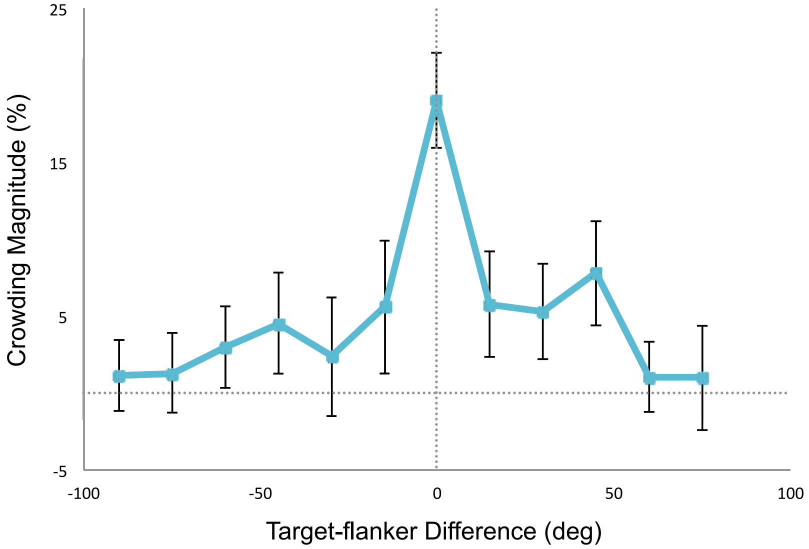

Figure 7. Crowding magnitude plotted as a function of the difference in tilt between target and flankers. Bars indicate ±1 *SEM* over observers.

orientation). These z-scores were calculated first for each individual observer separately for each base tilt and eccentricity.

Results

Repeated-measures ANOVA results revealed that eccentricity and base tilt of the target had no significant effect on either recognition performance, $F(1, 64) = 0.67$, $p > 0.05$, or on localization error, $F(1, 64) = 0.187$, $p > 0.05$. Hence, for the remainder of the results, we report data averaged over eccentricity and target base tilt.

We examined performance as a function of target–flanker similarity. Results are shown in Figure 7. As expected, target–flanker similarity had a substantial effect on crowding magnitude, and a clear peak in crowding magnitude can be observed at 0° difference, i.e., when target and flanker are most similar. Because crowding was predicted to be strongest at 0° tilt difference, we used a paired *t* test to compare crowding magnitude at tilt level 0° to the average crowding magnitude in the other conditions ($p < 0.001$). This indicates that crowding magnitude at 0° was increased compared to crowding magnitude in the other tilt conditions.

In an analogous fashion, for the analysis of localization error, we examined performance as a function of target–flanker similarity. Results are shown in Figure 8. Similar to crowding magnitude for recognition performance, target–flanker similarity had a substantial effect on localization error, and a clear peak can be observed near 0° difference. A paired *t* test comparing localization error at tilt level 0° to the average error in the other conditions ($p = 0.06$) indicated that localization error at 0° was increased compared to the average error in the other tilt conditions.

To enable comparison of the changes in recognition performance and localization error caused by crowding, results for each observer were converted into a z-score. The results are shown in Figure 9. Coinciding peaks in recognition (crowding magnitude) and localization performance (localization error) can be observed at 0° difference, i.e., when target and flanker are most similar. A repeated-measures ANOVA on the z-scores of recognition and localization performance revealed no significant interaction between similarity level and parameter, $F(11, 44) = 1.15$, $p = 0.345$, indicating that, when expressed in terms of a z-score, the crowding magnitudes for recognition and localization are not different.

Discussion

The finding of our main experiment is that flanking objects affect both recognition and localization performance and that crowding magnitude in both domains is approximately equal. As anticipated based on previous studies, we observed that crowding magnitude (recognition) was largest for identical target and flankers and decreased with increasing target–flanker difference (Figure 7). The localization error showed a very similar dependence on target–flanker difference (Figure 8). Moreover, when expressed as a z-score (Figure 9), crowding magnitudes for recognition

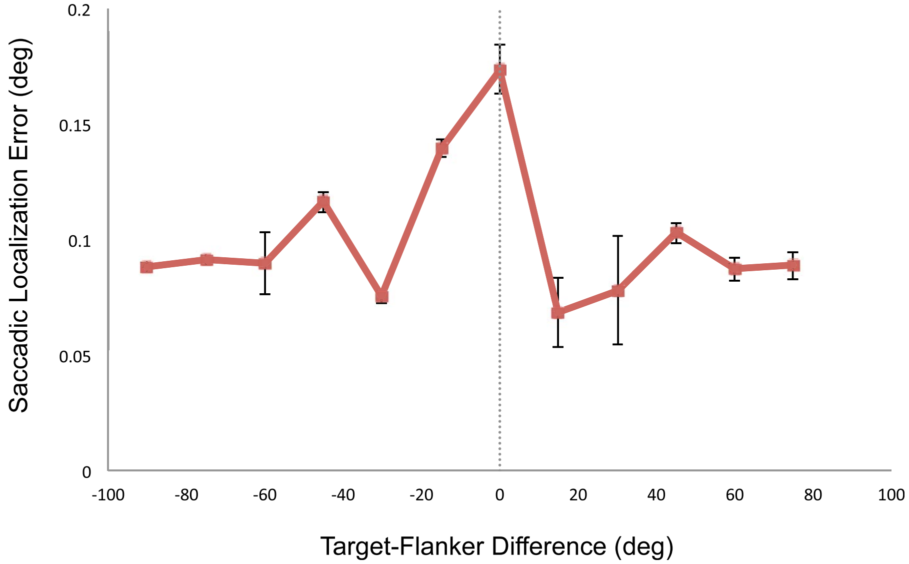

Figure 8. Saccadic localization error plotted as a function of the difference in tilt between target and flankers. Bars (where visible) indicate ±1 *SEM* over observers.

performance and for localization error turned out to be approximately equal. In turn, this suggests that identity and position uncertainty are similarly affected by target–flanker differences.

These results are consistent with the predictions based on the model study of van den Berg et al. (2012). They are also consistent with a report by Greenwood et al. (2012), who concluded that bound position and orientation are affected by crowding in an all-or-none fashion. Our results—although measured in an entirely different manner—are congruent with this finding.

In our study, saccadic localization error showed a similar pattern of results as present in identification performance, therefore suggesting that the perceived position of the target was affected by target-flanker similarity. Our results therefore imply that in this

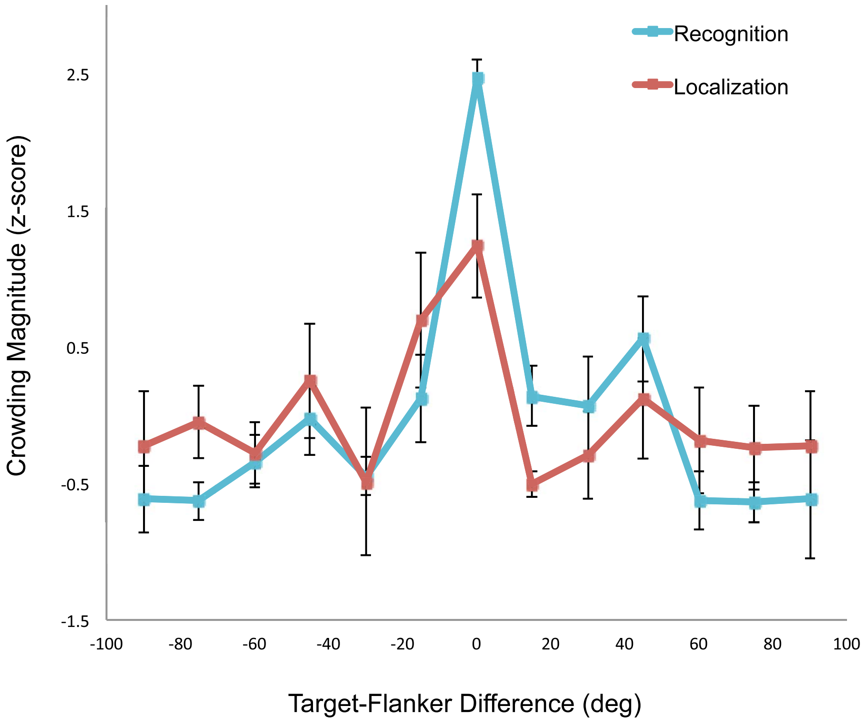

Figure 9. Comparison of average crowding magnitude (blue) and localization error (red) expressed in terms of a z-score and plotted as a function of target–flanker similarity. Bars indicate ±1 *SEM* over observers.

kind of task, saccadic localization follows the perceived rather than the physical location of the target.

Our finding that crowding affects saccadic localization performance is also relevant for understanding the potential role of reduced crowding around saccade initiation (Harrison, Mattingley et al., 2013a; Harrison, Retell et al., 2013). Although—as in our first experiment—orientation-based target selection would have occurred in advance of saccadic planning, target localization should definitely take place just prior to this process. Our finding that target discrimination and localization are similarly influenced by crowding argues against the notion that any modified crowding around saccade initiation might affect performance. As we noted previously, it simply appears to occur too late in the process.

The pattern of results (Figures 7 through 9) suggests that the relationship between target–flanker similarity on the one hand and crowding magnitude and saccadic localization error on the other hand does not follow a monotonically increasing and decreasing function. The reason for the deviations from monotonicity could be the "tilt illusion," which causes exaggeration of perceived target–flanker tilt differences. For crowding in the identity domain, the relationship with the tilt illusion has previously been described (Solomon, Felisberti, & Morgan, 2004). Our present results indicate deviations from a smooth function also for saccadic localization, suggesting that this process is similarly affected by the tilt illusion.

A limitation of the present experiment is that we used a fixed target contrast instead of basing it on individual contrast thresholds. Also, our present experiment may underestimate the actual localization errors due to crowding because of the very presence of the flankers. Although distant from the target, flankers were presented in a regular array. Therefore, participants might have used the flanker position for planning saccades to the target, thereby diminishing the influence of perceived target position. Using flankers with a randomized position around the target might diminish this possibility. Another reason for underestimation could be that participants habituate to making saccades to the limited number of four positions (±8° and ±10° of eccentricity), thereby potentially limiting the magnitude of the saccadic localization error. A final limitation of the present experiment relates to the fact that observers had to choose between two objects only. This is a very limited choice compared to the demand usually posed by natural scenes, which tend to provide much more complex sceneries and choices. Future experiments should therefore consider using more variegated visual stimuli and response requirements.

General conclusion

The main aim of this paper was to determine whether identity and localization performance are similarly affected by crowding, using eye movements as a response measure to integrally assess these aspects. Moreover, given the natural link between peripheral vision, eye movement planning, and crowding and the possibility that saccade-confounded image statistics explain visual crowding (Nandy & Tjan, 2012), the question of whether making eye movements influences performance in crowding tasks and vice versa seemed a relevant and natural one too.

We performed two experiments. A first experiment revealed that crowding is preserved when observers make an eye movement response rather than a manual keyboard response. This indicated that, in our task, preparing eye movements did not affect performance. This is good news for two reasons: First, it indicates that eye movements are a valid measure to assess crowding. Second, it indicates that most of the findings in the existing crowding literature—which primarily have been measured using manual responses—remain valid for understanding *real-life* visual behavior (in which eye movements are the most frequent way by which we select visual information).

Our main experiment used saccades to simultaneously assess changes in recognition and localization errors induced by crowding. We observed that localization errors and recognition performance are similarly affected by target–flanker similarity. Both traditional crowding magnitude (recognition performance) and localization error were larger for flankers that were similar in orientation to the target than for flankers that differed in orientation. Simultaneous assessment of the magnitude of—and the relationship between—errors in crowding in different feature domains is crucial for understanding the underlying mechanisms.

Current feature integration or pooling accounts of crowding (Pelli & Tillman, 2008) are predominantly concerned with the feature aspects of objects. To our knowledge, no specific predictions have been made for position other than averaging (Freeman, Chakravarthi, & Pelli, 2012). Our results argue against an averaging account of crowding of position information (Greenwood et al., 2009; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). Straightforward averaging of the positions of the five Gabors on each side of the stimulus would have given the exact location of the target and reference and thus not resulted in systematic localization errors as a function of target–flanker orientation difference. Instead of pooling, grouping has also been evoked to explain crowding (Banks, Bodinger, & Illige, 1974; Banks, Larson, & Prinzmetal, 1979; Livne & Sagi, 2007; Manassi, Sayim, & Herzog, 2012; Saarela, Westheimer, & Herzog, 2009) with stronger grouping—

and thus crowding, as we find—for higher target–flanker similarity. Position errors would be a consequence of the putative grouping mechanism having "jumbled" target and flanker features, in which case, the increased position uncertainty might be an epiphenomenon. A substitution or source confusion account (Ester, Klee, & Awh, 2014; Strasburger & Malania, 2013) comes close to explaining our present results in that it would predict both localization errors and—as a result of these—recognition errors in the presence of flankers. Whether it could quantitatively account for our present results remains to be determined.

In a previous model study of our group (van den Berg et al., 2012), we showed that perceptual uncertainty about both stimulus positions and identity could explain crowding (assuming that the uncertainty about these aspects depends on flanker proximity). Our present results support the notion that crowding is due to a combination of spatial and identity uncertainty (van den Berg et al., 2012).

Keywords: saccadic localization, eye movements, crowding, location uncertainty, identity uncertainty, peripheral vision

Acknowledgments

This work was supported by The Graduate School of Medical Sciences (GSMS), University of Groningen, The Netherlands. We thank three anonymous reviewers for their useful comments.

Commercial relationships: none.
Corresponding author: Funda Yildirim.
Email: fundayildirim@gmail.com.
Address: Laboratory of Experimental Ophthalmology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

Footnote

[1] Note that the term "target–flanker difference" refers to the difference in orientation between the base orientation of the target and the flanker orientation. In the experiments, the target and the reference were defined by a tilt offset relative to this base orientation For instance, when both target base tilt and flanker tilt were 45°, both the target and reference would still be 5° different from the flankers (e.g., 40° and 50°). For the sake of simplicity of the analyses, the target–flanker differences were grouped based on the difference between the base orientation of the target and the flanker orientation (so 0° in the particular case mentioned here).

References

Allik, J., Toom, M., & Luuk, A. (2003). Planning of saccadic eye movements. *Psychological Research, 67*(1), 10–21, doi:10.1007/s00426-002-0094-5.

Banks, W. P., Bodinger, D., & Illige, M. (1974). Visual detection accuracy and target-noise proximity. *Bulletin of the Psychonomic Society, 4*(4), 411–414.

Banks, W. P., Larson, D. W., & Prinzmetal, W. (1979). Asymmetry of visual interference. *Perception & Psychophysics, 25*(6), 447–456, doi:10.3758/BF03213822.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226,* 177–178.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436. Retrieved from http://booksandjournals.brillonline.com/content/journals/10.1163/156856897x00357.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc., 34*(4), 613–617. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12564564.

Dakin, S., Carlson, A., & Greenwood, J. (2011). Crowding is tuned for perceived (not physical) location. *Journal of Vision, 11*(9):2, 1–13, http://www.journalofvison.org/content/11/9/2, doi:10.1167/11.9.2. [PubMed] [Article]

Dakin, S., Cass, J., Greenwood, J., & Bex, P. (2010). Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *Journal of Vision, 10*(10):14, 1–16, http://www.journalofvision.org/content/10/10/14, doi:10.1167/10.10.14. [PubMed] [Article]

De Pisapia, N., Kaunitz, L., & Melcher, D. (2010). Backward masking and unmasking across saccadic eye movements. *Current Biology, 20*(7), 613–617, doi:10.1016/j.cub.2010.01.056.

De Vries, J. P., Hooge, I. T. C., Wiering, M. A., & Verstraten, F. A. J. (2011). Saccadic selection and crowding in visual search: Stronger lateral masking leads to shorter search times. *Experimental Brain Research, 211*(1), 119–131, doi:10.1007/s00221-011-2660-9.

Deubel, H. (2008). The time course of presaccadic attention shifts. *Psychological Research, 72*(6), 630–640, doi:10.1007/s00426-008-0165-3.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research,* *36*(12), 1827–1837. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8759451.

Ester, E. F., Klee, D., & Awh, E. (2014). Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology, Human Perception and Performance, 40*(3), 1022–1033, doi:10.1037/a0035377.

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences, 10*(8), 382–390, doi:10.1016/j.tics.2006.06.011.

Freeman, J., Chakravarthi, R., & Pelli, D. G. (2012). Substitution and pooling in crowding. *Attention, Perception & Psychophysics, 74*(2), 379–396, doi:10.3758/s13414-011-0229-0.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20–25, doi:10.1016/0166-2236(92)90344-8.

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences, USA, 106*(31), 13130–13135, doi:10.1073/pnas.0901352106.

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2010). Crowding changes appearance. *Current Biology, 20*(6), 496–501, doi:10.1016/j.cub.2010.01.023.

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2012). Crowding follows the binding of relative position and orientation. *Journal of Vision, 12*(3):18, 1–20, http://www.journalofvision.org/content/12/3/18, doi:10.1167/12.3.18. [PubMed] [Article]

Harrison, W. J., Mattingley, J. B., & Remington, R. W. (2013a). Eye movement targets are released from visual crowding. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(7), 2927–2933, doi:10.1523/JNEUROSCI.4172-12.2013.

Harrison, W. J., Mattingley, J. B., & Remington, R. W. (2013b). Releasing crowding prior to a saccade requires more than "attention": response to van Koningsbruggen and Buonocore. *The Journal of Neuroscience, 33,* 1–4.

Harrison, W. J., Retell, J. D., Remington, R. W., & Mattingley, J. B. (2013). Visual crowding at a distance during predictive remapping. *Current Biology, 23*(9), 793–798. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23562269.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188–194, doi:10.1016/j.tics.2005.02.009.

Hooge, I. T. C., Beintema, J. A., & van den Berg, A. V. (1999). Visual search of heading direction. *Experimental Brain Research, 129*(4), 615–628. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10638435.

Hooge, I. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics, 58*(7), 969–976. Retrieved from http://link.springer.com/article/10.3758/BF03206825.

Hunt, A. R., & Cavanagh, P. (2011). Remapped visual masking. *Journal of Vision, 11*(1):13, 1–8, http://www.journalofvision.org/content/11/1/13, doi:10.1167/11.1.13. [PubMed] [Article]

Itti, L., & Koch, C. (2001). Feature combination strategies for saliency-based visual attention systems. *Journal of Electronic Imaging, 10*(1), 161–169. Retrieved from http://dx.doi.org/10.1117/1.1333677.

Korte, W. (1923). Uber die Gestaltauffassung im indirekten Sehen [On the apprehension of Gestalt in indirect vision]. *Zeitschrift Für Psychologie, 93,* 17–82.

Lisi, M., & Cavanagh, P. (2014). The infinite regression illusion reveals dissociation between perception and action. *Journal of Vision, 14*(10): 1221, http://www.journalofvision.org/content/14/1/1221, doi:10.1167/14.10.1221. [Abstract]

Liu, T., Jiang, Y., Sun, X., & He, S. (2009). Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. *Current Biology, 19*(2), 127–132, doi:10.1016/j.cub.2008.11.065.

Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision, 7*(2):4, 1–12, http://www.journalofvision.org/content/7/2/4, doi:10.1167/7.2.4. [PubMed] [Article]

Manassi, M., Sayim, B., & Herzog, M. (2012). Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision, 12*(10):13, 1–14, http://www.journalofvision.org/content/12/10/13, doi:10.1167/12.10.13. [PubMed] [Article]

Maus, G. W., Fischer, J., & Whitney, D. (2011). Perceived positions determine crowding. *PloS One, 6*(5), e19796, doi:10.1371/journal.pone.0019796.

Nandy, A. S., & Tjan, B. S. (2012). Saccade-confounded image statistics explain visual crowding. *Nature Neuroscience*, *15*(3), 463–469, S1–S2, doi:10.1038/nn.3021.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4*(7), 739–744, doi:10.1038/89532.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442. Retrieved from http://booksandjournals.brillonline.com/content/journals/10.1163/156856897x00366.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*(10), 1129–1135, doi:10.1038/nn.2187.

Saarela, T. P., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision, 9*(2):5, 1–11, http://www.journalofvision.org/content/9/2/5, doi:10.1167/9.2.5. [PubMed] [Article]

Solomon, J. A., Felisberti, F. M., & Morgan, M. J. (2004). Crowding and the tilt illusion: Toward a unified account. *Journal of Vision, 4*(6):9, 500–508, http://www.journalofvision.org/content/4/6/9, doi:10.1167/4.6.9. [PubMed] [Article]

Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11):8, 1024–103, http://www.journalofvision.org/content/5/11/8, doi:10.1167/5.11.8. [PubMed] [Article]

Strasburger, H., Harvey, L., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics, 49*(6), 495–508, doi:10.3758/BF03212183.

Strasburger, H., & Malania, M. (2013). Source confusion is a major cause of crowding. *Journal of Vision, 13*(1):24, 1–20, http://www.journalofvision.org/content/13/1/24, doi:10.1167/13.1.24. [PubMed] [Article]

Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty: Its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology, 53,* 471.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision, 11*(5):5, 1–23, http://www.journalofvision.org/content/11/5/5, doi:10.1167/11.5.5. [PubMed] [Article]

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research, 32*(7), 1349–1357, doi:10.1016/0042-6989(92)90227-A.

van den Berg, R., Johnson, A., Martinez Anton, A., Schepers, A. L., & Cornelissen, F. W. (2012). Comparing crowding in human and ideal observers. *Journal of Vision, 12*(6):13, 1–15, http://www.journalofvision.org/content/12/6/13, doi:10.1167/12.6.13. [PubMed] [Article]

van den Berg, R., Roerdink, J. B. T. M., & Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology, 6*(1), e1000646, doi:10.1371/journal.pcbi.1000646.

van Koningsbruggen, M. G., & Buonocore, A. (2013). Mechanisms behind perisaccadic increase of perception. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(28), 11327–11328, doi:10.1523/JNEUROSCI.1567-13.2013.

Yanulevskaya, V., Marsman, J. B., Cornelissen, F., & Geusebroek, J.-M. (2011). An image statistics-based model for fixation prediction. *Cognitive Computation, 3*(1), 94–104, doi:10.1007/s12559-010-9087-7.