



Similar Electrophysiological Correlates of Texture Segregation Induced by Luminance, Orientation, Motion and Stereo

MICHAEL BACH,*‡ THOMAS MEIGEN†

Received 28 June 1996; in revised form 26 November 1996

Certain local features induce preattentive texture segregation. Recently, components in the visual evoked potential (VEP) associated with preattentive texture segregation (tsVEPs) have been demonstrated. To assess the similarity and dissimilarity of visual processing across visual dimensions, we compared VEPs and tsVEPs in texture segregation by luminance, orientation, motion and stereo disparity. We found tsVEPs across these four visual dimensions to be remarkably similar when compared to the “low-level” VEPs. The tsVEPs were always negative; their implicit time, peak latency and amplitude were (in msec/msec/ μ V): 91/234/–5.7, luminance; 84/257/–3.9, orientation; 80/295/–8.3, motion; and 95/310/–5.0 for stereo. The cross-correlation function, as a quantitative measure for similarity, on average was higher for the tsVEPs by a factor of 4.2 as compared to the low-level VEPs ($P < 0.0001$). The results suggest (1) that the tsVEPs represent activity of neural mechanisms that have generalised to some degree across visual dimensions; and (2) that these hypothetical generalisation mechanisms might exist already in the primary visual cortex. © 1997 Elsevier Science Ltd.

Texture segregation VEP Orientation contrast Motion Stereo Human

INTRODUCTION

Our visual system continuously analyses the visual surround using parallel cortical processing. As an early step to segregate a figure from its background, neural mechanisms process the visual input across the entire visual field without focal attention. These mechanisms rely on certain local features (“visual dimensions”), which include luminance, orientation, motion and stereo disparity. If there is a sufficiently strong spatial gradient in one of these visual dimensions, global structures “pop out”, “group”, or “segregate” preattentively (Neisser, 1967; Beck, 1972, 1983; Julesz & Bergen, 1983; Treisman, 1985; Julesz, 1986; Nothdurft, 1993). To assess how texture segregation differs between visual dimensions, we compared texture segregation-specific components in the human visual evoked potential (“tsVEPs”) across four visual dimensions.

Evoked potentials provide a tool to study neuronal processing in humans. A VEP response to a stimulus with global segregation will contain both “low-level” VEP-

components and tsVEP-components. By “low-level” we denote those components in the VEP that are associated with mechanisms specific for each visual dimension and that are evoked even if no gradients are present. The tsVEPs and low-level VEPs can be isolated by appropriate linear combinations of the composite VEPs to different stimuli (Fig. 2; Bach & Meigen, 1990, 1992; Lamme *et al.*, 1992, 1993a,b; Meigen & Bach, 1993).

While this had not been formally tested, it seemed likely that low-level VEPs differ between visual dimensions (e.g. Regan, 1989). If the tsVEPs were tied closely to the visual dimension, they should differ likewise between dimensions. Alternatively, we hypothesised that the tsVEPs might resemble each other more closely if they were associated with common processing of spatial gradients irrespective of the visual dimension. To test these alternatives, we compared low-level VEPs and tsVEPs for four visual dimensions: orientation, motion, luminance and stereo disparity. The stimulus conditions for these visual dimensions were additionally tested psychophysically to ensure that the spatial gradients were strong enough to induce a preattentive pop-out.

METHODS

Subjects

Seven visually normal observers served as subjects in the electrophysiological experiments, six of these also in

*Elektrophysiologisches Labor, Universitäts-Augenklinik Freiburg, Killianstr. 5, D-79106 Freiburg, Germany.

†Universitäts-Augenklinik Würzburg, Josef-Schneider Str. 11, D-97080, Würzburg, Germany.

‡To whom all correspondence should be addressed [Fax: ++49-761-270-4052; Email: bach@uni-freiburg.de].

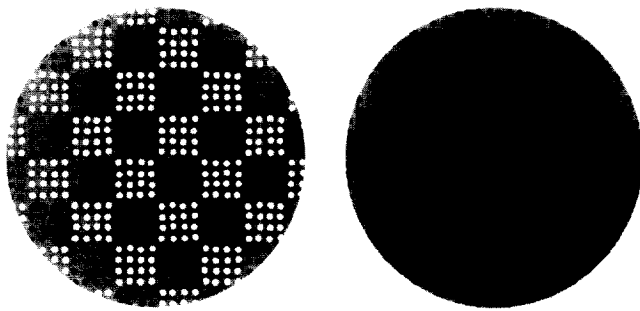


FIGURE 1. Examples of actual stimuli for the visual dimension luminance (top) and orientation (bottom). The stimuli for "motion" and "stereo" cannot convincingly be presented on paper. In the psychophysical experiments, the position of a single check had to be detected.

the psychophysical experiments. They wore appropriate refraction if necessary, acuity was ≥ 1.2 . The subjects gave their informed consent to participate in the experiment. Four of the seven subjects were naïve as to the specific aim of the experiment.

Stimuli

The basic paradigm was the one used by Bach & Meigen (1990, 1992); Lamme *et al.* (1992, 1993a,b, 1994) and Meigen & Bach (1993). The stimuli (Fig. 1) were presented on a visual display unit (HCM38, AEG) with a resolution of 832×832 pixels at a frame rate of 68.4 Hz.

For each visual dimension we introduced spatial gradients by choosing two different "variants" (Table 1, examples in Fig. 1, left). Different visual stimuli were defined by either arranging both variants in a global structure (checkerboard for electrophysiology or single check for psychophysics) or by filling the stimulus field homogeneously with one variant only.

For the electrophysiological experiments, "homogeneous" fields of one variant only (Fig. 1, right) were presented in addition to the arrangement in checkerboard patterns (Fig. 1, left). Changing between such patterns conceptually has effects at two levels, namely at the level of the local variant (e.g., horizontal/vertical lines) and at the level of the global structure (checkerboard onset/offset). The appropriate combinations of responses to different stimuli to either extract the tsVEP or the low-level VEP are detailed in data analysis below.

For the visual dimension "luminance", we used an arrangement of bright and dark disks so that the spatial arrangement resembles that in the other dimensions (Table 1). For the visual dimension "orientation", we used oriented line segments with a 90 deg orientation gradient. To avoid luminance artifacts, the lines were drawn at ± 45 deg relative to the screen, and the screen was tilted by 45 deg (Bach & Meigen, 1992). For the visual dimension "motion", we used bright disks on a dark background. Horizontal oscillatory motion of half of the disks, in a checkerboard arrangement, evoked a striking preattentive checkerboard. For the visual dimension "stereo disparity", we again used bright disks on a dark background. Disparity was introduced by using LCD-shutter glasses (SEGA) synchronised to the screen. For control, we tilted the screen by 90 deg, producing vertical disparities. In this condition, the subjects perceived no checkerboard pattern and no significant tsVEPs were evoked. This implies that the effect of monocular cues that might lead to a non-stereo texture segregation was negligible in the stereo condition.

In all cases, the local elements had a maximum extent of 0.2 deg and were spatially jittered around their lattice

TABLE 1. Stimulus details

Visual dimension	Luminance	Orientation	Motion	Stereo disparity
Variant 1	Bright disks	Vertically oriented line segments	Spatially oscillating disks	Disks with crossed disparity
Variant 2	Dark disks	Horizontally oriented line segments	Stationary disks	Disks with uncrossed disparity
Parameter details	Disk luminance 2.4/57 cd/m ² Background 30 cd/m ² Diameter 0.2 deg	Line luminance 57 cd/m ² Background 2.4 cd/m ² Length 0.2 deg	Line luminance 57 cd/m ² Background 2.4 cd/m ² Diameter 0.2 deg Oscillation frequency 8.55 Hz, amplitude ± 0.07 deg (velocity ± 1.2 deg/sec)	Line luminance 57 cd/m ² Background 2.4 cd/m ² Diameter 0.2 deg Disparity 0.04 deg
Spatial arrangement	Inter-element distance 0.4 deg Random jitter between ± 0.05 deg	Inter-element distance 0.4 deg Random jitter between ± 0.05 deg	Inter-element distance 0.4 deg Random jitter between ± 0.05 deg	Inter-element distance 0.4 deg Random jitter between ± 0.05 deg

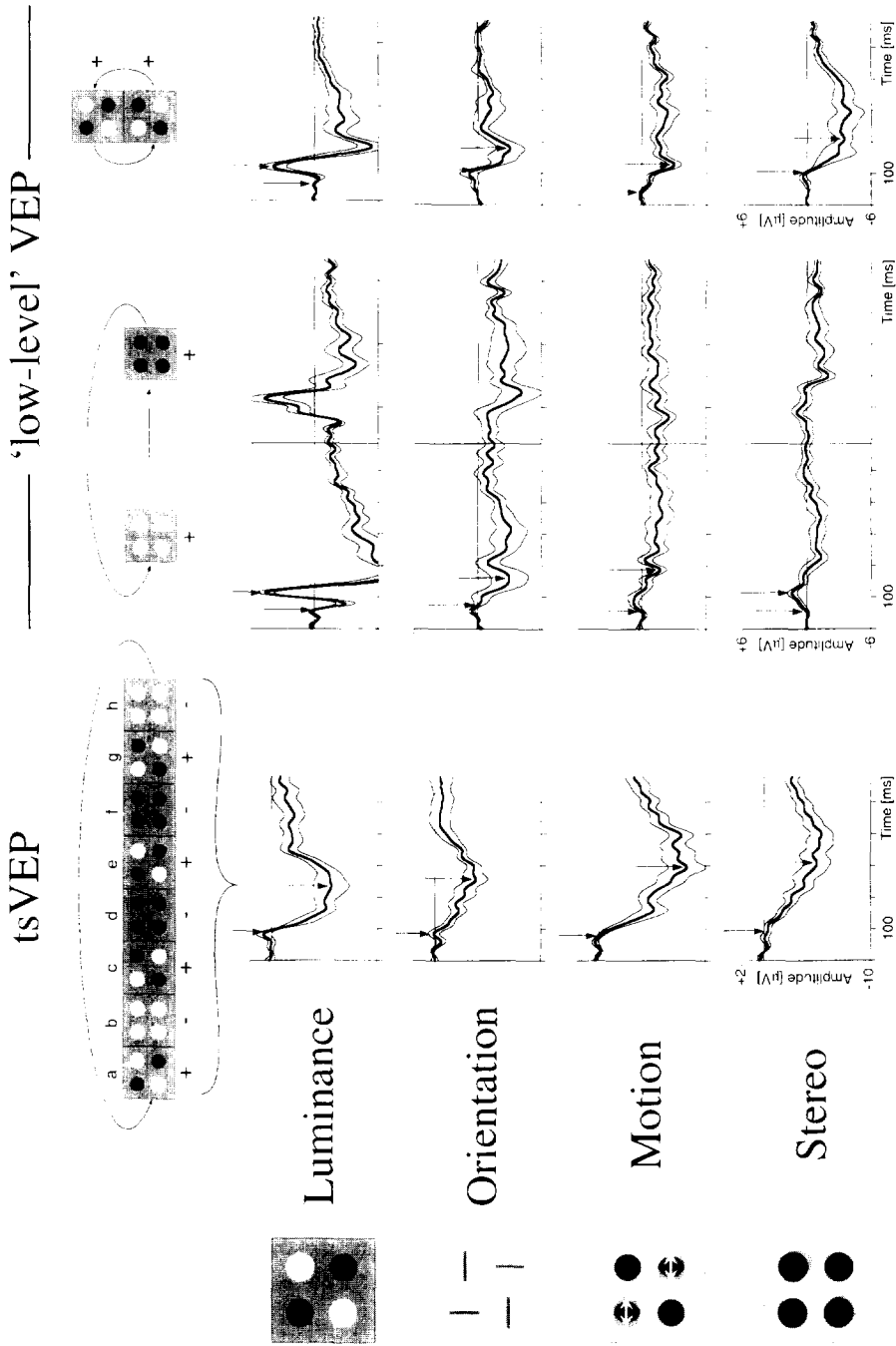


FIGURE 2. Visual evoked potentials associated with texture segregation (tsVEP, left) contrasted with low-level VEPs (right). On the very left the four visual dimensions are depicted symbolically. The next column depicts the respective tsVEP traces, which were extracted from the raw responses using the linear combination indicated by “+” or “-” in the stimulus pattern sequence at the top. On the right, there are two columns of low-level VEPs representing an “onset-offset” or a “reversal” response type of the respective visual dimension. The first arrow in each trace indicates implicit time, the second arrow denotes the “major peak”. Between the four visual dimensions, the tsVEPs have similarities (each shows a negativity at 230–310 msec) and dissimilarities. The similarities of the tsVEPs are striking when compared to those between the low-level VEPs. This suggests that the tsVEPs reflect a mechanism that has generalized to some degree across visual dimensions.

position by a random distance between -0.05 and $+0.05$ deg to further reduce local luminance artifacts. Furthermore, they were separated from each other by "white space" to reduce artifactual interactions (e.g., motion of adjacent spots stimulating detectors of orientation gradients). Owing to the positional jitter, the white space had a random length between 0.1 and 0.3 deg, resulting in an averaged white space of 0.2 deg and an averaged inter-element distance of 0.4 deg.

Electrophysiological recording

The VEP was recorded from an Oz-FPz derivation using gold-cup electrodes. Signals were amplified and filtered (first-order bandpass, 0.3–70 Hz, Toennies "Physiologic Amplifier") and digitised to a resolution of 12 bits at a sampling interval of 2.92 msec with a 386 AT-compatible computer. The computer averaged the sweeps if their amplitude did not exceed ± 100 mV and displayed them on-line while simultaneously generating the stimuli.

Electrophysiological procedure

We presented the various texture stimuli in an interleaved block design: each stimulus appeared 10 times, then the next stimulus followed. This cycle was repeated until a total of 120 sweeps for each condition was accumulated. The entire recording session lasted about 2 hr. To induce appropriate fixation and accommodation, the subjects reported random digits that appeared in the centre of the screen for 300 msec in random intervals between 2 and 10 sec.

Data analysis

We operationally defined the tsVEP as the difference between the response to the onset and offset of the "global checkerboard", independent of the local variant changes that induced the onset and offset. We presented an appropriate sequence of eight stimuli (Fig. 2, top left, a–h) that contained all possible transitions between "global checkerboard" and "globally homogeneous". Responses to the onset of global checkerboards are added with positive weights, responses to their offset are added with negative weights (indicated by the "+" or "-" symbols near the pattern sequences at the top of Fig. 2). For each local position the transition between the two variants (Table 1) occurs equally often with positive and negative weights, cancelling all low-level VEPs. Consider the top left element: from (h) to (a) it becomes dark; its response is cancelled by the negatively weighted transition from (c) to (d). Similar reasoning applies to all loci and transitions.

Low-level VEPs were extracted in two different ways, *onset/offset* and *reversal*, since we have not found a generic way to combine all possible local interactions in one sequence such that all local variants change and the global structure remains constant. *Onset* is defined as a transition from "all variant 2" to "all variant 1" (cf. Table 1), *offset* is the opposite transition; thus the onset/offset trace is likely to contain two, possibly different,

responses. For the visual dimension "motion" the offset transition was equivalent to "motion offset" for all disks (Table 1). As motion offset evokes a much smaller response than motion onset (e.g., Bach & Ullrich, 1994) and may be regarded as a "non-stimulus" compared to the offset conditions for the other visual dimensions, we focused our data analysis on the onset. In the case of *reversal*, a global checkerboard as defined by an arrangement of the two variants is continuously present, but the variants are exchanged. The two reversal responses were averaged, thus the reversal trace shows only one response. The icons in Fig. 2 under the heading "low-level" VEP represent schematically these stimulus transitions for the case of luminance.

Implicit times (equivalent to onset latencies) were defined as the zero-extrapolation of the regression line fitted to the slope of the first significant peak. This was performed with an automatic procedure as follows: the zero level was defined as the mean of the first 50 msec. Starting at 50 msec, the first point of the trace deviating from zero by more than 1.5 SEM defined t_0 , the first landmark for the regression interval. From this point, the trace was followed until its slope became zero, defining a peak at t_1 . Finally, a regression line was calculated for the time interval that comprised the centre 80% of the time interval (t_0, t_1). This reduction to the centre of the interval reduces the influence of the rounded parts at the borders of the interval, which would lead to an underestimation of the slope. We arrived at the parameters of this procedure after some trial and error with the aim of finding physiological plausible results with one algorithm for all traces. Peak latencies and amplitudes were measured from the "major peak", its polarity depending on the specific visual dimension in the case of the low-level VEPs.

To obtain a quantitative estimate of similarity of responses across all conditions, we calculated the cross-correlation function between the traces obtained in the various visual dimensions and took its maximum value, allowing for a differential time delay up to 100 msec.

Psychophysical procedure

The stimuli were basically identical to those in the electrophysiological experiments, but in place of the checkerboard arrangement, only a single check with 4 by 4 elements of variant 1 was presented against a background of variant 2 (cf. Table 1). The presentation interval of 117 msec (identical for all visual dimensions) was followed by a mask that contained a grid of alternating variants. In a four-alternative forced-choice design subjects detected the location of a single check in one of four positions (right, top, left, bottom).

RESULTS

In the psychophysical experiments, we found 99.7%, 100%, 95%, and 75% correct responses for luminance, orientation, motion and stereo, respectively, averaged across six subjects. Obviously, the stimuli were far above

TABLE 2. Latencies and amplitudes of tsVEPs

Visual dimension	Luminance	Orientation	Motion	Stereo disparity
tsVEP				
Implicit time [msec]	91	84	80	95
Peak latency [msec]	234	257	295	310
Amplitude at peak latency [μV]	-5.7	-3.9	-8.3	-5.0
Low-level VEP, onset				
Implicit time [msec]	58	75	55	56
Latency of major peak [msec]	114	161	187	114
Amplitude at peak latency [μV]	4.6	-3.0	-1.7	1.4
Low-level VEP, reversal				
Implicit time [msec]	53	111	42	105
Latency of major peak [msec]	108	178	129	211
Amplitude at peak latency [μV]	3.9	-3.1	-3.1	-3.5

threshold to induce a preattentive pop-out, though less so for stereo.

Significant tsVEPs were evoked by spatial gradients in all four visual dimensions tested (Fig. 2). All tsVEPs are negative, the amplitude varies by a factor of about two between visual dimensions, implicit time is about 90 msec and peak latency varies from 230 to 310 msec (Table 2). The significance level of the negative peak was $P < 0.001$ for every dimension. As a measure of similarity between the tsVEPs of two visual dimensions we obtained the peak value of their cross-correlation function. The peak values for all possible combinations across visual dimensions were (using the nomenclature L, luminance; O, orientation; M, motion; S stereo; units are μV^2): L-O 3.3, L-M 7.7, L-S 5.0, O-M 4.8, O-S 3.2, M-S 9.1; the average was $5.5 \mu V^2$.

Low-level VEPs exhibited a considerable variety of response shape across all four visual dimensions (Fig. 2, right). While our stimuli differ somewhat from classical VEP stimuli, there is a clear similarity between the luminance response (Fig. 2, top right) and the well-known P100 response (e.g., Regan, 1989). The amplitude of the low-level VEPs varied markedly between visual dimensions. Cross-correlation functions were obtained for the onset and the reversal response (nomenclature as above). Onset: L-O 3.2, L-M 2.0, L-S 1.0, O-M 1.0, O-S 0.5, M-S 0.3; the average is $1.3 \mu V^2$. Reversal: L-O 1.0, L-M 1.3, L-S 2.0, O-M 0.7, O-S 1.1, M-S 1.8; the average was $1.3 \mu V^2$. Averaged across the two response types of low-level VEPs, this results in a value of $1.3 \mu V^2$.

Both the tsVEPs and low-level VEPs vary across visual dimensions. The dependency on visual dimension, however, is much less for the tsVEPs. The low-level VEPs vary in polarity, number of peaks, latency, and markedly vary in amplitude. The tsVEPs, in contrast, are all negative, unimodal, have less variation in latency and vary less in amplitude. The averaged cross-correlation values for tsVEPs were 4.2 times larger than those obtained for the low-level VEPs ($P < 0.0001$). We thus take note of a striking similarity across visual dimensions between the tsVEPs when compared with the variety between the low-level VEPs.

DISCUSSION

As an electrophysiological correlate of the preattentive percept, we found significant tsVEPs for all four visual dimensions tested. This extends previous work based on the visual dimension "orientation" (Bach & Meigen, 1990, 1992; Lamme *et al.*, 1992, 1993b) and "motion" (Lamme *et al.*, 1993a, 1994) to "luminance" and "stereo disparity".

Through current density analysis and dipole localization, tsVEPs have been localized to area V1 for orientation (Lamme *et al.*, 1992a,b) and motion (Lamme *et al.*, 1993a, 1994). The similarity of tsVEPs across the four dimensions suggests that V1 is also a likely candidate locus for texture segregation mechanisms based on the visual dimensions luminance and stereo. Neurons that perform a preliminary step of gradient detection have been described for orientation contrast in primate V1 (Knierim & van Essen, 1992; Lamme, 1995) and cat (Blakemore & Tobin, 1972; Kastner *et al.*, 1995), and for motion contrast in primate V1 and V2 (Allman *et al.*, 1991; Lamme, 1995) and in cat (Kastner *et al.*, 1995).

As to our question how texture segregation differs between visual dimensions, we note that the tsVEPs across visual dimensions are very similar when compared to the large variety of low-level VEPs. When similarity was quantified with the cross-correlation function, its value was markedly higher among the tsVEPs as compared to the low-level VEPs. This suggests that the processes reflected by the tsVEPs possess certain similarities across the visual dimensions or may even be identical. Interestingly, psychophysical findings and models of texture segregation have also suggested similar processing for a wide range of visual dimensions (Treisman, 1985; Nothdurft, 1993; Wolfe, 1994). One would expect that the processing mirrored in the tsVEP occurs later than "low-level processing". Indeed, implicit times vary from 55 to 75 msec for the low-level onset-VEPs (Table 2). Given the difficulties in objectively assessing implicit time (see Methods), and with regard to the implicit time of the low-level reversal-VEP that range from 42 to 111 msec, we would not want to over-interpret these relatively small differences. Peak latencies of the

tsVEPs vary over a wider range, from 234 msec for luminance over orientation and motion to 310 msec for stereo (Table 2). While some of these differences may be traced to incomplete match of saliency, we note that the sequence of peak latencies is in keeping with our knowledge of visual processing: computation of luminance gradients can already be performed in the LGN (Nothdurft, 1990) with a consequently short latency. Extraction of orientation needs lateral processing and occurs in V1. Detection of motion requires lateral processing combined with a time delay, which to a first approximation could be estimated to 58 msec, half a period of the 17 Hz oscillatory motion. This would account for the higher latency of the motion tsVEPs. Stereo disparity as a hyperacuity may require extended local processing.

By simplifying models from the field of visual search (Treisman & Sato, 1990; Wolfe, 1994) and assuming that an identical process is activated by gradients in any visual dimension, we advance the following working hypothesis: filter stages (possibly in the LGN for luminance) and in the visual cortex (for orientation, motion and stereo) transform the gradients in each visual dimension into corresponding activity maps. These maps are integrated, possibly additively, into a "master map" or "saliency map", upon which border mechanisms and, later, Gestalt mechanisms can operate. This hypothesis generates specific predictions (like additivity across visual dimensions), testable and falsifiable by psychophysical or electrophysiological experiments, which might help to further our understanding of texture segregation.

REFERENCES

- Allman, J., Miezin, F. & McGuinness, E. L. (1991). Effects of background motion on the response of neurones in the first and second cortical visual areas. In Edelman, G. M., Gall, W. E. & Cowan, M. W. (Eds) *Signal and sense: local and global order in perceptual maps* (pp. 131–142). New York: Wiley-Liss.
- Bach, M. & Meigen, T. (1990). Electrophysiological correlates of texture-segmentation in human observers. *Investigative Ophthalmology & Visual Science (ARVO Suppl.)*, *31*, 104.
- Bach, M. & Meigen, T. (1992). Electrophysiological correlates of texture segregation in the human visual evoked potential. *Vision Research*, *32*, 417–424.
- Bach, M. & Ullrich, D. (1994). Motion adaptation governs the shape of motion-evoked cortical potentials (motion VEP). *Vision Research*, *34*, 1541–1547.
- Beck, J. (1972). Similarity grouping and peripheral discriminability under uncertainty. *American Journal of Psychology*, *85*, 1–19.
- Beck, J. (1983). Textural segmentation, second-order statistics, and textural elements. *Biological Cybernetics*, *48*, 125–130.
- Blakemore, C. & Tobin, E. A. (1972). Lateral inhibition between orientation detectors in the cat's visual cortex. *Experimental Brain Research*, *15*, 439–440.
- Julesz, B. (1986). Texton gradients: the texton theory revisited. *Biological Cybernetics*, *54*, 245–251.
- Julesz, B. & Bergen, J. R. (1983). Textons, the fundamental elements in preattentive vision and perception of textures. *The Bell System Technical Journal*, *62*, 1619–1645.
- Kastner, S., Nothdurft, H. C. & Pigarev, I. (1995). Neuronal responses to orientation and motion contrast in feline striate cortex. *Society for Neuroscience Abstracts*.
- Knierim, J. J. & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, *67*, 961–980.
- Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, *15*, 1605–1615.
- Lamme, V. A., van Dijk, B. W. & Spekreijse, H. (1992). Texture segregation is processed by primary visual cortex in man and monkey. Evidence from VEP experiments. *Vision Research*, *32*, 797–807.
- Lamme, V. A., van Dijk, B. W. & Spekreijse, H. (1993a). Contour from motion processing occurs in primary visual cortex. *Nature*, *363*, 541–543.
- Lamme, V. A., van Dijk, B. W. & Spekreijse, H. (1993b). Organization of texture segregation processing in primate visual cortex. *Visual Neuroscience*, *10*, 781–790.
- Lamme, V. A., van Dijk, B. W. & Spekreijse, H. (1994). Organization of contour from motion processing in primate visual cortex. *Vision Research*, *34*, 721–735.
- Meigen, T. & Bach, M. (1993). Perceptual ranking vs VEP for different local features in texture segregation. *Investigative Ophthalmology and Visual Science*, *34*, 3264–3270.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Nothdurft, H. C. (1990). Texture discrimination by cells in the cat lateral geniculate nucleus. *Experimental Brain Research*, *82*, 48–56.
- Nothdurft, H. C. (1993). The role of features in preattentive vision: comparison of orientation, motion and color cues. *Vision Research*, *33*, 1937–1958.
- Regan, D. (1989). *Human brain electrophysiology. Evoked potentials and evoked magnetic fields in science and medicine*. Amsterdam: Elsevier.
- Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics and Image Processing*, *31*, 156–177.
- Treisman, A. & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 459–478.
- Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*, 202–238.

Acknowledgements—This work was supported by the Deutsche Forschungsgemeinschaft. We also want to thank our colleagues who offered valuable critical comments on the manuscript: M. Fahle, S. Kastner, Ch. Nothdurft and J. Wolfe.