Vision Research 50 (2010) 509–521

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Electrophysiological correlates of figure–ground segregation directly reflect perceptual saliency

Sirko Straube [a,*], Cathleen Grimsen [a], Manfred Fahle [a,b]

[a] Department of Human Neurobiology, University of Bremen, Hochschulring 18, 28359 Bremen, Germany
[b] The Henry Wellcome Laboratories for Vision Sciences, Dept. of Optometry & Visual Science, City University, Northampton Square, London, EC1V 0HB, United Kingdom

ARTICLE INFO

Article history:
Received 30 June 2009
Received in revised form 27 October 2009

Keywords:
Cue combination
Event-related potentials
Figure–ground segregation
Human
Identification
Time–frequency analysis

ABSTRACT

In a figure identification task, we investigated the influence of different visual cue configurations (spatial frequency, orientation or a combination of both) on the human EEG. Combining psychophysics with ERP and time–frequency analysis, we show that the neural response at about 200 ms reflects perceptual saliency rather than physical cue contrast. Increasing saliency caused (i) a negative shift of the posterior P2 coinciding with a power decrease in the posterior θ-band and (ii) an amplitude and latency increase of the posterior P3. We demonstrate that visual cues interact for a percept that is non-linearly related to the physical figure–ground properties.

1. Introduction

When looking at a visual scene, we simultaneously identify different objects without caring about the visual features underlying this percept. However, we often recognize objects based on a conjunction of cues instead of using only one single visual cue (e.g. depth, color or orientation in space). The percept of the object is coherent, meaning that the information from various cues is integrated by the visual system.

To what extent is it easier to perceive an object defined by multiple cues instead of one defined by a single cue? Several behavioral studies dealing with cue combination tried to answer this important question with the result that combination seems to depend on task and cue type. Therefore, the amount of behavioral benefit from cue combination is still under debate, with the majority of studies observing cue interaction (Abele & Fahle, 1995; Kubovy & Cohen, 2001; Kubovy, Cohen, & Hollier, 1999; Meinhardt & Persike, 2003; Meinhardt, Persike, Mesenholl, & Hagemann, 2006; Meinhardt, Schmidt, Persike, & Roers, 2004; Nothdurft, 2000; Persike & Meinhardt, 2006; Rivest & Cavanagh, 1996; van Mierlo, Brenner, & Smeets, 2007), while others find independent processing (Leonards & Singer, 2000; Pashler, 1988; Phillips, 2001; Phillips & Craven, 2000; Treisman & Gelade, 1980; Tsujimoto & Tayama, 2004).

A deeper understanding of the underlying processes could be provided by electrophysiological measurements which allow to link physical stimulus properties and behavioral measurements with the timing of the neural response and are therefore ideal to examine the effect of cue combination on figure–ground segregation. It is known from a lot of studies using the electroencephalogram (EEG) that segregation of textures and figures causes a *similar* segregation-specific negative potential shift for a number of cues, suggesting that objects are identified by the visual system with a high temporal and spatial congruence across cue types. This shift occurs between 100 ms and 300 ms in the event-related potential (ERP) and was termed the texture-segregation visual evoked potential (tsVEP – Bach & Meigen, 1992, 1997; Bach, Schmitt, Quenzer, Meigen, & Fahle, 2000; Caputo & Casco, 1999; Fahle, Quenzer, Braun, & Spang, 2003). Similarly, a contour-specific negative response was found for contour integration paradigms (Mathes & Fahle, 2007; Mathes, Trenner, & Fahle, 2006), whereas later parts (maximally at 290 ms) were associated with closure processes (Doniger et al., 2000, 2001). Even when the contour was not physically present (i.e. it was illusory), an early negative modulation of the ERP was reported (Herrmann & Bosch, 2001; Murray, Imber, Javitt, & Foxe, 2006; Murray et al., 2002). To a first approximation, the ERP seems to be negatively modulated when figures or textures are segregated from the background. However, it has been shown in several of these studies that segregation-specific ERP modulations can be divided into early and late components (see also Heinrich, Andres, and Bach (2007)), and are strongly influenced by local–global surround conditions and task

* Corresponding author. Address: Department of Human Neurobiology, University of Bremen, Hochschulring 18, D-28359 Bremen, Germany. Fax: +49 421 218 62985.
E-mail address: sirko.straube@uni-bremen.de (S. Straube).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.12.013

(Casco, Campana, Han, & Guzzon, 2009; Han, Ding, & Song, 2002) suggesting that multiple cortical areas are involved.

An alternative way of looking at the electrophysiological response to a given stimulus is the investigation of specific frequency modulations in the EEG over time. Here, segregation-specific activity has been found in the γ-band (Eckhorn et al., 1988) and it has been shown that an early evoked activation in the γ-band is sensitive to stimulus properties in object detection and discrimination (Busch, Debener, Kranczioch, Engel, & Herrmann, 2004; Busch, Schadow, Frund, & Herrmann, 2006; Senkowski & Herrmann, 2002), while an induced activation at higher latencies seems to be involved in object representation (Tallon-Baudry & Bertrand, 1999). Both effects, the segregation-specific negativity of the ERP and the modulation of the γ-band are closely related to figure–ground segregation, which suggests that both may be sensitive to cue combination and/or the corresponding effect on the saliency of an object.

In the present study we specifically investigated the effect of cue combination and the related saliency changes on the human EEG (both, ERPs and specific frequency bands). Combining signal-detection theory and electrophysiology, we tested how the identification of a figure is altered by the underlying cue configuration, specifically by comparing single cues with cue combination, and how these changes influence the electrophysiological data. We used a novel approach by controlling the saliency of the target (and equating it for the single cue conditions) allowing us to specifically investigate the contribution of shape identification on ERP components. The paradigm is inspired by work from Meinhardt and colleagues who found an interaction in psychophysical tests when combining orientation and spatial frequency as visual cues (Meinhardt et al., 2004, 2006). They also reported differences in cue interaction between detection and identification tasks. Here, we concentrated on shape identification by using a figure discrimination task requiring identification of the target. Segregation of the figure from its background was not sufficient to solve this task but a classification was required. According to the results of Meinhardt et al., we expected a synergy effect for cue combination on the behavioral level. Our study focused on the earliest electrophysiological changes (ERP and/or time–frequency analysis) that can be attributed either to the physical properties of the stimulus and/or the percept. Due to the fact that orientation and spatial frequency interact as visual cues during figure–ground segregation, the expected changes should occur in the time range of the segregation-specific negativity of the ERP (i.e. between 100 ms and 300 ms). As outlined above, we also considered the γ-band (evoked and induced activity) as a candidate, possibly reflecting the cue configuration of the target or the perceived saliency of the figure.

2. Materials and methods

2.1. Subjects

Twelve undergraduate students (6 men, 6 women) aged between 22 and 27 (mean 23.7, sd 1.3) participated in this study. All participants had normal or corrected-to-normal vision as assessed by means of the Freiburg Visual Acuity Test (Bach, 1996) and reported no history of neurological or psychiatric disorders. Each subject was informed about the nature and the purpose of this study and gave written consent to participate. The study was conducted in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) and approved by the local ethics committee.

2.2. Task and stimulus

The stimulus consisted of a matrix of 33 × 25 Gabor patches on a gray background, as demonstrated in Fig. 1A, which were presented at a distance of 70 cm on a Samsung Syncmaster 1100 MB

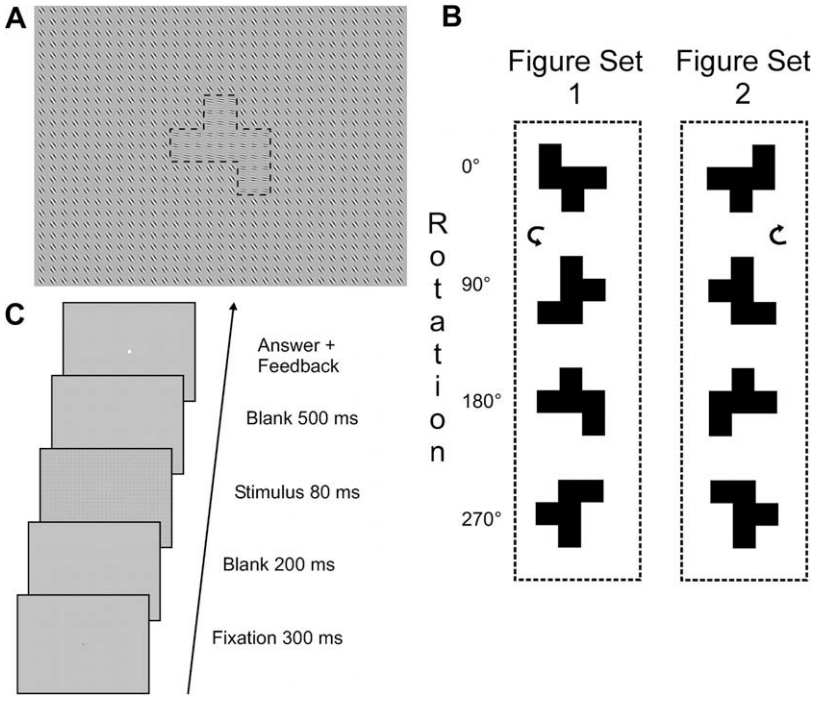

Fig. 1. Stimulus and task. (A) A matrix of Gabor elements containing a figure (shown here is the cue combination condition). The dashed line is shown for illustration purposes only and was not present in the original stimulus. All background elements had the same orientation and spatial frequency. Figure elements differed either in orientation, spatial frequency or both (cue combination). (B) Subjects had to discriminate between two mirror-symmetrical figures, which were presented randomly at one of four orientations. Figure of set 1 is rotated counterclockwise (indicated by arrow), while figure of set 2 is rotated clockwise, so that opposite pairs are always mirror images. (C) Sequence of one trial. The white square at the end of each trial indicated the answering period, which had no time limit. Auditory error feedback was given.

with a resolution of 1600 × 1200 pixels at a frame rate of 100 Hz. The Gabor patches had a center-to-center distance of 1° visual angle and a size defined by the width of the Gaussian envelope of σ = 10 arcmin. The target figure was part of this matrix with its elements differing from the background either in orientation, spatial frequency, or both (cue combination). Parameters of background elements were set to arbitrary values, exhibiting a spatial frequency of 3.5 cpd (cycles per degree) and an orientation of 36°.

The same task was used at all stages of the current study (except for the control experiment): Subjects had to discriminate between two mirror-symmetrical figures in a binary forced choice task (Fig. 1B), which were randomly rotated into one of four orientations. The figures were presented in a pseudo randomized order and subjects had to press one of two buttons to indicate whether they identified a figure of set 1 or 2. The correct answer was independent of the figure orientation presented and only considered the shape recognized. Moreover, the position of the figure was varied randomly, with a maximum center-to-center distance of 3° (center stimulus to center target). In summary, the exact position of the figure was unpredictable and identification of the main axis alone was not sufficient to give the appropriate answer, since both figures could take horizontal and vertical orientations and the shape was composed such that the whole figure equally spread in horizontal and vertical directions (see Fig. 1A and B). Subjects had to recognize the *whole* figure in order to solve the task and it was impossible to answer correctly from detecting only figure parts due to the asymmetrical form.

The procedure was as follows: A single trial started with a fixation period (300 ms – black fixation point) followed by a blank screen (200 ms). Subsequently, the stimulus (Gabor matrix as illustrated in Fig. 1A) appeared for 80 ms, again followed by a blank screen (500 ms). Finally, a small white square in the center of the blank screen indicated the answer period. There was no time limit for this period, so subjects were free to answer whenever they wanted to (Fig. 1C). By doing this, we allowed the subjects' to mentally rotate the recognized shape as long as necessary to find the correct answer. Auditory feedback was given for incorrect answers (2000 Hz tone for 100 ms). The background color of the monitor was gray during the whole sequence of one trial.

2.3. Procedure

2.3.1. Training

To ensure that all subjects were able to identify the figure properly, they were trained with a highly salient figure prior to the experiments. Training sessions were conducted for both single cue conditions (30 trials per session), with the figures' Gabor elements having an orientation of above 86° (i.e. 50° difference to the background) in the orientation condition and a spatial frequency above 5 cpd (i.e. 1.5 cpd difference to the background) in the spatial frequency condition, respectively. If the subject reached an accuracy of at least 90%-correct answers, the training session was finished, otherwise it was repeated. No subject needed more than one repetition.

2.3.2. Psychometric functions of single cues

The levels of difficulty in the main experiment were defined by the percent correct performance in the single cue conditions. Performance was estimated by measuring the psychometric functions for the identification of a figure defined purely by a difference in orientation or else spatial frequency. Any parameter–performance pair can be evaluated in this way due to the fact that the psychometric function is characterized by the location of a perceptual threshold and by its slope.

To achieve an accurate measure of the psychometric function, we sequentially used both an adaptive staircase procedure *and* the method of constant stimuli (MCS). First, we estimated the threshold – defined as the figure–ground difference corresponding to 75%-correct performance – and the slope of the psychometric function using the QUEST method (Watson & Pelli, 1983) with 50 trials. Second, we validated these results with the MCS by taking five values of the estimated function (parameters corresponding to a correct performance of 55%, 65%, 75%, 85% and 95%) presented pseudo randomized with 30 trials per configuration, which leads to 150 trials per MCS.

The psychometric functions thus measured were then used for the main experiment.

2.3.3. Main experiment

Three conditions were tested in the main experiment: identification of a figure defined by a figure–ground difference in (i) orientation (single cue), (ii) spatial frequency (single cue) or (iii) orientation *and* spatial frequency (cue combination). Each of these conditions was tested with three levels of difficulty, which were derived from the psychometric functions of the single cue conditions (see above). These levels were defined as stimuli corresponding to a correct performance of 55%, 76% and 98%, respectively. In the cue combination condition, figure–ground differences were defined by the superposition of the two single cue stimuli of the corresponding levels. For instance, on level 1 of the cue combination condition, the Gabor patches of the figure had a figure–ground difference corresponding to level 1 regarding *both* single cue conditions. Each configuration (condition × level) was repeated 100 times and presented in pseudo randomized order. Therefore, the main experiment consisted of three runs (one run = one condition, 300 trials per run), with the sequence counterbalanced between subjects. All participants were instructed only to blink during the answer phase to avoid blink artifacts in the time span of the ERP.

During the main experiment we simultaneously measured the EEG and performance, which allowed a direct comparison between psychophysics and electrophysiological data (see Section 3).

2.3.4. Control experiment – figure versus background

In a separate session, five of the subjects that took part in the main experiment performed a control experiment. The purpose of this experiment was to evaluate whether the figure in our stimulus by itself caused a segregation-specific negative shift in the ERP and whether the rarely observed P2 component (see Sections 3 and 4) is also observed in the pure background stimulus with a slightly modified task.

Stimulus, sequence and positioning of the figure were the same as in the previous experiments, but here the stimulus either contained a figure or else none. Subjects reported the presence of the figure in a yes–no detection paradigm. When a figure was present, it consisted of the single cue differences (either spatial frequency or orientation) corresponding to level 3. Trials without the figure contained only the background elements described above. The control experiment again consisted of 300 trials (100 trials per cue condition plus 100 trials without any figure).

2.3.5. Electrophysiological recording

The EEG was recorded in the main experiment from 25 recording sites (F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2, O9, O10) chosen from standard electrode positions (American Electroencephalographic Society, 1994) using Ag/AgCl sintered electrodes placed in an electrode cap (Easycap, Herrsching-Breitbrunn). For the control experiment, we reduced this configuration to 11 recording sites (Fz, C3, Cz, C4, Pz, PO3, POz, PO4, O1, Oz, O2). The average of both earlobe electrodes (A1 and A2) served as the reference and electrode impedance was kept below 10 kΩ. Eye movements, such as blinks, were monitored with a combined electrode pair above and lateral to the left eye. The EEG activity was amplified using a Nihon Kohden system (Neurofax EEG-1100).

During recording, a time constant of 0.3 s (cutoff frequency: 0.5 Hz) and a high-frequency cutoff of 120 Hz were used. The EEG was digitized at a sampling rate of 500 Hz.

2.4. Data analysis

2.4.1. Psychophysics

The relationship between figure–ground difference and perceptual saliency (measured in percent correct) is not linear due to the sigmoidal form of the psychometric function. Close to threshold, a small variation in physical figure–ground difference has a strong impact on the observers' performance. In contrast, this impact will be much weaker close to performance boundaries, i.e. close to both floor and ceiling of performance. Therefore, we rescaled the percent correct values into units of the sensitivity measure d' to linearize the measured object saliency according to the underlying sensory process. The relationship between d' and percent correct in a 2-alternative-forced choice task is given by

$$d' = \sqrt{2} * \Phi^{-1}(pc). \quad (1)$$

Here, Φ is the normal distribution function, and therefore $\Phi^{-1}(pc)$ gives the z-score of the percent correct value (Macmillan & Creelman, 1991, p. 124 ff).

In case of perfect performance (100%-correct) the value of d' becomes infinite. In order to compute a finite value of d' in this case, we set its value to a maximum of 4.0, which corresponds to a correct performance of 99.8% according to the relationship in Eq. (1).

To characterize the perceptual benefit of cue combination, we compared our results with an independent summation assumption. If orientation and spatial frequency are processed by independent neural pathways, the increase in performance is predicted by signal-detection theory according to

$$d'_{\perp} = \sqrt{\left(d'_f\right)^2 + \left(d'_{\phi}\right)^2}, \quad (2)$$

(Ashby & Townsend, 1986; Green & Swets, 1988, p. 271 ff; Macmillan & Creelman, 1991, p. 240 ff). The parameter d' denotes the distance between the means of the noise distribution and the particular signal + noise distribution. The two single-cue distributions (index f for spatial frequency cue, index ϕ for orientation cue) are orthogonal to each other in case of independent summation ($d'_{\perp}$) (Tanner, 1956). Therefore, the resulting perceptual object saliency is defined by the Euclidean distance between them, given by Eq. (2).

In the current study, we applied a discrimination task to measure the ease of shape recognition using d': The performance measured does neither reflect figure detection, because detecting or recognizing only parts of the figure was not sufficient to solve the task (see Section 2.2), nor does it reflect discrimination of mirror-symmetrical figures, because subjects were familiarized with the task (during training), so they could perform the discrimination with ease – once the shape has been recognized. The only manipulation that was used to influence the subjects' performance was a change of the underlying cue configuration (cue condition and level).

2.4.2. Event-related potentials

To investigate ERPs, 30 Hz (slope 24 dB/oct) low-pass filtered averages were used. We examined a time span from −100 ms to 500 ms relative to stimulus onset. Trials with blink artifacts, large eye movements, extensive muscle activity or other noise transients within this time span were rejected automatically on all recorded channels through an amplitude limitation of 70 μV (two subjects at 100 μV) and by visual inspection. Further control of eye movement was not necessary, since saccades to specifically search for the figure were not possible due to the short stimulus duration (80 ms). Only trials with correctly identified stimuli were included in the analysis. Remaining trials after artifact rejection for each level in mean and standard deviation (sd) were: 67 trials (sd 9 trials) of level 1, 83 trials (sd 8 trials) of level 2 and 92 trials (sd 2 trials) of level 3.

The mean signal of the 100 ms time window prior to stimulus onset served for baseline correction. The ERPs were sorted according to stimulus level and to the observers' answers. Filtering, artifact rejection and ERP generation were carried out using BESA 5.1.8 (MEGIS Software GmbH, Munich). Grand-average ERPs, amplitude and latency measurements of components were computed with in-house software using Matlab (Release 12.1, The MathWorks Inc., Massachusetts). Amplitudes were defined as distance to baseline.

2.4.3. Time–frequency analysis

A Morlet based wavelet transform with a width of six cycles was used for the inspection of power changes within defined frequency bands (4–80 Hz). The core routine was provided by Torrence and Compo (1998). Trials with artifacts identified in the ERP analysis were again not included in the time–frequency analysis. Since the wavelet transform at 4 Hz spans 1500 ms (with six cycles width), we computed a larger time span (between −1500 ms and 1250 ms relative to stimulus onset) to avoid border artifacts. A shorter time span was then used for the analysis (see Section 3). Only trials with correctly identified stimuli were included in the analysis. In contrast to the ERP analysis, the data were not filtered. We investigated normalized median power values of total activity (evoked and induced – for details see Herrmann, Grigutsch, and Busch (2005)). The procedure was as follows: For each subject separately, we computed the power values (in μV^2) in each trial for each frequency (f) and summarized all trials by taking the median power at each frequency and point in time. Moreover, we normalized the power value (P_f) at each time point t by the mean of the baseline power $\left(\overline{P}_f^0\right)$ according to

$$P_f^{norm}(t) = \frac{P_f(t) - \overline{P}_f^0}{\overline{P}_f^0} \quad (3)$$

Therefore the resulting normalized power value $\left(P_f^{norm}\right)$ at time t has no unit and represents the activity relative to baseline. As baseline, we used the time window 750 ms prior to stimulus onset. Since the normalization factor $\left(\overline{P}_f^0\right)$ is frequency-dependent, the normalization also accounts for the fact that high frequencies have less power in the EEG than low frequencies. Hence, the normalized power values represent the frequency-specific increase compared to baseline power.

Finally, the normalized data were averaged over all subjects. We obtained nine time–frequency plots for each electrode for the nine experimental configurations (condition × level).

2.4.4. Statistics

Statistical analysis was performed using SPSS 12.0 (SPSS Inc., Chicago). All results (performance, ERPs and time–frequency analysis) were validated by using repeated measurement ANOVAs. Wherever appropriate, p-values were adjusted by Greenhouse–Geisser corrections. Pairwise comparisons were conducted by using post-hoc t-tests. The correlations computed in this study report Pearson's correlation coefficient.

3. Results

3.1. Psychophysics

The measurement of the psychometric functions for both single cue conditions in the main experiment revealed similar thresholds (points of 75%-correct performance) for all subjects, which lay at 50.8° (sd 2.2°) for the orientation and at 4.7 cpd (sd 0.2 cpd) for the spatial frequency condition. Fig. 2 shows an increase in performance

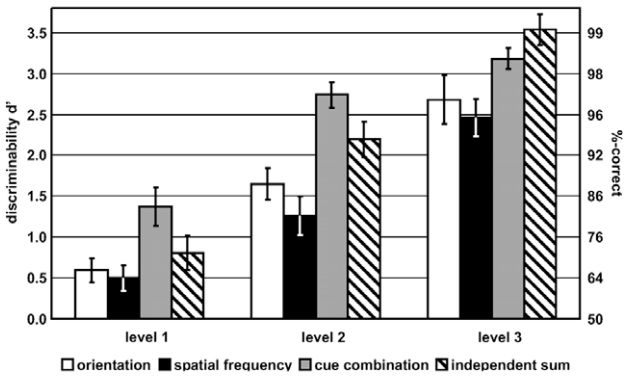

Fig. 2. Psychophysical results. Mean and standard errors of performance in the three experimental conditions (orientation, spatial frequency and cue combination) and performance predicted by the independent summation model (estimated from single cue performances of each subject). Shown are d'-values (left axis) and corresponding values of %-correct (right axis). Performance differs significantly between levels and conditions (levels: $p < 0.001$; conditions: $p < 0.001$) and is significantly higher for cue combination than predicted by independent summation ($p < 0.05$) for levels 1 and 2.

from levels 1 to 3 in all conditions. Furthermore, both single cue conditions had similar saliencies at all levels, with level 1 beneath threshold ($d' < 1.0$), level 2 slightly above threshold ($1.0 < d' < 2.0$) and level 3 way above threshold ($d' > 2.0$) close to the performance limit. A two factorial ANOVA for repeated measurements showed a significant main effect for both level ($F(2,22) = 70.0$, $p < 0.001$) and condition ($F(2,22) = 14.2$, $p < 0.001$), whereas we observed no significant interactions ($F(4,44) = 2.3$, $p = 0.08$). Pairwise comparisons of the three conditions on each level yielded no differences between the two single cue performances (two-tailed t-test – level 1: $p = 0.35$; level 2: $p = 0.19$; level 3: $p = 0.57$), while there was a significant improvement for the cue combination (one-tailed t-test – true for both single cues: level 1: $p < 0.01$; level 2: $p < 0.001$; level 3: $p < 0.05$). Hence, the combination of both cues increased object saliency compared to single cues on *all* levels. On levels 1 and 2, the perceptual improvement by cue combination was significantly higher ($p < 0.05$, one-tailed t-test) than what would be expected by the applied independent summation model (see Section 2) indicating the reported synergy effect (Meinhardt et al., 2004, 2006). At level 3 the improvement reached a ceiling due to the performance limit of 100%-correct.

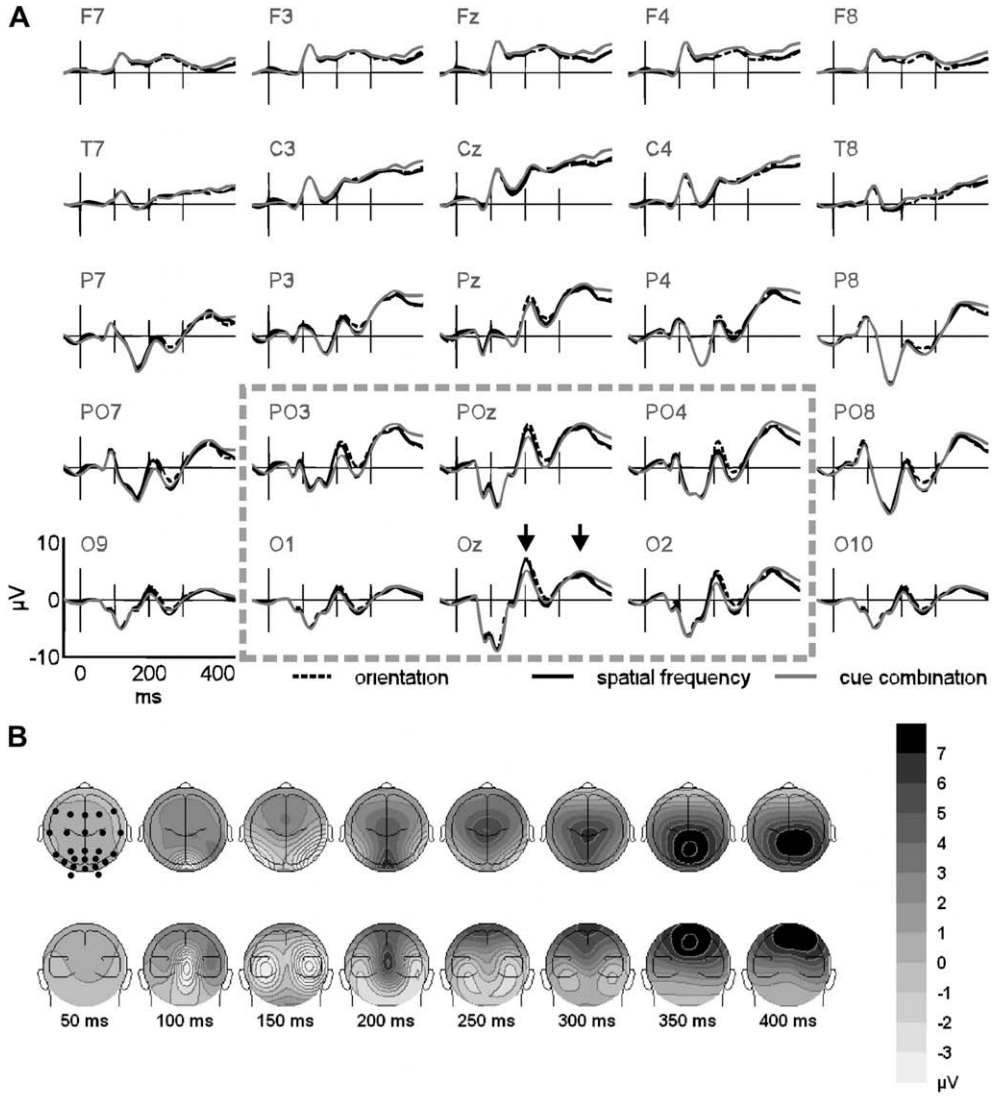

Fig. 3. Grand-average ERPs ($n = 12$) of level 2. (A) All electrodes for all three conditions. Arrows mark posterior P2 and P3 components. Dashed-dotted line (gray) marks electrodes used for P2 and P3 analysis. (B) Topographical timeline of the grand-average ERP for the orientation condition. Shown are top and back view of the head. The electrodes used in this study are depicted as black dots on the first top view (50 ms). The two electrodes that seem to lie outside the skull (O9 and O10) are actually situated left and right of the inion.

3.2. Event-related potentials

All three conditions elicited very similar ERPs. An example is illustrated near the perceptual threshold (level 2) in Fig. 3. Comparison of the single cue conditions with the cue combination showed two pronounced differences. The first was a negative shift around 200 ms which was most prominent over occipital and parieto-occipital electrodes, influencing mainly the posterior P2 component, while the second effect was an increase of the posterior P3 component. These effects were also validated statistically (see next subsections). No effects in amplitude or latency modulation were observed for posterior P1, N1 and N2 or anterior components.

A direct comparison of posterior ERPs is enabled in Fig. 4 (all conditions, levels and control experiment). In each cue condition, the most salient figure (level 3) elicited the smallest P2-amplitude (Fig. 4A–C). These ERP-characteristics were also observed in the control experiment (Fig. 4D), although here task and stimulus were slightly modified (see Sections 2 and 4). Background and figure-ERP mainly differed in a negative shift of the figure-ERP, which maximally influenced the posterior P2 component.

3.3. Amplitudes and latencies of the P2 component

The first component that was modulated by the stimulus manipulations applied in the present study was the posterior P2 which was defined as the positive peak in a time window between 180 ms and 250 ms after stimulus onset. To characterize the P2 component for each experimental configuration (cue condition and level), we measured both amplitude and latency at each electrode for each subject, separately, within this time window. We defined a region of interest (ROI – see Fig. 3) according to where the P2 was most prominent, and calculated the mean amplitude and latency of the associated electrodes for each subject. The means of all subjects are illustrated in Fig. 6A. Latencies did not differ across levels and conditions. Similarly, means and small standard errors indicate that the P2 occurred strictly time-locked to stimulus onset at about 208 ms (mean across configurations: 208.4 ms, sd 1.6 ms).

We investigated the topography of the negative amplitude shift of the P2 by testing the voltage maps at 208 ms with two three factorial ANOVAs, one for lateralization and one for anterior–posterior effects. To have approximately equidistant electrode positions we omitted PO-electrodes in this analysis. First, we investigated possible lateralization by pooling electrode sites according to lateral position (i.e. {F7, T7, P7, O9}; {F3, C3, P3, O1}; {Fz, Cz, Pz, Oz}; {F4, C4, P4, O2}; {F8, T8, P8, O10}). We found a main effect for electrode position ($F(2,18) = 24.3$, $p < 0.001$), but not for condition ($F(2,22) = 1.1$, $p = 0.35$) or level ($F(2,22) = 0.2$, $p = 0.79$). No interactions were observed (electrodesite × level: $F(3,30) = 2.8$, $p = 0.07$; electrodesite × condition: $F(3,34) = 1.5$, $p = 0.22$; condition × level: $F(4,44) = 1.2$, $p = 0.31$; electrodesite × level × condition: $F(4,39) = 2.2$, $p = 0.09$). Post-hoc t-tests revealed that the central electrode site differed from all other sites ($p < 0.001$), while the lateral electrodes did not differ between corresponding contralateral sites (mediolateral sites: $p = 0.29$; temporal sites: $p = 0.97$), but again from all other sites ($p < 0.01$). Hence, we observed no effect of lateralization for the peak of the P2 component. Second, posterior–anterior differences were investigated by pooling electrode sites according to anterior–posterior positions (i.e. {F7, F3, Fz, F4, F8}; {T7, C3, Cz, C4, T8}; {P7, P3, Pz, P4, P8}; {O9, O1, Oz, O2, O10}). We found no main effect of electrode site ($F(1,16) = 0.5$, $p = 0.57$), level ($F(2,22) = 0.2$, $p = 0.79$) or condition ($F(2,22) = 1.1$, $p = 0.35$), but observed an interaction of electrode site with level ($F(3,28) = 25.5$, $p < 0.001$) and condition ($F(2,22) = 4.5$, $p < 0.01$), while not with both ($F(3,33) = 0.7$, $p = 0.74$). Further inspection of the interactions revealed that the negative shift of the P2 compo-

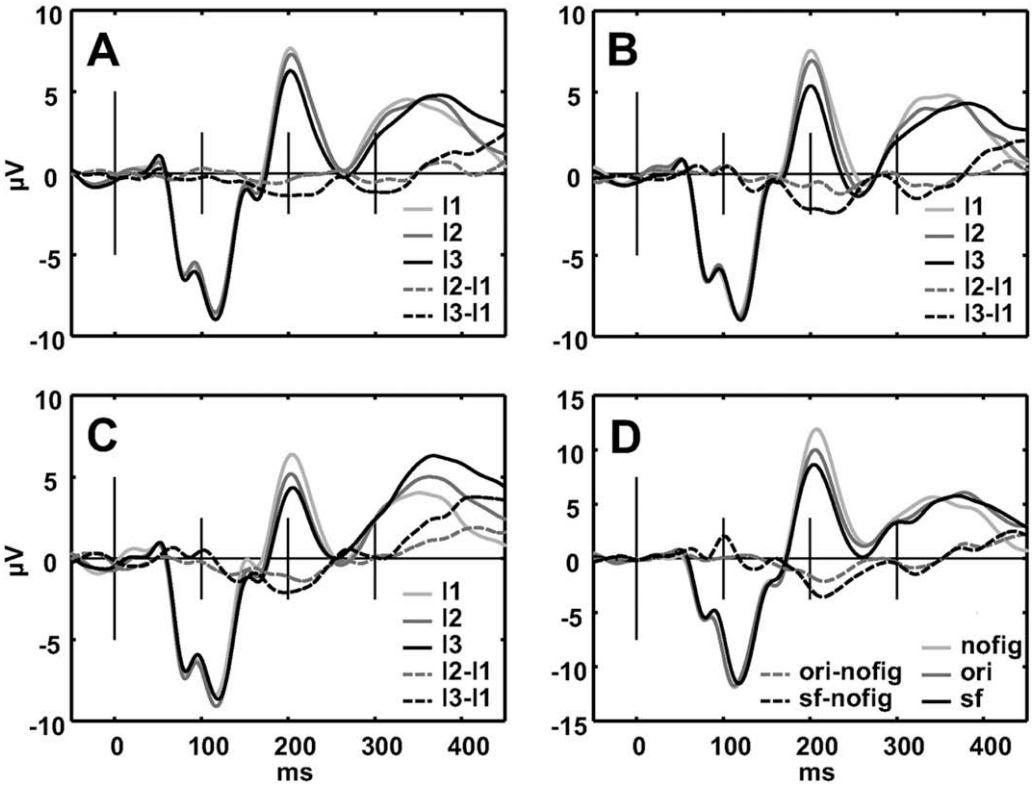

Fig. 4. ERPs for each experimental condition (solid lines) and corresponding differences (dashed lines) at electrode Oz. (A–C) Main experiment ($n = 12$): levels 1–3 for each cue configuration, i.e. (A) orientation cue, (B) spatial frequency cue and (C) cue combination. (D) Control experiment (figure versus background; $n = 5$): shown are ERPs for conditions with no figure (nofig) and with a figure defined by the single cues orientation (ori) or spatial frequency (sf). Note that the scale is changed on the y-axis.

nent is reversed into a positive shift at frontal electrode sides. Results of the topographical analysis are illustrated by the difference maps between level 3 and level 1 (see Fig. 5), indicating that the negative shift was not lateralized and was most pronounced at posterior electrode sites.

P2-amplitudes were now analyzed in the ROI depicted in Fig. 3. The amplitude of the P2 decreased in all conditions from levels 1 to 3. At each level, its smallest amplitude always occurred in the cue combination condition. A two factorial ANOVA for repeated measurements showed a significant main effect, both for level ($F(2,22) = 16.2$, $p < 0.001$) and condition ($F(2,22) = 3.7$, $p < 0.05$), whereas there were no significant interactions ($F(4,44) = 0.8$, $p = 0.56$). The P2-amplitudes did not differ between the two single cue conditions at any level as revealed by a two-tailed t-test (level 1: $p = 0.84$; level 2: $p = 0.38$; level 3: $p = 0.09$). Pairwise testing for significant negative amplitude shift of cue combination relative to both single cue conditions on each level (one-tailed t-test) yielded no differences for level 1 ($p = 0.11$), while there were significant differences on the other two levels: On level 2, both single cue P2-amplitudes differed from those for cue combination ($p < 0.05$). Whereas this also held for the comparison between the orientation and the cue combination condition on level 3 ($p < 0.05$), the difference between spatial frequency and cue combination conditions did not reach statistical significance ($p = 0.08$). This heterogeneity of results was due to a large diversity in P2-amplitudes across subjects (see amplitude standard errors in Fig. 6A). In the further analysis we dealt with this fact by using an amplitude normalization technique.

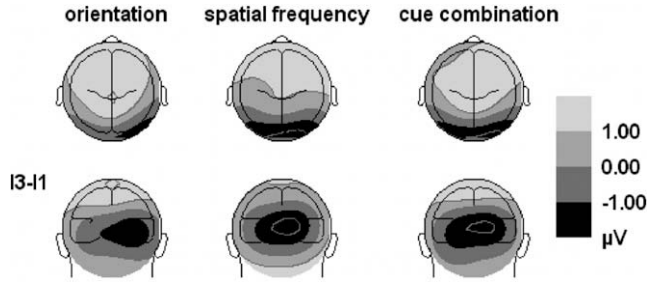

Fig. 5. Difference maps of topographies (level 3–level 1) of grand-average ERPs for all cue conditions at the peak of the posterior P2 component (208 ms).

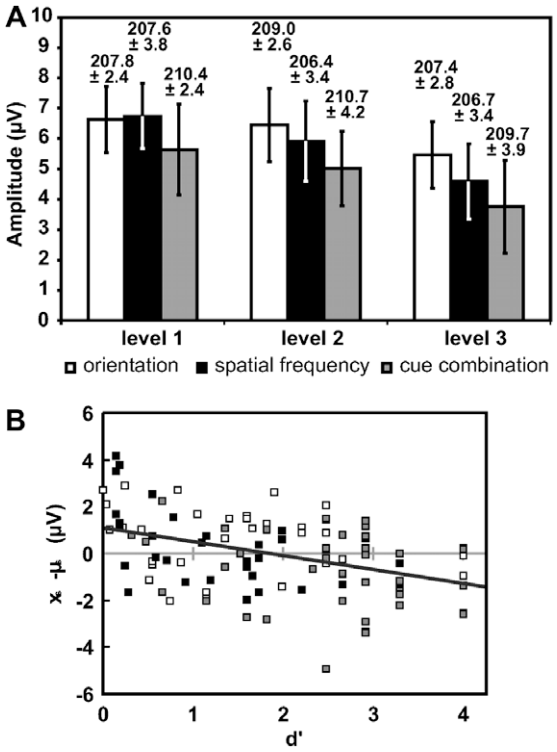

Fig. 6. Results of posterior P2-amplitude and latency analysis. The P2 component was defined as the positive peak between 180 ms and 250 ms. Each amplitude and latency measurement is the mean of marked electrodes in Fig. 3. (A) Mean amplitudes (all subjects) and standard errors for all levels and conditions. Amplitudes differ significantly between levels and conditions (levels: $p < 0.001$; conditions: $p < 0.05$). Corresponding latencies (and standard errors) are indicated above each bar (in ms). (B) Correlation between amplitude change of the posterior P2 component (y-axis) and figure saliency (x-axis), for all experimental conditions (single cue and cue combination). Each square denotes the P2-amplitude change and corresponding performance of a single subject in one experimental configuration (cue condition × level). For illustration purposes, each condition is color-coded according to the legend above. Amplitude change is measured as the difference of amplitude in the experimental configuration (x_s) to the individual mean amplitude (μ_s). The correlation coefficient is −0.42 ($p < 0.001$).

3.4. Relationship between P2 component and saliency

The measured performance reflects the perceived saliency of the figure, irrespective of the underlying cue configuration. The behavioral data (Fig. 2) show that all subjects benefited perceptually in the cue combination condition (as indicated by better performance). Hence, within a given level, the cue combination condition differed not only in its physical parameters (two cues), but also in its saliency. Is the observed effect at the posterior P2 component primarily related to the physical stimulus properties or else to perceived saliency?

To disclose the nature of the reduction in P2-amplitude, we re-examined the individual P2-amplitude changes in relation to individual performance in each particular experimental configuration (cue condition and level). The amplitude change was defined as the deviation from the individual mean (composed of all nine individual P2-amplitude measurements) and correlated with the individually measured d' value as a performance measure. Fig. 6B illustrates the significant correlation between amplitude reduction of the P2 component and increasing figure saliency (correlation coefficient −0.42, $p < 0.001$), revealing that a smaller P2-amplitude represents an increase in perceptual saliency.

3.5. Amplitudes and latencies of the P3 component

We also observed an effect of level and cue condition on the P3 component at parietal, parieto-occipital and occipital electrodes and evaluated these differences with the same method and ROI as was used in the analysis of the P2 component (see above and Fig. 3). The P3 component was defined as the positive peak in a time window between 300 ms and 500 ms. It had a mean amplitude of 7.0 μV (sd 3.3 μV) and a mean latency of 375.5 ms (sd 38.0 ms). A two factorial ANOVA for repeated measurements showed no main effect for condition (amplitude: $F(2,22) = 0.5$, $p = 0.62$; latency: $F(2,22) = 1.1$, $p = 0.36$), but a significant main effect for level regarding both amplitude ($F(2,22) = 4.0$, $p < 0.05$) and latency ($F(2,22) = 11.1$, $p < 0.001$). Additionally, the amplitude was modulated by an interaction between condition and level ($F(4,44) = 4.0$, $p < 0.01$), while the latency was not ($F(4,44) = 1.1$, $p = 0.36$). The P3 component varied both in amplitude *and* latency across experimental configurations. The results imply that this modulation is mainly based on saliency level, while the influence of the cue conditions is rather limited. To evaluate the relation between the amplitude of the P3 component and the saliency of the object, we correlated its amplitude changes with the results of the behavioral analysis using the same method as applied for the P2 component (see above and Fig. 6B). There was no significant corre-

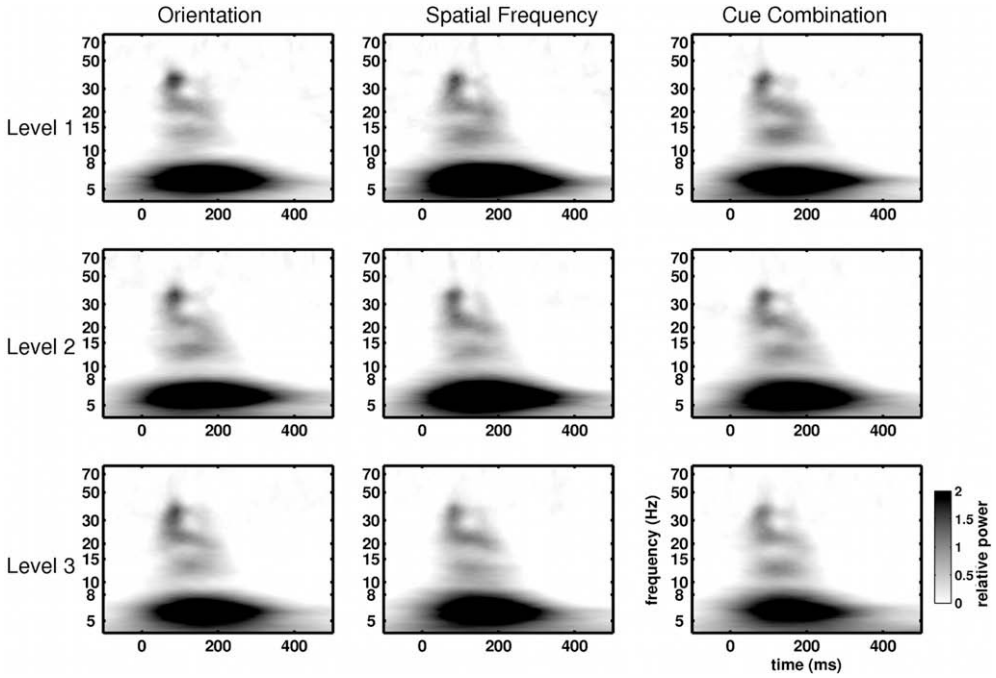

Fig. 7. Time–frequency results (average of all subjects) for all experimental configurations at the Oz-electrode. Normalized activity (see Section 2) shown in a time window between −100 ms and 500 ms relative to stimulus onset. There was no power suppression after stimulus onset, so only positive values occur.

Table 1
Summary of time–frequency analysis. The first column shows latency of peaks, computed from the normalized power of each subject. The second column shows results of a 2-factorial ANOVA with the normalized power values of these peaks. The last two columns show values of correlations of amplitude changes of these peaks with d' and amplitude change of the P2 component, respectively. For the θ-band, these data are illustrated in Fig. 8B and C. Not significant results are denoted by ns.

Frequency-band	Peak (ms) mean ± standard deviation	Main effect	Correlation with d'	Correlation with P2
θ (4–8 Hz)	188.2 ± 5.9	Condition ($p < 0.05$) + Level ($p < 0.01$)	−0.38 ($p < 0.001$)	0.34 ($p < 0.001$)
α (8–12 Hz)	127.6 ± 13.5	Condition ($p < 0.05$)	ns	ns
β (12–30 Hz)	140.3 ± 19.1	ns	ns	0.30 ($p < 0.01$)
Lower γ (30–50 Hz)	112.5 ± 23.3	ns	ns	0.37 ($p < 0.001$)

lation (see Section 4). In contrast, the peak of the P3 occurred significantly later with increasing saliency (correlation coefficient 0.45, $p < 0.001$).

3.6. Time–frequency analysis

Generally, we found a power increase in a broad frequency range up to 50 Hz after stimulus onset, most prominent at occipital and parieto-occipital electrodes, with maximum activity at Oz. Power changes at central and frontal electrodes were relatively small. We observed no significant power reduction compared to the pre-stimulus period. The averaged power of normalized single-subject data (see Section 2) for all experimental configurations (condition × level) demonstrates highly similar power progressions across cue conditions or performance levels at all frequencies (see Fig. 7).

The Oz-electrode, where the power increases were most pronounced, served as basis for our analysis. Here, we evaluated the peak of the power increase in each frequency band (for definitions see Table 1 and Herrmann et al. (2005)) after stimulus onset across the whole epoch in the normalized data of each subject. The standard deviation (sd) was computed across experimental configurations (for a summary of the results see Table 1). The strongest increase was observed in the θ-band (4–8 Hz), peaking at 188.2 ms (sd 5.9 ms). Power progressions for γ- and θ-band are illustrated in Fig. 8A.

To investigate whether the amplitude of these peaks is affected by the cue condition and/or saliency level, we applied a two factorial ANOVA for repeated measurements (see also Table 1). Only the peak of the θ-band was significantly modified by cue condition ($F(2,22) = 3.6$, $p < 0.05$) and level ($F(2,22) = 7.2$, $p < 0.01$). Post-hoc comparisons revealed a significant difference between level 3 and both level 2 ($p < 0.05$) and level 1 ($p < 0.001$), as well as a difference of spatial frequency and cue combination ($p < 0.05$) across levels. In the α-band, the cue condition significantly influenced the peak ($F(2,22) = 3.9$, $p < 0.05$): Pairwise comparisons indicated that single cues differed significantly ($p < 0.05$), with spatial frequency causing higher α-power.

3.7. Frequency modulations according to saliency

As with the analysis of the P2 component, we tried to relate the power changes to figure saliency. The individual change in power was correlated with d', i.e. the deviation of the normalized peak-power to the individual mean power in each frequency band. This method was also used for the correlation of the P2 component with the detection parameter d' (see above). In line with the results obtained from the two factorial ANOVA, we found a significant correlation only in the θ-band (correlation coefficient −0.38, $p < 0.001$), indicating that θ-activity decreased with increasing saliency (Fig. 8B). Since we observed similar effects for P2 and θ-band with a similar latency for both peaks, we tested the relation between the

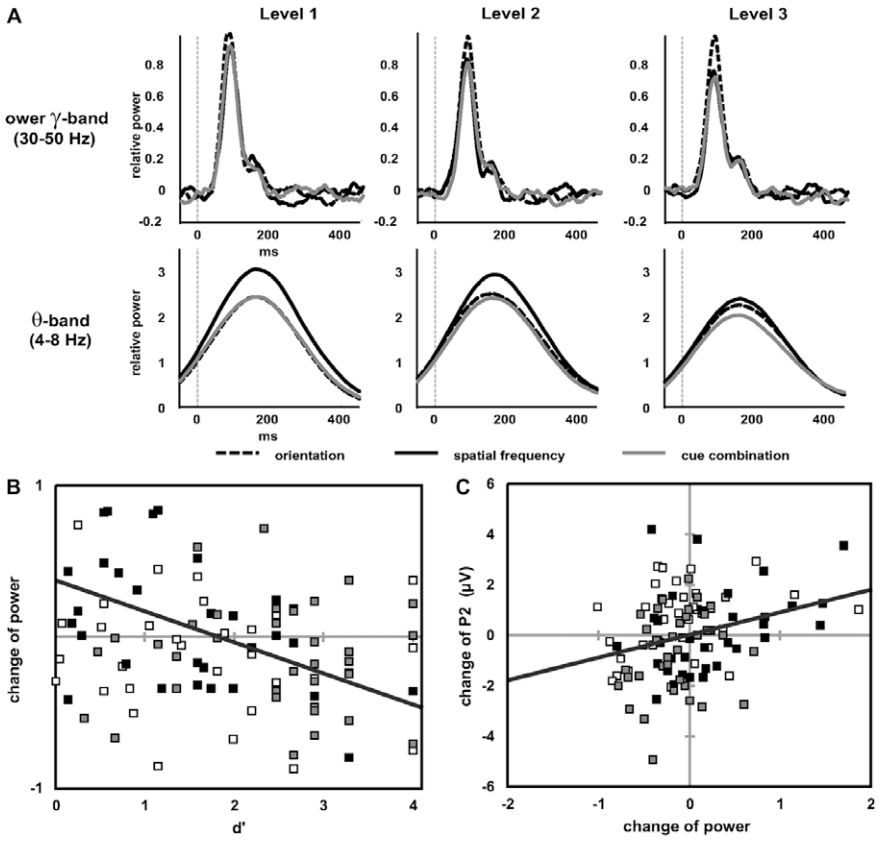

Fig. 8. Time–frequency analysis with averaged (all subjects) and normalized power at electrode Oz. (A) Power progression (normalized by pre-stimulus power) in the lower γ-band (upper panel) and θ-band (lower panel) for all levels and conditions. Stimulus onset is indicated by the dashed gray line. (B and C) Correlations to behavior and ERP. Condition is color-coded (as in Fig. 6) for orientation (white squares), spatial frequency (black squares) and cue combination (gray squares). (B) Changes of peak θ-power (mean 188.2 ms, sd 5.9 ms) were correlated with object saliency. Each square denotes the power change and corresponding performance of a single subject in one experimental configuration (cue condition × level). Power change is measured as the difference between the normalized power in the experimental configuration (x_s) and the individual mean of normalized power (μ_s). The correlation coefficient is −0.38 ($p < 0.001$). (C) Correlation of amplitude change of the P2 component (see Fig. 6B) and change of θ-power (see (B)). The correlation coefficient is 0.34 ($p < 0.001$).

two by correlating the change in θ-power with the change of the posterior P2-amplitude. This correlation was significant (correlation coefficient 0.34, $p < 0.001$), as becomes evident in Fig. 8C.

Surprisingly, we also observed such a significant correlation for the β- (correlation coefficient 0.30, $p < 0.01$) and lower γ-bands (correlation coefficient 0.37, $p < 0.001$), although the corresponding peaks occurred much earlier and we found no effect of cue condition or saliency onto these peaks (see Table 1).

4. Discussion

A definition of an object by cue combination improves its identification by the visual system. At least for spatial frequency and orientation this improvement is synergistic. Both behavioral and electrophysiological results imply that the visual system uses both cues for the recognition process. The first correlates of this process are observed as a negative amplitude shift, influencing mainly the peak amplitude of the posterior P2 component at about 200 ms. At this latency, we demonstrated that this shift is strongly correlated to the perceived saliency of the figure and therefore only indirectly related to the underlying physical cue configuration. The same effect can be demonstrated by a relative power decrease of the θ-band. In the following, we integrate the findings of our study in the present literature and further discuss the relationship between saliency and the underlying physical figure–ground differences, as well as the observed P2 effect. Finally, we briefly deal with the changes of the P3, whose characteristics slightly change near perceptual threshold.

4.1. Synergy through combination of cues

Our results confirm that the combination of spatial frequency and orientation improves the identification of a figure exceeding the predictions of an independent summation assumption (Meinhardt & Persike, 2003; Meinhardt et al., 2004, 2006; Persike & Meinhardt, 2006). This result was obtained, although subjects trained the single cue conditions (during QUEST and MCS measures) and were first confronted with the cue combination in the main experiment, indicating that the strong perceptual benefit of cue combination is a very robust and reliable effect.

The crucial question is at which processing stage both cues are combined. Our psychophysical results clearly show that shape recognition strongly benefits from cue combination, but is this effect caused by improved local feature contrast between figure and background elements or are less local processes like grouping and form completion directly affected by combination of the applied cues? Evidence for the latter comes from a study of Meinhardt and colleagues (2006), who compared the cue summation effect of orientation and spatial frequency for figure detection and identification. They show a higher benefit for figure identification and argue that this benefit is not exclusively a consequence of improved local feature contrast (which would be sufficient for detection), but involves further processes causing spatial form completion. In line with this view, we find the first ERP effect at an intermediate latency (208 ms) of known segregation-specific modulations in the ERP which are reported between 100 ms and

300 ms (Bach & Meigen, 1992, 1997; Bach et al., 2000; Caputo & Casco, 1999; Fahle et al., 2003). If the improvement of shape recognition by cue combination would rely on early segmentation processes, this effect should have been observable much earlier in time. Moreover, we show that this effect reflects the saliency of the figure indicating that shape recognition (which is based on *all* cues) has to be occurred at this point in time. The figure saliency most likely evolves from interactions between early and higher visual areas: Visual cues like orientation and spatial frequency are detected in the primary visual cortex (V1), while the processing of cue differences requires interactions between the detectors for figure and background regions. The earliest intermediate area, integrating information from V1, is area V2 which seems to be anatomically and functionally ideally suited for segregation processes (Shipp & Zeki, 2002a, 2002b). Computational models of texture segregation have demonstrated that the actual segregation process is accomplished by feedback from higher onto early visual areas (Bullier, 2001; Deco & Rolls, 2004; Itti & Koch, 2001; Roelfsema, Lamme, Spekreijse, & Bosch, 2002; Zwickel, Wachtler, & Eckhorn, 2007) and therefore occurs later in time. Saliency, which is a rather perceptual object property, is certainly related to these *re*activations of early visual areas by top-down control. Indeed, it has been shown that an early negative shift in the ERP during the processing of illusory contours was caused by the lateral-occipital complex (LOC) or at least modulated via feedback from the LOC (Murray, Foxe, Javitt, & Foxe, 2004; Murray et al., 2002), a region which is known to be involved in object recognition. Also the later negativity associated with closure has been attributed to LOC (Doniger et al., 2000, 2001; Sehatpour, Molholm, Javitt, & Foxe, 2006). Similarly, the ERP effect we find at about 200 ms reflects the saliency after form completion and cue combination.

4.2. Saliency as a non-linear function of physical figure–ground difference

The earliest electrophysiological difference between both performance levels and cue conditions is an amplitude change of the posterior P2-amplitude. Unfortunately, there are only a few studies that observe a posterior P2, since N1, N2 and P3 components are often overlapping (Luck, 2005, Chapter 1). The present study shows that the amplitude change of the P2 represents a change in perceptual saliency rather than directly representing the physical properties of the stimulus. Saliency is of course a product of figure–ground differences, but the influence of an increase in figure–ground difference (e.g. an increase in the orientation difference between figure and background elements) on saliency is variable, depending on where on the psychometric function this increase takes place (see also Section 2.4). Since saliency is a non-linear function of the figure–ground difference, the observed modulation of the P2 should also be correlated with changes in the physical figure–ground difference. This is the case, as shown in Fig. 9A and B, where we correlated the change of the P2 with the adjusted single cue difference, both in the single cue and in the cue combination conditions. The figure shows parallel regression lines for *both* comparisons of single cue and cue combination. If the amplitude of the P2 would be a direct indicator of the physical figure–ground difference, the regression line for the cue combination condition should be steeper than that for the single cue condition, since the value of the second cue also increases in the cue combination condition from left to right (i.e. in Fig. 9: values on the left are mainly from level 1 and values on the right are mainly from level 3). Hence, the difference between figure and ground increases more from left to right in the cue combination condition than in the single cue condition. The finding of parallel regression lines,

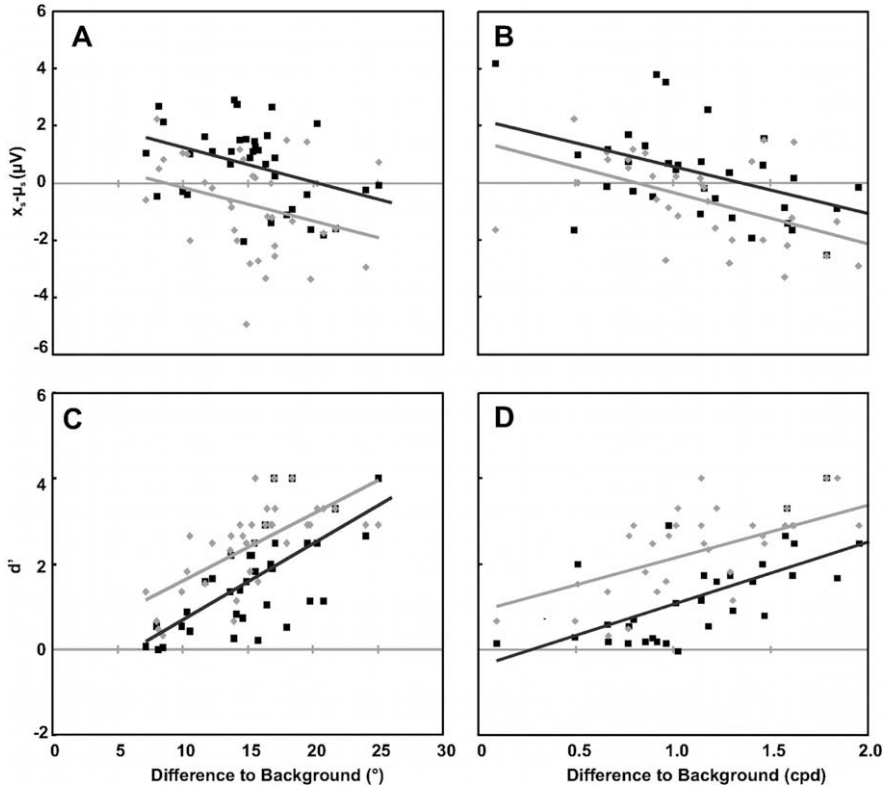

Fig. 9. Electrophysiological and behavioral changes with increasing figure–ground differences. (A and B) Correlation between amplitude change of the P2 component and figure–ground difference in the stimulus for single and combined cues. (A) Orientation cue in black (correlation coefficient −0.38; $p < 0.05$) and cue combination in gray (correlation coefficient −0.30; $p = 0.07$). (B) Spatial frequency cue in black (correlation coefficient −0.54; $p < 0.01$) and cue combination in gray (correlation coefficient −0.55; $p < 0.001$). (C and D) Correlation between saliency (d') and figure–ground difference in the stimulus for single and combined cues. (C) Orientation cue in black (correlation coefficient 0.68; $p < 0.001$) and cue combination in gray (correlation coefficient 0.69; $p < 0.001$). (D) Spatial frequency cue in black (correlation coefficient 0.70; $p < 0.001$) and cue combination in gray (correlation coefficient 0.65; $p < 0.001$).

however, does *not* support the notion that changes in the P2 directly reflect the figure–ground difference. The observed shift of the regression line is also observed in the correlation of *d'* with the physical figure–ground differences (Fig. 9C and D), supporting the conclusion that the shift in Fig. 9A and B is caused by a nearly constant benefit of saliency in the cue combination condition. Taken together, these findings clearly demonstrate that the amplitude of the P2 reflects the perceptual saliency as a non-linear function of the physical figure–ground difference.

4.3. The P2 component as a signature for saliency

Classically, the posterior P2 component has been related to object processing, influenced by spatial attention, feature selection and object memory (Anllo-Vento & Hillyard, 1996; Mecklinger & Muller, 1996; Tallon-Baudry, Bertrand, Peronnet, & Pernier, 1998). The timing (P2 latency: 208 ms) and direction (decreasing amplitude with increasing saliency) of the P2 differences we find are in line with previous studies reporting a segregation-specific negative amplitude shift between 100 ms and 300 ms (Bach & Meigen, 1992, 1997; Bach et al., 2000; Caputo & Casco, 1999; Fahle et al., 2003; Mathes & Fahle, 2007; Mathes et al., 2006). In line with the approach of these studies, when comparing ERP responses to background stimuli with those containing the figure (as we did in the control experiment – see Fig. 4D), it becomes likewise evident that the observed P2 effect is part of segregation-specific modulations in the ERP. The findings of the current study extend previous results by indicating that the negative shift is not only segregation-specific, but at some point in time directly correlated with perceptual saliency. In an overview article, Bach and Meigen (1998) reported a correlation between the tsVEP and saliency, when they increased saliency by changing the line length of a checkerboard stimulus. In their report, the normal VEP did not show a consistent modulation as a result of varying saliency, while the amplitude of the tsVEP increased. Here, we substantiate their conclusion by measuring individual saliencies and by systematically varying figure–ground difference and cue. The saliency-effect we find in the ERP is mostly unaffected by the number and type of cues. Contrary to Bach and Meigen, we observe a strong effect on the normal ERP (i.e. a modulation of the P2 component) which is blurred in the difference-ERPs (compare Fig. 4). Even when task and stimulus were slightly changed, as in the control experiment, we observed the occurrence of a P2 component which was diminished when the figure was present in the stimulus. The task in this control was much simpler, since only detection (yes/no) of the figure was necessary while recognition was not mandatory. Therefore, the task was easier to accomplish, so the task-related saliency was even higher than the adjusted level 3 of the identification experiment. This observation also suggests that the P2 effect found in this study is task independent and related to saliency, which is further supported by the fact that we found a very similar modulation of the P2 component in a figure detection task (Straube & Fahle, 2010).

Our findings support theories of a common saliency map, which is created by the combined responses of selectively modulated neurons (for a review see Treue (2003)). The P2 modulation we observed could well be a correlate of this saliency representation, since it is mainly affected by saliency irrespective of cue condition.

But why do we observe an amplitude reduction of the P2 when saliency *increases*? In the classical view, later negative modulations measured on the scalp are associated with excitatory postsynaptic potential activity, which means – for the results of the present study – that increasing saliency causes additional excitatory activity. This extra activity could be caused by the higher bottom-up signal (physical figure–ground difference) when the saliency increases. However, this interpretation is problematic when the latency of the effect is considered, because a modulation attributed to bottom-up processing should show up much earlier in the ERP. An alternative explanation is possible, when the observed amplitude decrease is interpreted as less neural activity. Then, the processes involved might be mediated by attention. Global attentional effects were controlled in our experiment by pseudo randomizing stimulus levels within conditions and order of blocks between subjects. However, selective attention might mediate enhanced population activity in order to improve the performance for weakly salient stimuli (Maunsell & Treue, 2006; Treue, 2003), whereas highly salient stimuli needed little or no attentional allocation in psychophysical tasks (see e.g. Nothdurft, 2000). Then the decrease of attentional allocation required to solve the task would be proportional to increasing figure saliency, which is exactly what we observe. Consistent with this view a recent study observed a similar effect on the posterior P2 component in a masking paradigm (Kotsoni, Csibra, Mareschal, & Johnson, 2007), the amplitude of the posterior P2 also decreasing with increasing d'. The authors interpret the P2 as a reactivation of primary and secondary visual areas by feedback from higher areas subserving appropriate representation of the stimulus, which perfectly fits with the interpretation that saliency evolves from interactions between early and higher visual areas (see Section 4.1). While not specifically quantifying the relationship to d', Kotsoni and colleagues suggest that an amplitude reduction represents a higher congruence between bottom-up and top-down signals and therefore less interference through feedback.

Another indirect support for our findings is given by a study, reporting a temporal blurring of the posterior P2 with increasing eccentricity (Shoji & Ozaki, 2006). Here, the saliency of the target (a circle) depended on the distractor type (squares, hexagons or octagons) and the P2-amplitude also declined with increasing saliency.

4.4. Relation of θ-power to ERP and saliency

We found the earliest increase of energy in the lower γ-band around 40 Hz, a well known phenomenon in object detection and discrimination tasks (Busch et al., 2004, 2006; Senkowski & Herrmann, 2002). Yet, this energy was stable across experimental conditions, so we conclude that the early onset γ-activity is not specific to the visual cues modified here and neither to object saliency. Later activations in the γ-band (e.g. induced γ-activity) were not observed in our paradigm. It has been suggested that induced γ-band activity is a candidate for active binding of visual features (Tallon-Baudry & Bertrand, 1999). However, this does not seem to be the case for simple visual cues such as orientation and spatial frequency.

At lower frequencies, we observed a modulation of the θ-band (4–8 Hz). Functionally, θ is related to short term memory and selective attention (Basar, Basar-Eroglu, Karakas, & Schurmann, 1999), whereas it is unlikely that it has a single functional role (Kahana, Seelig, & Madsen, 2001). In our study, θ-activity decreased with increasing saliency at a latency around 188 ms. The maximum of θ-activity is near the peak of the P2 component, which in turn has a strong θ-component, suggesting that both strongly depend on each other. Accordingly, a wavelet transformation of the ERP shows a strong θ-peak at a latency near the P2, indicating that most of the θ-activity is time-locked and therefore shows up in the ERP. However, the temporal resolution of the wavelet analysis is very poor at these low frequencies, and we could, nevertheless, clearly demonstrate that the effect of increased cue-independent saliency is related to a reduction in the θ-band.

Moreover, we found significant correlations of activity changes at other frequency bands with changes of the P2 even at *earlier* latencies (compare Table 1), although these were unrelated to saliency. These changes could probably be a hint of how the P2 component is generated, but this issue should be a subject of future studies.

4.5. The P3 component

The P3 represents a major endogenous component, which is moreover influenced by a number of experimental parameters (Fabiani, Gratton, & Coles, 2000; Key, Dove, & Maguire, 2005; Luck, 2005, chap. 1). In our study, the amplitude *and* latency of the P3 were influenced mainly by performance levels, which affect many endogenous parameters (in addition to exogenous ones), for example by making subjects less confident about their decision, rendering the task more difficult, requiring longer processing and increasing the proportion of guesses.

Classically, the amplitude of the P3 component has been related to working memory update (Polich, 2004; Polich & Kok, 1995), although such update processes seem not mandatory (Picton, 1992). The P3 amplitude decreases while its latency increases when the task becomes more difficult, indicating that the P3 is involved in stimulus classification and decision making processes (Picton, 1992; Polich & Kok, 1995). Although the ANOVA showed a significant effect of performance level on the P3 amplitude, the correlation of P3 amplitude with saliency was not significant, possibly because a large proportion of trials was near or below perceptual threshold ($d' = 1.0$) and might have been correctly guessed while not really recognized. In other words, below perceptual threshold the fraction of really recognized figures is small compared to the correctly guessed trials (50%-correct). Excluding all trials with $d' < 1.0$ and correlating the remaining trials with the P3 amplitude observed yielded indeed a significant correlation between amplitude and saliency (correlation coefficient 0.26, $p < 0.05$). Hence, the classical relationship of increasing P3 amplitude with decreasing task difficulty holds at least above perceptual threshold.

The latency of the P3 has been attributed to stimulus classification and processing (Polich, 2004; Polich & Kok, 1995), with longer latencies occurring in more difficult trials. Quite contrary, we observe longer latencies with easier trials. In contrast to classical studies investigating the P3, our target is hardly visible even on the easiest performance level. Visual noise distorts the P3 (McCarthy & Donchin, 1981), so the trend we observe here may reflect the emergence of the P3 out of noise with higher saliencies. Our results imply that the characteristics of the P3 near perceptual threshold have to be reviewed.

5. Conclusions

Orientation and spatial frequency interact as visual cues during the identification of a figure. This interaction, based on cue combination, is neither specifically reflected in the ERP nor in the power distribution of frequencies up to 80 Hz. Instead, the crucial feature of the neural response is a negative shift in the ERP, which occurs on an intermediate stage (about 200 ms after stimulus onset) correlated to saliency. This shift is measurable as an amplitude modulation of the posterior P2 component as well as a power reduction in the θ-band (4–8 Hz). A posterior P2 component is rarely described in the literature which may be based on the fact that other studies used highly salient stimuli. With the present work, we explored the direct relationship between the P2-amplitude and object saliency, probably reflecting the existence of a cue-independent saliency map and/or reflecting the fact that the more salient an object is, the less computation is required to detect or identify it. This saliency-effect on the ERP is robust across different cues and number of cues, maybe even across tasks, and should be considered in future studies as an important factor affecting the results. Furthermore, our study provides the basis for the development of an electrophysiological method to evaluate the strength of perceptual impressions in humans, either by utilizing the negative shift of the ERP (in particular the P2 component) or the power decrease of the θ-band.

Acknowledgments

The authors like to thank B. Mathes, D. Trenner, U. Ernst and G. Meinhardt for valuable suggestions. Supported by Grant 01GQ0705 (Bernstein program) of the German Federal Ministry of Education and Research (BMBF).

References

Abele, M., & Fahle, M. (1995). Interactions between orientation, luminance and color cues in figure–ground discrimination. *Perception, 24*, 11.

American Electroencephalographic Society (1994). Guideline thirteen: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology, 11*, 111–113.

Anllo-Vento, L., & Hillyard, S. A. (1996). Selective attention to the color and direction of moving stimuli: Electrophysiological correlates of hierarchical feature selection. *Perception and Psychophysics, 58*, 191–206.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93*, 154–179.

Bach, M. (1996). The Freiburg visual acuity test – Automatic measurement of visual acuity. *Optometry and Vision Science, 73*, 49–53.

Bach, M., & Meigen, T. (1992). Electrophysiological correlates of texture segregation in the human visual evoked potential. *Vision Research, 32*, 417–424.

Bach, M., & Meigen, T. (1997). Similar electrophysiological correlates of texture segregation induced by luminance, orientation, motion and stereo. *Vision Research, 37*, 1409–1414.

Bach, M., & Meigen, T. (1998). Electrophysiological correlates of human texture segregation, an overview. *Documenta Ophthalmologica, 95*, 335–347.

Bach, M., Schmitt, C., Quenzer, T., Meigen, T., & Fahle, M. (2000). Summation of texture segregation across orientation and spatial frequency: Electrophysiological and psychophysical findings. *Vision Research, 40*, 3559–3566.

Basar, E., Basar-Eroglu, C., Karakas, S., & Schurmann, M. (1999). Are cognitive processes manifested in event-related gamma, alpha, theta and delta oscillations in the EEG? *Neuroscience Letters, 259*, 165–168.

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews, 36*, 96–107.

Busch, N. A., Debener, S., Kranczioch, C., Engel, A. K., & Herrmann, C. S. (2004). Size matters: Effects of stimulus size, duration and eccentricity on the visual gamma-band response. *Clinical Neurophysiology, 115*, 1810–1820.

Busch, N. A., Schadow, J., Frund, I., & Herrmann, C. S. (2006). Time–frequency analysis of target detection reveals an early interface between bottom-up and top-down processes in the gamma-band. *NeuroImage, 29*, 1106–1116.

Caputo, G., & Casco, C. (1999). A visual evoked potential correlate of global figure–ground segmentation. *Vision Research, 39*, 1597–1610.

Casco, C., Campana, G., Han, S., & Guzzon, D. (2009). Psychophysical and electrophysiological evidence of independent facilitation by collinearity and similarity in texture grouping and segmentation. *Vision Research, 49*, 583–593.

Deco, G., & Rolls, E. T. (2004). A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research, 44*, 621–642.

Doniger, G. M., Foxe, J. J., Murray, M. M., Higgins, B. A., Snodgrass, J. G., Schroeder, C. E., et al. (2000). Activation timecourse of ventral visual stream object-recognition areas: High density electrical mapping of perceptual closure processes. *Journal of Cognitive Neuroscience, 12*, 615–621.

Doniger, G. M., Foxe, J. J., Schroeder, C. E., Murray, M. M., Higgins, B. A., & Javitt, D. C. (2001). Visual perceptual learning in human object recognition areas: A repetition priming study using high-density electrical mapping. *NeuroImage, 13*, 305–313.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., et al. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biological Cybernetics, 60*, 121–130.

Fabiani, M., Gratton, G., & Coles, M. G. H. (2000). Event-related brain potentials: Methods, theory, and applications. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 53–84). Cambridge: Cambridge University Press.

Fahle, M., Quenzer, T., Braun, C., & Spang, K. (2003). Feature-specific electrophysiological correlates of texture segregation. *Vision Research, 43*, 7–19.

Green, D. M., & Swets, J. A. (1988). *Signal detection theory and psychophysics* (rev., repr. ed.). Los Altos, CA: Peninsula Publ..

Han, S., Ding, Y., & Song, Y. (2002). Neural mechanisms of perceptual grouping in humans as revealed by high density event related potentials. *Neuroscience Letters, 319*, 29–32.

Heinrich, S. P., Andres, M., & Bach, M. (2007). Attention and visual texture segregation. *Journal of Vision, 7*, 6.

Herrmann, C. S., & Bosch, V. (2001). Gestalt perception modulates early visual processing. *Neuroreport, 12*, 901–904.

Herrmann, C. S., Grigutsch, M., & Busch, N. A. (2005). EEG oscillations and wavelet analysis. In T. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 229–259). MIT Press.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*, 194–203.

Kahana, M. J., Seelig, D., & Madsen, J. R. (2001). Theta returns. *Current Opinion in Neurobiology, 11*, 739–744.
Key, A. P., Dove, G. O., & Maguire, M. J. (2005). Linking brainwaves to the brain: An ERP primer. *Developmental Neuropsychology, 27*, 183–215.
Kotsoni, E., Csibra, G., Mareschal, D., & Johnson, M. H. (2007). Electrophysiological correlates of common-onset visual masking. *Neuropsychologia, 45*, 2285–2293.
Kubovy, M., & Cohen, D. J. (2001). What boundaries tell us about binding. *Trends in Cognitive Sciences, 5*, 93–95.
Kubovy, M., Cohen, D. J., & Hollier, J. (1999). Feature integration that routinely occurs without focal attention. *Psychonomic Bulletin and Review, 6*, 183–203.
Leonards, U., & Singer, W. (2000). Conjunctions of colour, luminance and orientation: The role of colour and luminance contrast on saliency and proximity grouping in texture segregation. *Spatial Vision, 13*, 87–105.
Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA, USA: MIT Press.
Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge, USA: Cambridge University Press.
Mathes, B., & Fahle, M. (2007). The electrophysiological correlate of contour integration is similar for color and luminance mechanisms. *Psychophysiology, 44*, 305–322.
Mathes, B., Trenner, D., & Fahle, M. (2006). The electrophysiological correlate of contour integration is modulated by task demands. *Brain Research, 1114*, 98–112.
Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences, 29*, 317–322.
McCarthy, G., & Donchin, E. (1981). A metric for thought: A comparison of P300 latency and reaction time. *Science, 211*, 77–80.
Mecklinger, A., & Muller, N. (1996). Dissociations in the processing of "what" and "where" information in working memory: An event-related potential analysis. *Journal of Cognitive Neuroscience, 8*, 453–473.
Meinhardt, G., & Persike, M. (2003). Strength of feature contrast mediates interaction among feature domains. *Spatial Vision, 16*, 459–478.
Meinhardt, G., Persike, M., Mesenholl, B., & Hagemann, C. (2006). Cue combination in a combined feature contrast detection and figure identification task. *Vision Research, 46*, 3977–3993.
Meinhardt, G., Schmidt, M., Persike, M., & Roers, B. (2004). Feature synergy depends on feature contrast and objecthood. *Vision Research, 44*, 1843–1850.
Murray, M. M., Foxe, D. M., Javitt, D. C., & Foxe, J. J. (2004). Setting boundaries: Brain dynamics of modal and amodal illusory shape completion in humans. *The Journal of Neuroscience, 24*, 6898–6903.
Murray, M. M., Imber, M. L., Javitt, D. C., & Foxe, J. J. (2006). Boundary completion is automatic and dissociable from shape discrimination. *The Journal of Neuroscience, 26*, 12043–12054.
Murray, M. M., Wylie, G. R., Higgins, B. A., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). The spatiotemporal dynamics of illusory contour processing: Combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *The Journal of Neuroscience, 22*, 5055–5073.
Nothdurft, H. (2000). Salience from feature contrast: Additivity across dimensions. *Vision Research, 40*, 1183–1201.
Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception and Psychophysics, 43*, 307–318.
Persike, M., & Meinhardt, G. (2006). Synergy of features enables detection of texture defined figures. *Spatial Vision, 19*, 77–102.
Phillips, W. A. (2001). Contextual modulation and dynamic grouping in perception. *Trends in Cognitive Sciences, 5*, 95–97.
Phillips, W. A., & Craven, B. J. (2000). Interactions between coincident and orthogonal cues to texture boundaries. *Perception and Psychophysics, 62*, 1019–1038.
Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology, 9*, 456–479.
Polich, J. (2004). Clinical application of the P300 event-related brain potential. *Physical Medicine and Rehabilitation Clinics of North America, 15*, 133–161.
Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: An integrative review. *Biological Psychology, 41*, 103–146.
Rivest, J., & Cavanagh, P. (1996). Localizing contours defined by more than one attribute. *Vision Research, 36*, 53–66.
Roelfsema, P. R., Lamme, V. A., Spekreijse, H., & Bosch, H. (2002). Figure–ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience, 14*, 525–537.
Sehatpour, P., Molholm, S., Javitt, D. C., & Foxe, J. J. (2006). Spatiotemporal dynamics of human object recognition processing: An integrated high-density electrical mapping and functional imaging study of "closure" processes. *NeuroImage, 29*, 605–618.
Senkowski, D., & Herrmann, C. S. (2002). Effects of task difficulty on evoked gamma activity and ERPs in a visual discrimination task. *Clinical Neurophysiology, 113*, 1742–1753.
Shipp, S., & Zeki, S. (2002a). The functional organization of area V2, I: Specialization across stripes and layers. *Visual Neuroscience, 19*, 187–210.
Shipp, S., & Zeki, S. (2002b). The functional organization of area V2, II: The impact of stripes on visual topography. *Visual Neuroscience, 19*, 211–231.
Shoji, H., & Ozaki, H. (2006). Topographic change in ERP due to discrimination of geometric figures in the peripheral visual field. *International Journal of Psychophysiology, 62*, 115–121.
Straube, S., & Fahle, M. (2010). The electrophysiological correlate of saliency: Evidence from a figure-detection task. *Brain Research, 1307*, 89–102.
Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences, 3*, 151–162.
Tallon-Baudry, C., Bertrand, O., Peronnet, F., & Pernier, J. (1998). Induced gamma-band activity during the delay of a visual short-term memory task in humans. *The Journal of Neuroscience, 18*, 4244–4254.
Tanner, W. P. (1956). Theory of recognition. *Journal of the Acoustical Society of America, 28*, 882–888.
Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society, 79*, 61–78.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.
Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology, 13*, 428–432.
Tsujimoto, S., & Tayama, T. (2004). Independent mechanisms for dividing attention between the motion and the color of dynamic random dot patterns. *Psychological Research, 68*, 237–244.
van Mierlo, C. M., Brenner, E., & Smeets, J. B. (2007). Temporal aspects of cue combination. *Journal of Vision, 7*, 8.1–11.
Watson, A. B., & Pelli, D. G. (1983). QUEST: A bayesian adaptive psychometric method. *Perception and Psychophysics, 33*, 113–120.
Zwickel, T., Wachtler, T., & Eckhorn, R. (2007). Coding the presence of visual objects in a recurrent neural network of visual cortex. *Bio Systems, 89*, 216–226.