Neural correlates of visual crowding

Vitaly Chicherova,⁎, Gijss Plomp, Michael H. Herzog

Abstract

In visual crowding, target discrimination strongly deteriorates when flanking elements are added. We have recently shown that crowding cannot be explained by simple low-level interactions and that grouping is a key component instead. We presented a vernier flanked by arrays of vertical lines. When the flanking elements had the same lengths as the vernier, offset discrimination was strongly impaired. When longer flanking elements were presented, crowding was weaker. We proposed that crowding is strong when the flanking group with the target (equal length flanking). When the target segregates from the flanking, crowding is weaker (long flanking). To understand the neurophysiological mechanisms of grouping in crowding, here, we adapted the above vernier paradigm to a high-density EEG study. The P1 component reflected basic stimulus characteristics (flanker length) but not crowding. Crowding emerged slowly and manifested as a suppression of the N1 component (after 180 ms). Using inverse solutions, we found that the N1 suppression was caused by reduced neural activity in high-level visual areas such as the lateral occipital cortex. Our results suggest that crowding occurs when elements are grouped into wholes, a process reflected by the N1 component.

Introduction

Perception of an object is strongly influenced by neighboring elements. For example, a vertical line can appear strongly tilted when flanked by tilted lines (the tilt illusion, Gibson, 1937). Crowding is an example where neighboring elements strongly deteriorate performance (Bouma, 1970; Flom et al., 1963; Levi, 2008; Pelli and Tillman, 2008; Strasburger et al., 1991). For example, a letter that is clearly visible when presented alone cannot be identified when flanked by other letters. The letter itself is visible but its identification is severely impaired.

One influential explanation proposes that crowding occurs by pooling of nearby features. An element is first analyzed by neurons with small receptive fields coding for basic features. The output of these neurons is then fed into neurons with larger receptive fields pooling features of nearby elements. This pooling leads to confusion of features (Greenwood et al., 2009; Pelli et al., 2004; Wilkinson et al., 1997). Other explanations focus on centroids and lateral inhibition (Levi and Carney, 2009; Westheimer and Hauske, 1975; Wilson, 1986).

We have recently shown that grouping, rather than pooling, plays a key role in both peripheral and foveal crowding (Malania et al., 2007; Manassi et al., 2012; Sayim et al., 2008). For example, we presented a vernier stimulus that comprised two vertical lines which were slightly offset horizontally. Observers were asked to indicate the offset direction (Fig. 1A). When the vernier was flanked by single lines of equal length, performance strongly deteriorated. This is a classical crowding effect. Interestingly, crowding diminished when single long or short flanking elements were presented and almost disappeared when arrays of short or long flankers were presented. These results cannot be explained with simple pooling models because longer and additional flanking should increase crowding strength, e.g., because more target-irrelevant information is pooled with the relevant target signal. We proposed that grouping determines crowding. Crowding is strong (equal length flanks), when the flanks group with the target. Crowding is weaker, when the target segregates from the flanks (Malania et al., 2007; Manassi et al., 2012; Sayim et al., 2008; see also Livne and Sagi, 2007).

Whereas there is a vast amount of studies on crowding, there are only a few human neuroimaging studies (Anderson et al., 2012; Bi et al., 2009; Fang and He, 2008; Freeman et al., 2011; Millin et al., 2013) and almost nothing is known about the time course of crowding (but see early EEG studies on vernier acuity (Steinman et al., 1985; Zak...
and Berkley, 1986)). To understand the neurophysiological mechanisms of crowding and contextual modulation in general, we adapted the above vernier paradigm to a high-density EEG study and found that crowding is reflected in late but not early visual processing, indicating that crowding occurs during higher-level visual processing when elements are grouped into wholes.

Fig. 1. Experiment 1a: foveal presentation. (A–B) In the target conditions, observers indicated whether the vernier was offset to the left or right (an offset to the right is shown here). In the control conditions, observers pushed randomly any button they wished. Performance in the target conditions was best when the vernier was flanked by long flankers, intermediate for short, and worst for equal length flankers. (C–D) Global field power (GFP). The P1 component in both the target (C) and control (D) conditions reflects flanker length. The N1 component reflects performance level in the target conditions. In the control conditions, this is not the case. The horizontal turquoise bars under the GFP curves indicate time periods of a significant main effect of Flanker length, the red bars — of significant interaction effects (see Supplementary Fig. 1 for the time course of the p-values). The main effect of Target/Control was not significant. (E) Amplitudes of the P1 component averaged over the time window 80–156 ms. Average topographical maps for each condition are shown above the bar plots. (F) Amplitudes of the N1 component averaged over the time window 176–202 ms. (G) Inverse solutions for the target conditions (LAURA inverse solution technique). A one-way ANOVA with the factor Flanker length was conducted in the source space in the N1 time window. Significant sources were attributed to two areas, one in the lateral occipital cortex (with parietal and temporal overlap) and a smaller area on the medial surface (precuneus), in the right hemisphere. The color scale shows current density differences between the long flanker condition and the equal length flanker condition indicating which sources contribute most strongly to the scalp differences (Talairach coordinates of the maximum: 48, −66, 1 mm). (H) Current densities averaged over all significant source points. As with N1 amplitudes, current densities were highest in the long condition, intermediate in the short, and lowest in the equal length condition. In the control conditions, current densities in the same source points were not significantly different. Error bars show the standard errors of mean. P-values for two-tailed paired t-tests are shown above the bar plots.
Experiment 1

Methods

Participants
All observers had normal or corrected-to-normal visual acuity in at least one eye as measured with the Freiburg Visual Acuity test, i.e., acuity values were above 1.0 (Bach, 1996). Handedness was assessed with the Edinburgh inventory (Oldfield, 1971). All participants provided written consent. The experiment was approved by the local ethics committee.

12 paid volunteers joined Experiment 1a. One observer was excluded due to high amplitude noise in the EEG and another one was excluded due to low performance in the long flanker condition. Ten observers (4 females), aged 22–40 years (mean (SD) = 27.4 (6.2)) were included in the analysis. All but two of the observers were right-handed. All but two of the observers were naive to the purpose of the experiments.

10 paid volunteers (3 females), aged 18–27 years (mean (SD) = 22.2 (2.6)), joined Experiment 1b. All but two of the observers were right-handed. All observers were naive to the purpose of the experiments.

Apparatus
Stimuli were generated with a PC and presented on an X-Y-display (HP-1332A, green phosphor P31). Stimulus lines were drawn by dots at a rate of 1000 kHz and a dot pitch of 250 μm. Refresh rate was 200 Hz. Stimulus luminance was 80 cd/m², as measured with a Minolta LS-100 luminance meter. The background was black. Observers were seated in a dimly-lit (0.5 lx), electrically isolated room at 150 cm from the monitor (foveal stimuli) or 75 cm (peripheral stimuli). Observers held one response button in the right hand, the other one in the left hand.

Stimuli
In Experiment 1a, a vernier target was presented foveally. The vernier consisted of two vertical lines, each 10.3′ (arcmin) long, with a 1.2′ vertical gap. The lower line was offset to the left or right relative to the upper line. Observers indicated whether the vernier was offset to the right or left by pressing the corresponding button. The offset direction changed randomly from trial to trial. The number of left and right offsets was balanced in each block. The vernier was flanked by 20 regularly placed vertical lines (flankers): ten to the left and ten to the right. Horizontal distance between the flankers as well as between the vernier and the flankers was 3.33′. Flanker length: 10.3′ (short), 21.8′ (equal length), and 43.5′ (long, Fig. 1A). Equal length flankers had the same length as the vernier. Response accuracy was emphasized over speed. The position of the stimuli randomly jittered from trial to trial within 1′ to compensate for potential inhomogeneities of the screen. In control conditions, the vernier was omitted, and observers were asked to push buttons randomly. Equal numbers of target and control trials were presented.

Unless stated otherwise, the stimulation sequence was as follows. The screen stayed blank for 500 ms. A fixation cross (1° diameter) appeared at the center of the screen. After a random interval of 700–1300 ms, the fixation cross disappeared and the stimulus was presented. After 200 ms, the stimulus disappeared and observers had to respond within 3 s. After the response, the next trial started.

In Experiment 1b, we repeated the same experiment with 10 new observers. This time, stimuli were presented in the periphery at 3.9′ to the right of the fixation cross. To account for peripheral vision, we increased line lengths and spatial separations between lines. The vernier lines were 40′ long, with a 4′ vertical gap, flanked by six lines on both sides. Horizontal distance between the flankers as well as between the vernier and the flankers was 23.33′ (the innermost flanker was 1.6′ to the right of the fixation). Flanker length: 40′ (short), 84′ (equal length), and 168′ (long). The position of the stimuli randomly jittered from trial to trial within 2′ to compensate for potential inhomogeneities of the screen.

Procedure
First, a behavioral training session was conducted (1 h) with all stimulus conditions to achieve stable performance in the EEG experiment. All conditions were trained equally and in a counterbalanced order. Based on the psychometric functions of the target conditions, we determined, for each observer individually, the vernier offset size corresponding to 90% correct responses in the easiest, long flankers, condition and used this value for the EEG session.

Second, the EEG session was conducted. Vernier offset size was fixed at the level as defined in the training sessions and was the same in all target conditions. All six conditions were randomly interleaved and presented in 10 blocks of 84 trials (140 per condition). The duration of one block was around 3 min. Short breaks were allowed between blocks.

EEG setup and preprocessing
The EEG was recorded with a Biosemi Active Two system with 192 Ag/AgCl sintered active electrodes. The cap size and placement were adjusted individually; the Cz electrode was positioned halfway between inion and nasion. The set of the electrodes uniformly covered the entire scalp. Electrooculogram (EOG) was recorded with four electrodes placed 1 cm above and below the right eye, and 1 cm laterally to the right and left lateral canthus. Vertical and horizontal EOG (VEOG and HEOG) channels were derived from EOG channels as difference potentials between vertically arranged channels and horizontally arranged channels respectively. The data were recorded with a sampling rate of 2048 Hz, filtered off-line with a bandpass of 0.2–30 Hz (2nd order Butterworth digital filter), and downsampled off-line to 512 Hz. The VEOG channel for the eye blink artifact rejection was filtered with a bandpass of 0.2–15 Hz. The VEOG and HEOG channels for the saccade artifact rejection were filtered with a bandpass of 0.2–30 Hz. EEG data were recorded with a CMS (Common Mode Sensor) reference and re-referenced off-line to the average of all electrodes (average reference).

For the EEG analysis, we used EEGLAB (Delorme and Makeig, 2004) and ERPLAB (http://erpinfo.org/erplab) toolboxes for MATLAB. We extracted EEG epochs from 200 ms before to 500 ms after stimulus onset. First, epochs with reaction times less than 300 ms and more than 3 s were rejected. Second, epochs contaminated by eye blinks and saccades were rejected. Eye blinks were detected from the VEOG channel using a “differential step” function in a 200 ms moving window (Luck, 2005, p. 162). Vertical and horizontal saccades were detected from the VEOG and HEOG channels using a 200 ms step function (Luck, 2005, p. 165). Thresholds for rejection were defined individually due to large variability across observers. Third, artifacts in EEG channels were detected using a modified version of the SCADS algorithm (statistical control of artifacts in dense array EEG/MEG studies, Junghöfer et al., 2000). We removed artifacts based on three different artifact measures: mean power spectral density in 20–40 Hz frequency range of the nonfiltered EEG (muscular-related artifacts), standard deviation of potential in an epoch (all types of high amplitude persistent artifacts), and difference between maximum and minimum values in an epoch (all types of high amplitude transient and persistent artifacts). On average, 15% of trials were rejected and 9% of data were interpolated. Event-related potentials (ERPs) were obtained by averaging together artifact-free epochs separately for each experimental condition including both trials with correct and incorrect responses. Stimulus onset was taken as the reference (0 ms). Average pre-stimulus potential (−200 to 0 ms) was subtracted from the ERPs.

ERP analysis
We determined global-field power (GFP) for each observer and each condition. GFP is the standard deviation of potentials across all
electrodes at a given time point (Lehmann and Skrandies, 1980). GFP reflects overall response strength in ERPs and is calculated as:

\[ GFP(t) = \sqrt{\frac{\sum_{i=1}^{n} (u_i(t) - \bar{u}(t))^2}{n}} \]

where \( u_i(t) \) is the potential at electrode \( i \) at time point \( t \), \( \bar{u}(t) \) is the average across all electrode potentials at time \( t \), and \( n \) is the number of electrodes.

Repeated-measures ANOVAs with the factors Flanker length (short, equal length, long) and Target/Control were conducted at each time point of the GPPs. The effect of interest is the interaction effect of the two factors (Flanker length × Target / Control). The interaction term reflects processes characterizing different crowding strengths in the three target conditions. The low-level properties (flanker length and presence of the vernier) are represented by the main effects of the Flanker length and Target/Control factors.

Statistical significance of the interaction at each single time point was assessed with a false discovery rate (FDR) control method based on the Benjamini–Hochberg (BH) algorithm (Benjamini and Hochberg, 1995), which was shown to correctly control FDR in a realistic simulated ERP data (Groppe et al., 2011). A p-value threshold for rejection of the null hypothesis was calculated such that, on average, \( q = 10\% \) of the rejected null hypotheses were false alarms (type I errors). The false-discovery rate is well interpretable at any level of \( q \) (proportion of discoveries that are false). Even with \( q = 10\% \), the vast majority (90\%) of the discoveries are true effects. Time points from 50 to 500 ms were taken into calculation of the FDR thresholds (50 ms is the latency of the earliest visual response in the cortex). Whenever ANOVAs were performed, Huynh–Feldt correction for non-sphericity was used if necessary (Huynh and Feldt, 1976). Degrees of freedom are reported multiplied by the correction factor.

To estimate brain areas generating the ERP effects in the target conditions, we used the Local Auto Regressive Average inverse solution (LAURA, Grave de Peralta Menendez et al., 2001; Michel et al., 2004; Plomp et al., 2009, 2010). Source analysis was performed in the Cartool software (https://sites.google.com/site/flmbl/cartool). In the MNI 152 brain model, we placed evenly 4022 source points and used the model identical to Plomp et al. (2009, 2010). We report Talairach coordinates (Talairach and Tournoux, 1988). MNI coordinates were converted to Talairach coordinates using non-linear transformation (http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach).

**Results**

**Experiment 1a**

In the fovea, we presented a vernier flanked by arrays of lines of three different lengths: short, equal length and long (Fig. 1A). Observers indicated whether the vernier was offset to the right or left. As in previous studies (Malania et al., 2007; Manassi et al., 2012), performance was indicated whether the vernier was offset to the right or left. As in previous studies showing that the P1 amplitudes did not depend on the presence of the target vernier. Hence, the P1 component reflects flanker length in accordance with previous studies showing that the P1 component is sensitive to stimulus size and other low-level properties (e.g., Busch et al., 2004). The main effect of Flanker length was significant also during later time periods but we did not analyze them because the main effects of ANOVA were not the primary interest of the study.

Crowding effects. Contrary to the P1, the N1 component reflected performance level and, hence, crowding (176–202 ms, interaction effect: F(1.4, 12.9) = 6.9, p = 0.01). To further analyze the significant interaction, we performed one-way ANOVAs separately on the target and control conditions. Amplitudes of the N1 component were highest for the long, intermediate for the short, and lowest for the equal length flankers in the target conditions (one-way ANOVA on the target conditions: F(2, 18) = 10.1, p = 0.001, Fig. 1F). The N1 components of the control conditions were of similar amplitudes (one-way ANOVA on the control conditions: F(1.8, 16.2) = 2.5, p = 0.12). There were significant interaction effects also during later time windows (the P3 component) but we here focus only on the earliest detectable effects of crowding.

Mean reaction times in the target conditions were above 900 ms and similar across the different flanker configurations, and thus, did not affect the P1 and N1 components (Table 1 in Supplementary data, p > 0.05, pair-wise t-tests). There was no significant response bias (51%, 50%, and 53% left responses in the short, equal length, and long flanker target conditions, respectively). We used electrical source imaging (Grave de Peralta Menendez et al., 2001; Michel et al., 2004) to estimate current source densities throughout the brain in the target conditions in the N1 time window (176–202 ms). We estimated current densities for each observer’s ERPs averaged in the N1 time window. To locate the brain areas underlying the crowding effect, we calculated one-way ANOVAs with the factor Flanker length in the target conditions. We here used a one-way ANOVA since there was no significant difference among flanker-levels in the control conditions, and the one-way ANOVA in the target conditions is a more sensitive test than the interaction effect in a two-way ANOVA. We found two areas in the right hemisphere that reliably contributed to the N1 differences (Fig. 1G). The first area was in the middle occipital gyrus (lateral occipital cortex) overlapping with the inferior parietal and posterior middle temporal cortex. The second, smaller activation patch was on the medial surface in the precuneus and the posterior cingulate gyrus. Current densities, on average, were highest in the long, intermediate in the short, and lowest in the equal length condition (F(2, 18) = 16.3, p = 0.001, Fig. 1H) indicating that these sources are underlying the N1 differences. In the control conditions, source activities were not significantly different (F(2, 18) = 0.2, p = 0.8). The right lateralization of the effect is consistent with the right-lateralized N1 topography (see topographical maps in Fig. 1F).

**Experiment 1b**

Usually, crowding is studied in peripheral vision. To show that peripheral and foveal crowding share similar electrophysiological correlates, we performed the same experiment with peripherally instead of foveally presented stimuli at 3.9° of eccentricity for ten new observers (Fig. 2). Psychophysical results were similar to the foveal presentations (Fig. 2A). Accuracy was best in the long flanker condition, intermediate in the short, and worst in the equal length flanker condition (92 ± 2%, 76 ± 2%, and 68 ± 3% respectively, F(1.8, 15.9) = 34.3, p = 3 × 10–6). While ERP differences were not as pronounced as with
foveally presented stimuli, qualitatively the results were very similar. The N1 component (188–202 ms) reflected the performance level in the target conditions (interaction of the factors Flanker length and Target/Control; F(1.9, 17.1) = 11.7, p = 0.001). While the N1 component was significantly different in the target conditions (one-way ANOVA on the target conditions: F(2, 18) = 5.7, p = 0.01), there was no effect in the control conditions (one-way ANOVA on the control conditions: F(1.4, 12.9) = 0.16, p = 0.8). A source localization analysis similar to Experiment 1a, however, did not reveal significant effects. However, the fact that crowding was observed at N1 latencies with topographies comparable to Experiment 1 supports the idea that crowding in foveal and peripheral vision has similar electrophysiological correlates.

In the target conditions of both experiments, N1 amplitudes reflected the performance level in a clear cut fashion, i.e., N1 amplitudes were highest for the best performance level (long flankers), medium for the medium performance level (short flankers), and lowest for the lowest performance level (equal length flankers). When we designed the experiment, we did not expect such a clear cut result but rather that N1 amplitudes depended on both performance level and flanker length. For this reason, we added the control conditions and aimed for an interaction effect (if any). In hindsight, we could have omitted the control conditions and just have planned a simple one-way ANOVA on the target conditions. Since target and control conditions were presented interleaved form trial to trial, we decided to present the experiment as originally planned.

Experiment 2

Methods

Apparatus, stimulus configuration, EEG setup, and analysis were similar to Experiment 1. Stimuli were presented foveally.

Participants

Eleven paid new volunteers participated, first, in a length discrimination task and, then, in a vernier discrimination task. One observer was excluded due to high-amplitude noise in the EEG. Ten observers (3 females), aged 19–24 years (mean (SD) = 21.8 (2.0)) were included in the analysis. All observers were right-handed. All observers were naïve to the purposes of the experiment. All but one of the observers were naïve to the vernier discrimination task.

Length discrimination task

We presented long and equal length flankers together with the vernier (target conditions) or without it (control conditions). Observers pushed the left button for long flankers and the right button for equal length flankers or vice versa (counterbalanced), irrespectively of whether or not the vernier was present. The vernier offset in the target conditions was set to 0°. We call the conditions ‘target’ and ‘control’ to be consistent with Experiment 1, although the vernier was not a target in this task. All four conditions were randomly interleaved and presented in 6 blocks of 80 trials (120 trials per condition). A two-way repeated-measures ANOVA was conducted on the averaged GFP potentials in the N1 time window as found in Experiment 1a, with the factors Flanker length (equal length, long) and Target/Control.

Vernier offset discrimination

We presented the vernier flanked by long or equal length flankers. As in Experiment 1, observers indicated the offset direction. We chose three different vernier offset sizes individually for each observer to obtain good, medium, and low performance levels (mean ± SD: 38 ± 15°, 12 ± 3° and 3.0 ± 1.1°). All six conditions were randomly interleaved and presented in 10 blocks of 72 trials (120 trials per condition). The ERP analysis was similar to the length discrimination task, except for the different set of factors.
Results

In Experiment 2, we show that the N1 does not simply reflect the difficulty of the task but rather crowding, i.e., spatial processing. To this end, we used three vernier offset sizes leading to three different performance levels. As expected, we did not find any effect of task difficulty on N1 amplitudes. Next, we show that the N1 is sensitive to attention, i.e., whether or not observers perform the vernier task. Passive viewing of the stimuli leads to attenuated N1 effects.

Length discrimination task

Observers discriminated the length of flankers (Figs. 3A–B). Observers were at ceiling performance (99.2% correct responses) when discriminating the length of flankers. We analyzed average GFP amplitudes in the N1 time window, as determined in Experiment 1a (176–202 ms), and performed an ANOVA on the averaged amplitudes. The interaction was significant (F(1, 9) = 12.8, p = 0.006) along with a significant effect of Flanker length (F(1, 9) = 8.8, p = 0.02) and a non-significant effect of Target/Control (F(1, 9) = 1.0, p = 0.4).

Vernier offset discrimination and figural layout

In Experiment 1, we found that performance levels correlated well with the N1 amplitudes. Hence, it may be that N1 amplitudes simply reflect whether a task is easy or hard to perform independent of what underlies task difficulty. Here, we show that the N1 amplitudes reflect the spatial aspects underlying crowding and not task difficulty. Observers performed a vernier offset discrimination task as in Experiment 1. We used three vernier offset levels leading to three different performance levels in both the equal length and long flanker conditions (Fig. 3D). Even though performance levels were clearly different (main effect of task difficulty: F(1.2, 10.4) = 314, p = 3 × 10^{-9}), N1 amplitudes were almost identical for the three offset sizes (Figs. 3C and E, main effect of difficulty: F(2, 18) = 0.4, p = 0.6; interaction effect: F(1.8, 16.5) = 0.1, p = 0.9). Hence, task difficulty per se is not reflected...
Crowding is usually thought to be mediated by local interactions of neurons sensitive to the target and nearby flanking elements. The pooling hypothesis proposes that neural responses to nearby elements are pooled. Nearby flankers deteriorate the representation of the target because they add target-relevant information (Levi et al., 1985; Pelli et al., 2004; Wilkinson et al., 1997). However, we have recently shown that the global stimulus configuration is an equally important factor in crowding which cannot easily be explained by pooling (Malania et al., 2007; Manassi et al., 2012; Saarela et al., 2009; Sayim et al., 2008; see also Chakravarthi and Pelli, 2011; Livne and Sagi, 2007; Wolford and Chambers, 1983; Yeotikar et al., 2011). We proposed that target-flanker grouping plays a crucial role in crowding. A target that groups with the flankers becomes part of a global structure, loses its individuality, and this deteriorates access to the target's features (see also He et al., 1996; Parkes et al., 2001). If the target segregates from the flankers it stands out of the global structure, its features are easily accessible, and hence crowding strength is reduced (see also, Banks and Prinzmetal, 1976; Wolford and Chambers, 1983). For example, if the vernier is flanked by short lines, crowding decreases monotonically the more flankers are added (Malania et al., 2007; Manassi et al., 2012). We proposed that the more lines are presented the stronger they group because of the Gestalt cue of similarity (Malania et al., 2007; Manassi et al., 2012). Performance improved for up to 4 flankers (inter-flanker distance 3.33 arcmin) on each side of the vernier (Malania et al., 2007). The extension of an array of 4 flankers was 14 arcmin, i.e., much beyond the interaction area of a single flanker which is about 4–5 arcmin in the fovea (Levi et al., 1985; Westheimer and Hauske, 1975). Hence, global rather than local aspects determine crowding strength. In addition, these results suggest a close relationship between grouping and crowding. Here, we investigated the neural mechanisms of grouping in crowding.

Our behavioral findings reproduce previous results from our group (Malania et al., 2007; Manassi et al., 2012; Saarela et al., 2009). In the EEG recordings, the P1 component, around 125 ms, reflected flankers length with highest amplitudes for the long flankers, intermediate for equal length, and lowest for short flankers. Hence, the P1 reflects low-level properties of the stimulus such as its size and energy. Our results are in accordance with previous studies where the P1 component also reflected low-level characteristics of the stimuli such as luminance, eccentricity, and size (Busch et al., 2004; Johannes et al., 1995).

The N1 component, around 190 ms, did not reflect stimulus size but correlated with crowding. The N1 was highest in the long, intermediate in the short, and lowest in the equal length flanker condition. Using EEG source localization, we found that the differences in the N1 amplitudes in the three flanker conditions were generated by parts of the lateral occipital cortex and the neighboring parietal and middle temporal areas (Figs. 1G and 3F). The lateral occipital area includes the lateral occipital complex (LOC), a high-level visual area involved in object recognition (Grill-Spector et al., 2001). In our experiments, primary visual cortex (V1) was not significantly involved in crowding. Hence, it seems that the effects of grouping in crowding are not present in the first sweep of information processing, reflected by the P1 component, but occur in a dynamic fashion in high-level visual and non-visual areas, reflected by the N1 component.

Our results are in line with other studies where spatial grouping is best reflected in the modulation of the N1. For example, the N1 increases when a texture is embedded in another texture compared to a homogenous texture (Bach and Meigen, 1992, 1998; Caputo and Casco, 1999) or when symmetric dot patterns are presented compared to random dot patterns (Noria et al., 2002). Likewise, N1 amplitudes are higher for illusory figures, such as Kanizsa squares, compared to when inducers do not give rise to a good figure (Herrmann and Bosch, 2001; Murray et al., 2002). The same holds true for contour integration (Machilsen et al., 2011; Mathes et al., 2006; Tanskanen et al., 2008). All these studies suggest that the N1 increases when global, complex structures emerge from integrating single elements into wholes. Likewise in

in the N1 amplitudes. Replicating the results of Experiment 1, N1 amplitudes were clearly higher in the long than in the equal length conditions (main effect of Flanker length: F(1, 9) = 28.3, p = 0.0004). We propose that the N1 amplitudes reflect the figural layout of the stimulus, i.e., whether elements can be grouped in one figure (equal length condition) or split up in sub-figures, e.g., two textures of long lines and the central vernier as in the long condition. This grouping determines crowding strength. It is important to note that crowding strength is not reflected simply in the performance level because crowding strength depends on the spatial layout of the stimulus and the vernier offset size. Hence, crowding strength can only be compared across flanker configurations when vernier offset sizes are identical, as in Experiment 1.

Do the N1 amplitudes reflect also the figural aspects of the stimuli when there is no crowding involved? There are also N1 amplitude differences when observers perform the length discrimination task, which does not involve crowding. However, the differences are much more pronounced when crowding is involved. In the vernier discrimination task, the N1 amplitude difference between the long and equal length flanker target conditions is 0.95 μV, similar to what we observed in Experiment 1a (1.01 μV). In the length discrimination task, this difference is strongly and significantly reduced (0.57 μV, t-test between the vernier discrimination and length discrimination tasks: p = 0.05).

**Suppression**

It is an open question whether neural activity is reduced in crowding (equal length flanker condition) or enhanced in “uncrowding” (long flanker condition). In the long flanker target conditions, the N1 amplitudes were similar in both tasks, indicating that there was no enhancement in uncrowding (2.75(0.52) μV in the vernier discrimination task; 2.84(0.52) μV in the length discrimination task; t-test: p = 0.5, Figs. 3B and E). In the equal length target conditions, the N1 amplitudes were lower in the vernier discrimination than in the length discrimination task indicating that the N1 component is suppressed during crowding (1.80(0.39) μV in the vernier task; 2.27(0.48) μV in the length task; t-test: p = 0.01). Hence, our results suggest that the N1 component is suppressed in crowding.

**Source localization**

Source localization in the vernier discrimination task confirmed the results of Experiment 1a (Figs. 3F–G). First, we localized sources that showed a significant effect of Flanker length during the N1 component (176–202 ms, Fig. 3F). On average, these sources were more active in the long flanker conditions compared to the equal length flanker conditions (main effect of Flanker length: F(1, 9) = 16.4, p = 0.003) in accordance with the GFP results (Fig. 3G). The source activities did not depend on task difficulty (main effect of difficulty: F(1, 16.1) = 2.4, p = 0.13; interaction: F(2, 18) = 0.2, p = 0.8). For each significant source point, we calculated the mean difference between the long and equal length flanker conditions: (Long Easy + Long Medium + Long Hard − Equal Easy − Equal Medium − Equal Hard) / 3. The largest effects were localized in the same areas as in Experiment 1a (lateral occipital cortex with temporal and parietal overlaps, mostly in the right, but also partly in the left hemisphere).

**Discussion**

In crowding, performance deteriorates when a target element is neighbored by flankers. Contrary to backward masking, where the target itself is rendered invisible, only the identification of target features is deteriorated in crowding. For example, vernier offset discrimination is strongly affected by flanking lines while the vernier itself is clearly visible (Levi et al., 1985; Westheimer and Hauske, 1975). In this respect crowding is similar to contextual modulation, spatial illusions, and visual search where context strongly modulates perception.

In crowding, performance deteriorates when a target element is neighbored by flankers. Contrary to backward masking, where the target itself is rendered invisible, only the identification of target features is deteriorated in crowding. For example, vernier offset discrimination is strongly affected by flanking lines while the vernier itself is clearly visible (Levi et al., 1985; Westheimer and Hauske, 1975). In this respect crowding is similar to contextual modulation, spatial illusions, and visual search where context strongly modulates perception.
our experiment, a much more complex structure is present in the short and long flanker conditions (a vernier and two groups of flankers) compared to the equal length condition where all elements are bound into one homogeneous array (Fig. 1A).

The brain areas underlying the N1 modulation in crowding we found (LOC and posterior temporal and parietal areas) are similar to areas found in other studies on perceptual grouping. For contour integration and illusory contours, the N1 was modulated by sources in the LOC, posterior temporal and occipito-parietal cortex (Murray et al., 2002; Tanskanen et al., 2008). fMRI studies showed that grouping is processed in extrastriate and posterior temporal areas (Hirsch et al., 1995; Kubilius et al., 2011; Mendola et al., 1999). The parietal cortex was found to be important for the individuation of objects and attentional selection of relevant information (Bisley and Goldberg, 2003; Chambers et al., 2004; Xu, 2009). It seems that the lateral occipital, temporal and parietal cortex interact with each other to mediate grouping and selection of a target in a cluttered scene (Fang et al., 2008; Zaretskaya et al., 2013).

Similarly in our study, we found that the lateral occipital, posterior temporal, and inferior parietal areas are the major sources of the N1 modulation in crowding.

Our findings are in line with previous fMRI work showing that high-level visual areas are involved in crowding (Anderson et al., 2012; Bi et al., 2009; Freeman et al., 2011). Anderson et al. (2012) found that higher visual areas reflected crowding better than earlier ones and Freeman et al. (2011) showed that crowding decreases the correlations between activities in high- and low-level visual areas.

The N1 effects are not related to task difficulty itself but to spatial processing of the vernier embedded in the flankers, i.e., to crowding. When we presented the vernier with large or small offsets, performance strongly changed with offset size but N1 amplitudes did not (Figs. 3C–E). Likewise, earlier studies found that the N1 component is not affected by task difficulty (Bankó et al., 2011; Senkowski and Herrmann, 2002; Vogel and Luck, 2000). Our results imply that the N1 component reflects spatial processing of grouping and crowding and not simply the task difficulty.

As mentioned, we propose that N1 amplitudes reflect the effects of grouping in crowding. Even though we can only speculate about the underlying neural mechanisms, our results suggest that the vernier is actively suppressed in crowding. In the equal length condition, N1 amplitudes were larger when observers discriminated the length of the flankers compared to vernier offset discrimination (Fig. 3). In the long condition, there was no such effect, i.e., N1 amplitudes were comparable. This suppression may be related to attention effects in crowding (He et al., 1996; Intriligator and Cavanagh, 2001; Petrov and Meleshkevich, 2011; Strasburger, 2005). Resolution of attention in the fovea is around 3–5 arcmin (Intriligator and Cavanagh, 2001) which is close to the inter-element distance of our stimuli (3.33 arcmin). In the equal length flanker condition, attentional selection of the vernier is impaired leading to reduced N1 amplitudes and diminished activation of the lateral occipital and parietal cortex. Long flankers, however, ungroup from the target, and alternatively, feature-based selection of the target can overcome the coarse resolution of spatial attention (Intriligator and Cavanagh, 2001).

Finally, our results provide a bridge between foveal and peripheral crowding. In previous studies, crowding of vernier targets showed similar characteristics in the fovea and periphery (Levi et al., 1985; Malania et al., 2007; Manassi et al., 2012). In the current work, we found that peripheral crowding and foveal crowding also have the same electrophysiological correlate — the N1 component. The N1 is reduced for crowded stimuli both in the fovea and in the periphery. This result implies similarity of underlying mechanisms of vernier crowding across eccentricities and suggests that, with vernier stimuli, crowding can be studied well in the fovea.

To summarize, the P1 component reflects flanker length and the overall luminance of the stimuli (pixel energy). The N1 component correlates highly with crowding (Experiment 1). As Experiment 2 shows, the N1 does not simply reflect performance level, i.e., task difficulty, because the N1 amplitudes were almost identical in the easy, medium, and hard long flanker conditions. The same holds true for the equal length flanker conditions. We argue that the N1 amplitudes reflect whether or not elements can be grouped into one figure. In the long flanker condition, the elements group in three sub-figures and, hence, N1 amplitudes are higher than in the equal length flanker conditions where all lines make up a homogeneous array. This is very well in accordance with the notion that crowding depends on grouping. When target and flankers group, crowding is strong (equal length flankers). When target and flankers ungroup (long flankers), crowding is weak. Whereas the spatial aspects of crowding are reflected also when observers do not perform the vernier discrimination, these effects are much weaker than when observers perform the vernier discrimination task, where crowding plays a role. We argue that crowding is due to active suppression in the equal length flanker condition.

Acknowledgment

Vitaly Chicherov is supported by the SNF grant “Basics of visual processing: what crowds in crowding?” of the Swiss National Science Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.02.021.

References


Fang, F., He, S., 2008. Crowding alters the spatial distribution of attention modulation in human primary visual cortex. J. Vis. 8, 6.1–6.3.


Cannon, J.J., 1937. Adaptation, after-effect, and contrast in the perception of tilted lines. II. Simultaneous contrast and the areal restriction of the after-effect. J. Exp. Psychol. 20, 553–569.
