Neuromagnetic Correlates of Perceived Contrast in Primary Visual Cortex

J. D. Haynes, G. Roth, M. Stadler, and H. J. Heinze

1Department of Neurology II, Otto-von-Guericke University, D-39120 Magdeburg, Germany; 2Institute of Neuroscience, University of Plymouth, Plymouth PL4 8AA, United Kingdom; and 3Brain Research Institute, 4Institute for Psychology and Cognition Research, and 5Department of Neuropsychology and Behavioral Neurobiology, University of Bremen, D-28334 Bremen, Germany

Submitted 17 September 2002; accepted in final form 1 January 2003

Haynes, J. D., G. Roth, M. Stadler, and H. J. Heinze. Neuromagnetic correlates of perceived contrast in primary visual cortex. J Neurophysiol 89: 2655–2666, 2003. First published January 15, 2003; 10.1152/jn.00820.2002. When a target grating is flashed into a larger, surrounding grating, its contrast is perceived to be lower when both gratings are oriented collinearly rather than orthogonally. This effect can be used to dissociate the perceived contrast from the physical contrast of a target grating. We recorded the transient electric potentials and magnetic fields evoked by flashed target gratings and compared them with psychophysical judgments of perceived contrast. Both early (100 ms) and late (150 ms) transients were reduced in amplitude when targets were flashed into a collinear rather than orthogonal surround, thus paralleling the reduction in perceived contrast. Although targets in orthogonal backgrounds required 40% lower physical contrast to match the perceived contrast of collinear targets, the amplitudes of electrophysiological transients of matching stimuli were almost identical. Thus the responses correlated better with perceived than with physical target contrast. This holds especially for the late transient response. Source localization indicated that the transients in question may originate in primary visual cortex. Our results therefore identify the activity of primary visual cortex as one possible neural correlate of perceived contrast.

INTRODUCTION

There is growing evidence that the ability to discriminate visual contrast is limited by neuronal response rates in primary visual cortex. Psychophysical thresholds for contrast increments are consistent with a hypothetical “contrast transducer function” (CTF) of sigmoidal shape, which combines an accelerating regime at low contrast with a decelerating regime at high contrast (Legge 1981; Legge and Foley 1980). This is in good agreement with the actual contrast-response characteristics of single cells in primary visual cortex and both psychophysical CTFs and single-cell responses are often modeled by similar functions (Albrecht and Hamilton 1982; Chao-iy and Creutzfeldt 1984; Geisler and Albrecht 1997; Sclar et al. 1990). An even better match appears to exist between psychophysical thresholds for contrast discrimination and the population response of primary visual cortex, as inferred either from serial recording of multiple units (Geisler and Albrecht 1997) or from BOLD activation in functional MRI (fMRI) (Boynton et al. 1999).

The neural basis of the perceived magnitude of contrast has not been studied as extensively as that of contrast discrimina-

tion. Although models of perceived contrast and of contrast discrimination both assume sigmoidal CTFs with similar exponents in the high-contrast regime, the two functions may differ in important details (Cannon 1985; Cannon and Fullenkamp 1991a; Legge 1981). The comparison is further complicated by the fact that there is no agreement in the literature on how perceived magnitude of contrast is related to physical contrast (Bryngdahl 1966; Cannon 1979, 1985; Fiorentini and Maffei 1973; Franzen and Berkley 1975; Geogeson 1991; Kulikowski 1976).

Here we measure the perceived contrast of flashed stimuli with the help of a contrast matching paradigm and simultaneously measure the electrophysiological transients evoked by such stimuli in human visual cortex. It is well known that the perceived contrast of a target grating can be reduced significantly by the presence of a surrounding grating of higher contrast, a phenomenon also known as “contrast–contrast” (Cannon and Fullenkamp 1991b; Chubb et al. 1989; Ejima and Takahashi 1985; Snowden and Hammett 1998; Solomon et al. 1993; Xing and Heeger 2000, 2001). This masking effect reaches maximal strength when target and surrounding gratings exhibit identical spatial frequency, orientation, and speed (Cannon and Fullenkamp 1991b; Chubb et al. 1989; Solomon et al. 1993; Takeuchi and DeValois 2000), and diminishes with differences along any of these dimensions. As perceived target contrast is reduced more by collinear than by orthogonal surrounds, targets with collinear surrounds require additional physical contrast to match the perceived contrast of targets with orthogonal surrounds. In analogy to color vision, we speak of “contrast metamers” when targets of different physical contrast are matched in terms of perceived contrast. By dissociating physical and perceived contrast, contrast metamers let us establish whether the electrophysiological responses of visual cortex reflect physical contrast, perceived contrast, or neither. Using this approach it is possible to study the neural representation of perceived contrast without requiring any assumptions on the shape of the perceived magnitude function.

METHODS

Subjects

Eight right-handed subjects (2 male, 6 female; age range, 21–27 years) with normal vision (uncorrected) participated in the experi-

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
ment. All subjects had prior experience with electroencephalography (EEG) and magnetoencephalography (MEG) recordings. The experimental procedures were in conformity with the Declaration of Helsinki.

**Visual stimuli**

Masking background stimuli were two large (9.8° × 9.8°) square areas of high-contrast grating (0.79 Michelson luminance contrast) oriented at either 45° or 135° into which 16 small squares of gray were simultaneously inserted (Fig. 1A). Target stimuli were 16 square patches of square-wave grating (6 cpd, 1.1°) with variable Michelson luminance contrasts of 0.13, 0.20, 0.32, and 0.50 that were oriented either collinearly or orthogonally to the masking background. The 16 targets were presented simultaneously within the background. Mean luminance was 194 cd/m² over the entire display. In accordance with most previous studies of lateral masking, we chose the collinear targets to be in-phase with the surround. To decrease border effects, an isoluminant band of 0.1° was inserted between targets and masks. The border was not smoothed, following the results of Cannon and Fullenkamp (1991b), showing that the “sharpness” of the edge had no influence on masking. All displays contained a central fixation spot. We chose only two orientations because a larger set of orientations cannot be produced with cathode ray tube (CRT) or LCD displays without changing basic physical properties of the stimuli. The cardinal orientations 0° and 90° were not used because vertical gratings are distorted by the limited bandwidth of the video signal, whereas horizontal gratings are not, which is likely to lead to artifactual differences between the stimulus categories (Brainard et al. 2002). Square-wave gratings were chosen because the spatial resolution of the display in the recording chamber was fixed and did not enable a finer grading that would have been necessary to generate sine-wave gratings at our preferred spatial frequency. Because the main focus of our study was not on lateral masking in itself but on the manipulation of perceived contrast, we did not use additional stimuli with mask contrast of zero. Despite individual differences when targets and masks have similar contrasts, the inhibitory nature of lateral masking for targets that are (as in our case) presented in a surround of much higher contrast is well established (Cannon and Fullenkamp 1993; Xing and Heeger 2001).

Stimuli were generated using an LCD projector (SHARP XG-SV1E) that projected via two mirrors onto a rear projection screen in the shielded recording chamber. This was necessary to prevent electrical interference in the electromagnetic recordings. To provide linear luminance transfer characteristics, a combination of hardware and software gamma correction was applied. The hardware gamma setting of the graphics card (GEFORCE 2 GTS) was set to the point closest to linearity. Additionally a look-up-table (LUT) was used to fine tune the luminance transfer. Stimuli and LUTs were programmed using MATLAB. The electronic time delay between voltage changes in the video signal and luminance changes in the display, which is typical for LCD projectors (Brainard et al. 2002), was measured to be 18 ms and was corrected for.

**Procedure**

The subject triggered each trial with a keypress. A trial began with the presentation of the mask alone for 2,000 ms. The mask remained present throughout the entire duration of a trial. A sequence of eight identical stimuli, either collinear or orthogonal, was flashed into the holes in the background, each for 250 ms with a (uniformly distributed) random interstimulus interval of 650–850 ms. After a gap of 2,000 ms, a second sequence was presented with the same timing but using the other orientation. Two thousand milliseconds later, the mask disappeared, and the subject was required to give a response indicating which of the two stimulus trains had the higher contrast. In this modified two-alternative forced choice paradigm, the collinear stimuli were “standard” stimuli against which orthogonal stimuli of varying contrasts were tested, allowing us to estimate the orthogonal stimulus that was a subjective “match” to a given collinear stimulus. Our design departs from previous studies on lateral masking in two ways. Normally, targets and masks are presented synchronously, which was changed in our case to reduce the interference between transient mask and target responses. Second, we presented the stimuli in groups of eight, a change that was made to increase the number of electrophysiological transients recorded for each condition.

Following pilot studies, the high-contrast collinear stimulus (0.50) was paired with (compared with) orthogonal stimuli with contrasts 0.20, 0.32, or 0.50 (high-contrast condition). The low-contrast collinear stimulus (0.32) was paired with orthogonal stimuli with con-
trasts of 0.13, 0.20, or 0.32 (low-contrast condition). Due to this design, the total number of presentations was different for each stimulus category (600 for the 2 collinear stimuli, 400 for the middle 2 contrast levels of the orthogonal stimuli, and 200 for the lowest and highest orthogonal stimuli). Conditions, contrast levels, and orientations were randomly intermixed. Subjects were instructed to maintain fixation throughout each trial. Fixation was monitored during the sessions using an infrared camera. During recording setup, subjects practiced the task for a minimum of 10 min, leading to preadaptation before onset of the data acquisition.

EEG/MEG data acquisition

We simultaneously recorded 148-channel MEG– and 32-channel EEG–evoked transient responses at a sampling rate of 254 Hz filtered with a band-pass of 0.1–100 Hz. MEG was acquired with a BTI Magnes 2500 whole-head MEG system with 148 magnetometers (Biological Magnetic Technologies, San Diego, CA). EEG was recorded using a 32-channel Synamps amplifier (NeuroScan, Henderson, VA) with an electrode cap (Electrocap International, Eaton, OH) covering the channels Fz, Cz, Pz, Oz, Iz, Fp1, Fp2, F3, F4, F7, F8, T7, T8, C3, C4, P3, P4, O9, O10, P7, P8, FC1, FC2, CP1, CP2, PO3, PO4, PO7, PO8, plus extra electrodes for right horizontal EOG, right vertical electro-oculogram, and left mastoid. EEG channels were referenced to right mastoid and referenced off-line to the average of right and left mastoids. MEG was subjected to an on-line noise reduction process that removed a weighted sum of environmentally induced magnetic noise (1st order spatial gradients of the field) recorded by eight remote reference channels that do not pick up brain activity. Locations of EEG electrodes and MEG sensors were registered using a Polhemus 3Space Fastrak system with a common reference system defined relative to three anatomical landmarks (nasion and left/right preauricular points). These were coregistered with the individuals’ anatomical T1 magnetic resonance scans. To further enhance the individual peaks and remove contribution of low-frequency noise, the data were digitally high-pass filtered at 3 Hz, which does not significantly alter amplitudes of the early visual evoked components (Skuse and Burke 1990). Then data were sorted into stimulus-locked epochs of 600 ms length with 100 ms of pretrigger and subjected to artifact rejection, which removed epochs with peak-to-peak amplitudes exceeding a criterion of $3.0 \times 10^{-12}$ T for MEG or 100 $\mu$V for EEG data.

We chose to record transient rather than steady-state visual evoked responses for several reasons: 1) transient visual evoked responses (VERs) have the advantage of minimizing the contribution of motion processing, which is typically observed for counterphase reversing gratings used in steady-state designs (Murray and Kulikowski 1983); 2) they allow segregation of timecourses into separable components corresponding to different processing stages; and 3) they allow more straightforward equivalent current dipole modeling.

Analysis of behavioral data

The orthogonal contrast that matches the collinear standard stimulus was estimated by interpolating the data with a Weibull cumulative distribution function (Weibull 1951) for each condition and subject. The matching contrast can be found where the Weibull function takes a value of $P = 0.5$, which is where it is equally likely that the subject will judge either stimulus to be stronger (Fig. 1B).

Topography analysis

A 177-dimensional topography vector was calculated for each time-point sampled between 0 and 300 ms. This vector consisted of the amplitudes of the 148 magnetic plus 29 electric channels (excluding the 2 eye-channels and the left mastoid) scaled to a unit length of 1. For each subject, the normalized topography vectors from all conditions and timepoints were fed into a cluster analysis with a fixed number of 10 clusters (Euclidian distance metric; clustering algorithm based on average distance).

Source localization

There is considerable disagreement in the literature on the striate and extrastriate contributions to the early deflections of evoked electric and magnetic fields (e.g., Aine et al. 1995, 1996; Foxe and Simpson 2002; Ikeda et al. 1998; Jeffreys and Axford 1972a,b; Maier et al. 1987; Noachtar et al. 1993; Portin et al. 1998; Schroeder et al. 1991, 1998; Seki et al. 1996). This can be mainly attributed to differences in spatial and temporal stimulus parameters (quadrant, eccentricity, spatial frequency, onset vs. reversal vs. offset responses) to which early components respond sensitively. Studies of action potentials, local field potentials, and current source densities in multiple visual areas and cortical layers reveal that synaptic activity is temporally extended, occurs synchronously in multiple areas, shows signs of feedback and polarity inversion at different processing stages, and depends critically on spatial and temporal stimulus properties (Creutzfeldt and Kuhnt 1971; Lamme et al. 1998; Maunsell and Gibson 1992; Schroeder et al. 1991, 1998). Thus when using novel stimuli, it is impossible to infer the generators by simply referring to the literature. Instead, it is necessary to localize the dominant generators using equivalent current dipole modeling. According to its receptive-field tuning, V1 should contribute strongly to early deflections for stimuli with high spatial frequencies as in our case. Because the anatomical representation of a stimulus in V1 can be well estimated using retinotopic considerations (Aine et al. 1996; Horton and Hoyt 1991), we placed two seed points to the lateral side of the calcarine sulcus of either hemisphere at a depth of 2 cm from the occipital pole (Fig. 2A). Then, two equivalent current dipoles (ECDs) with fixed orientation and fixed location were fit within a radius of 5 mm of the seeds to the whole time period between 0 and 300 ms using the software package CURRY (Neuroscan, El Paso, TX). A three-sphere boundary element model (BEM) was used, and dipoles were fit to EEG and MEG data concurrently. Because the localization accuracy cannot be precisely known, we chose the 10 mm diameter of our bounding spheres to match previous estimates of combined EEG and MEG dipole localization accuracy (reviewed in Liu et al. 2002). Two, instead of eight, dipoles (1 for each target) were chosen because the expected distance between the cortical representations of the individual targets was below the expected localization accuracy. The stimulus parameters (especially spatial frequency) were set after several pilot experiments to the values that produced the most pronounced occipital-centered fields. To reduce noise due to contributions of nonvisual areas, we restricted the MEG channels to a circular set of 72 sensors centered around Pz. Most channels outside of this set had very low global field power.

Contrast response functions

For quantitative analysis of contrast responses, we measured peak amplitudes of the major evoked electric and magnetic responses for each contrast level, stimulus type, and subject. For MEG components, left and right hemisphere deflections were averaged across the peak negative and peak positive channels. For EEG components, the peak left and right hemisphere channels were averaged. To combine data across subjects, we normalized the response amplitudes for each subject and component individually by dividing the data by their mean. These data were then used to plot CRFs for collinear and orthogonal stimuli. For two subjects (kd83 and mn22), the electric and
magnetic P230 were too weak to be measured, so for this component, the results will be based on the data of only six subjects.

We also computed CRFs for the peaks in the timecourses of estimated dipole strengths to obtain a more pure estimate of striate cortical activity. The responses to the four levels of orthogonal stimuli were fit using the hyperbolic ratio-function, which is a standard model of contrast-dependent responses in primary visual cortex (Albrecht and Hamilton 1982; Boynton et al. 1996, 1999; Chao-yi and Creutzfeldt 1984; Geisler and Albrecht 1997; Sclar et al. 1990)

\[
R(C) = R_{\text{max}} \cdot \left[ \frac{C^{p+q}}{(C^{q} + \sigma^{q})} \right]
\]

\(R\) denotes response amplitude, \(C\) denotes stimulus contrast, \(R_{\text{max}}\) denotes maximum response, and \(\sigma\), \(p\), and \(q\) are parameters jointly determining slope and inflection point of the function.

**FIG. 2.** A: bounding spheres of the 2-dipole fit in the calcarine sulcus shown on top of an anatomical T1-weighted MR image. The center of the sphere shows the dipole seeds, and the sphere shows the region to which the final solution was restricted. The diameter of the bounding sphere was chosen to match previous estimates of electroencephalographic (EEG) and magnetoencephalographic (MEG) dipole localization accuracy (Liu et al. 2002). B: evoked electric and magnetic brain responses to orthogonal (red) and collinear stimuli (green) of same contrast (0.50) over the occipital pole (subject ka81) (top left). EEG and MEG responses are highly similar and left occipital and right occipital magnetic responses are mutually inverted as would be expected for a single dominant generator. Collinear responses are clearly reduced for all 3 major components [P80/M80 (I), N130/M130 (II), P230/M230 (III)]. Stimulus onset was at 0 ms. Top right: response topographies for the first peak (I) plotted for the time indicated by the dashed line (90 ms). Top: EEG data; bottom: MEG data. Figures are individually scaled to demonstrate that the topography is similar for both conditions. The circle, square, and diamond symbols indicate the scalp location of the channels shown on the left. Bottom left: results of cluster analysis for the same subject shown for the time between 30 and 230 ms. Each cluster is coded with a color and each column is a sample point. Each row is a stimulus category and is coded by its orientation relative to the background (collinear, C, or orthogonal, O) and its contrast. The peaks of the 3 major electric and magnetic responses are shown above the figure. Bottom right: evoked responses for the channel indicated by the square for high-contrast parallel stimuli presented in either the 1st (solid line) or 2nd intervals (dashed line) of trials. C: measured (“data”) and fitted (“fit”) response topographies for the peaks of the 2 striate components (subject ka81). Left: P80/M80 data (“early”); right: N130/M130 data (“late”). The quality of the fits is striking, especially regarding that dipoles (orange arrows) were placed on a priori retinotopical assumptions. The N130/M130 is an inversion of the P80/M80 response. Data are shown as 2-dimensional projections with posterior right at the bottom right.
Prediction of psychophysics by physiology

To assess the degree to which striate activity predicts perceived magnitude of contrast, it was necessary to compare the striate responses for the collinear standard stimuli with the according orthogonal matching stimuli. The responses to the orthogonal matching stimuli were interpolated using the hyperbolic ratio model fitted for each subject.

RESULTS

Psychophysics

Although our stimulus design departed from the literature, we measured a strong masking effect. Figure 1B shows psychophysical results for one subject in the high-contrast condition. The proportion of trials in which the orthogonal stimulus train of varying contrast (0.20, 0.32, 0.50) was judged as stronger than the collinear stimulus train with 0.50 contrast is plotted on the ordinate. The data show that for physically equal stimuli, the orthogonal stimulus was judged as higher than the collinear in 80% of the trials. An orthogonal stimulus of about 0.33 was judged as equal to the collinear standard. This occurs at \( \rho_{\text{orthogonal}} = 0.5 \), where both stimuli were chosen equally often and neither of the two stimuli appears to have a higher contrast. The difference between collinear standard and matching contrast indicates a reduction in perceived contrast of the collinear stimulus due to the orientation dependency of lateral masking. This reduction was present for both contrast levels and all subjects. Across all subjects, matching orthogonal grating contrasts were 0.18 for the 0.32 standard stimulus (1-sample t-test: \( t_{17} = -7.5, P < 0.001 \)) and 0.33 for the 0.50 standard stimulus (1-sample t-test: \( t_{17} = -9.1, P < 0.001 \)). Thus in both high- and low-contrast conditions, an orthogonal stimulus required only approximately 60% of the contrast of a collinear stimulus to match.

Physiology (waveforms and topography)

Electrical and magnetic responses followed a similar three-phase waveform (Fig. 2B, top left). Occipital EEG channels showed a sequence of positive-negative-positive deflections: (I) A positive component with onset latency around 80 ms and peak latency of 80–130 ms (P80); (II) a negative component with onset latency of 100–160 ms and peak latency of 130–180 ms (N130); and (III) a positive component with onset latency of 160–200 ms and peak latency of 190–260 ms (P230). The MEG channels showed highly similar temporal profiles, but with opposite polarities for left and right hemisphere channels, reflecting bipolar fields. We will label the MEG components in analogy to their electric counterparts with “M” (M80, M130, M230). All three major deflections show a reduction of amplitude for the collinear compared with the orthogonal stimulus with the same physical contrast. However, there is no change in latency between orthogonal and collinear stimuli of the same contrast.

Figure 2B (top right) shows the EEG and MEG response topographies for one subject to the high-contrast orthogonal and collinear stimuli. The data are individually scaled to show topographies for one subject to the high-contrast orthogonal stimuli of the same contrast.

During the first two major components (I and II), the clusters are largely identical for all conditions at a given latency. There are minor latency differences that show up as a shorter latency or phase advance for high-contrast stimuli. This has been described for single cells and evoked responses by previous authors (Albrecht 1995; Carandini and Heeger 1994; Heinrich and Bach 2001; Tyler and Apkarian 1985). Beginning with the last major deflection (III), the clusters are again distributed randomly. Thus during the first two deflections, the evoked brain topographies are highly similar, differing only by a linear scaling factor.

The bottom right of Fig. 2A also shows a graph comparing evoked responses during the first and second intervals. It demonstrates that there is no (observable) change in amplitudes (neither reduction nor enhancement) from the first to the second interval, as could have been expected following previous experiments on the effect of contrast adaptation on single cell responses and transient and steady-state evoked potentials (Bach et al. 1988; Carandini and Ferster 1997; Göpfert et al. 1999; Heinrich and Bach 2001; Ohzawa et al. 1982; Rebai and Bonnet 1989). Thus it seems likely that we are recording at a steady-state of adaptation where the temporal precedence of interval 1 over interval 2 does not play a major role. Long-term biphasic changes in event-related potential (ERP) amplitudes as previously described by Rebai and Bonnet (1989) do not seem to influence our data in a systematic fashion.

EEG and MEG contrast-response functions

The CRFs (Fig. 3) show the typical linear to slightly expansive shape when plotted as a function of log contrast, except for the EEG P230 component, which shows an initial decrease. This is in good accordance with previous studies on the relationship between stimulus contrast and response amplitudes as measured by EEG and fMRI (Boynton et al. 1999; Campbell and Maffei 1970; Göpfert et al. 1999; Tyler and Apkarian 1985). Response amplitudes for collinear stimuli are strongly suppressed compared with orthogonal stimuli at all contrast levels and for all components. We found no evidence for the “over-saturation” effects that have been observed in other studies, where response amplitude decreases at high-contrast levels (Chao-yi and Creutzfeldt 1984; Tyler and Apkarian 1985). This may be due to the fact that we did not study contrast levels higher than 0.5. Studies previously finding over-saturation were also mostly performed using steady-state stimulation, which may have introduced adaptation effects.

Source analysis

We used dipole modeling to assess the contribution of striate cortex to the individual components. The variance accounted for by two striate dipoles (Fig. 4) is 87.7 ± 6.39% (SD) for the P80/M80 peak, 88.1 ± 4.21% for the N130/M130 peak, and 70.1 ± 15.74% for the P230/M230 peak. Thus the early two components are dominated by striate responses, whereas the later component seems to have significant influences from other areas. The good fit between the data and our model-guided forward solution for the striate dipoles for the early two components is shown for one subject in Fig. 2C. Figure 4 shows the time courses of normalized dipole strengths and percentages of variance explained for each subject. The quality
of the fit is surprising given the complex stimulus shown and the fact that dipole locations were chosen on a priori anatomical assumptions. It is also surprising that the N130/M130 (“late striate”) component was so clearly an inversion of the P80/M80 (“early striate”) component as can also be seen in Fig. 2C. Similar results have been reported by other authors and may point to reentrant processing effects (Aine et al. 1995). The third component P230/M230 also had significant striate contributions, but the decrease in quality of fit suggests strong contributions from extrastriate visual areas. The residual fields after removing activity generated by the calcarine dipoles did not show any systematic pattern across subjects and were thus not further analyzed. The lack of any clear contribution of MT may be owed to the fact that the paradigm was tailored to yield strong striate effects. The spatial frequency we used is very close to the mean spatial frequency tuning of monkey and human V1, which is approximately 4 cpd. Beyond V1, cells respond preferentially to lower spatial frequencies (Foster et al. 1985; Geisler and Albrecht 1997; Singh et al. 2000). Besides the main effects, the data also point toward the existence of an earlier striate effect (around 20–60 ms), which is rather weak (as can be seen from the dipole strength) but nonetheless explains a considerable amount of variance for four of the subjects (nn22, nt68, kd83, and rg45). This effect was too weak to allow quantitative analysis, but it did not seem to differentiate between collinear and orthogonal stimuli.

FIG. 3. Contrast response functions for all components normalized and averaged across subjects. Error bars show SE.
To assess how well perceived contrast is predicted by activity in primary visual cortex, we restricted our analysis to the striate dipole timecourses of the first two deflections that had a dominant origin in V1. Figure 5A shows a CRF of the peak dipole amplitudes for the early striate component for collinear and orthogonal stimuli. The filled circle and filled square show the respective responses for two perceptually matching stimuli, S (collinear) and M (orthogonal). If perceived contrast is encoded in V1, the responses should be identical. For early and late striate components and both contrast levels, the responses to perceptually matching stimuli are similar, and the response amplitudes to the two physically identical stimuli are very different (Fig. 5B). Thus stimuli that are perceived to have the same contrast generate similar V1 responses, whereas stimuli that are perceived to have a higher contrast evoke stronger responses. The early striate component already shows a good match for perceptually matching stimuli, but there is a slight deviation at high contrasts. The second striate component on the other hand provides a good match at both contrast levels studied.

Further analysis shows that there is a strong correlation for individual subjects between the reduction of perceived contrast and the reduction of the response to the collinear standard stimulus (P80/M80, low contrast: Pearson’s $r = 0.87$; P80/M80, high contrast: Pearson’s $r = 0.77$; N130/M130, low contrast: Pearson’s $r = 0.87$; N130/M130, high contrast: Pearson’s $r = 0.62$). However, this correlation could be contaminated by differences between individual subject’s contrast response functions. Two subjects with the same reduction in perceived contrast might have quite different reductions in their individual contrast transfer functions between the standard and matching contrasts. For this reason, we chose a different measure to assess the quality of the match between response amplitude and perceived contrast at the individual subject level. Figure 6 shows a scatterplot of psychophysical matching contrasts and matching contrasts predicted from the contrast response function. The quality of our prediction shows up not as the correlation between the two variables but as the distance between the data points and this unity line. The scatterplot demonstrates a good match between psychophysically measured and predicted matching contrasts at the level of individual subjects. The match is slightly better for the late than for the early striate component, as shown by the smaller root mean square values for the difference between predicted and measured matching contrasts.

Table 1 shows the parameters estimated for Eq.1 collapsed across subjects, which can be used for a further comparison with previously published psychophysical data (Boynton et al. 1999; Legge 1981; Legge and Foley 1980). The exponent $p$ that governs the response behavior for mid- to high-contrast stimuli ($C \gg \sigma$) is similar for both early and late striate response components (about 0.3) and comparable with those found in previous studies on contrast discrimination (Legge and Foley 1980) and magnitude scaling (Cannon 1985). It is also similar to the exponents fit to fMRI contrast-response functions in human striate cortex (Boynton et al. 1999). We also explored how well the hyperbolic ratio function fit would extrapolate beyond the range of contrasts measured. In some
psychophysical models, the inflection point in the contrast representation function is used to explain the dip of the threshold versus contrast (TvC) function that occurs at low contrasts around 0.01 (Legge 1981; Legge and Foley 1980). We also computed the inflection point \( R'(C) \) for our contrast response functions. Our estimate is in a similar range as human psychophysics for the late striate component (0.03 for N130/M130) but not for the early striate component (0.13 for P80/M80).

**DISCUSSION**

Using a lateral masking paradigm that allowed the dissociation of physical and perceived contrast, we have demonstrated that activity of primary visual cortex correlates closely with perception. Previously, there has only been sparse evidence that perceived contrast may be correlated with the strength of V1 activity. Fiorentini and Maffei (1973) and Franzen and Berkley (1975) showed a correspondence between the slope of contrast response functions of steady-state evoked potentials and contrast representation data based on direct scaling. However, there are considerable discrepancies between perceived contrast functions obtained with different methods, such as magnitude scaling (Cannon 1979, 1985; Franzen and Berkley 1975), contrast halving (Kulikowski 1976), and luminance

![Figure 5](https://example.com/figure5.png)

**FIG. 5.** A: contrast response function (CRF) for early striate component (subject ka81). The amplitude of the 1st peak in the dipole time-course is plotted as a function of contrast. The collinear CRF is shown by open and filled circles and the orthogonal CRF by open diamonds. The neural response to the perceptually matching orthogonal stimulus (filled square) of a high-contrast colinear standard is interpolated from the contrast response data of the orthogonal stimuli for each subject by fitting a hyperbolic ratio function (Eq. 1). In the case of a perfect match between response amplitude and perceived contrast, the normalized dipole strengths for the standard and matching stimuli (filled circle and filled square) should be identical. B: mean response amplitudes averaged across subjects of striate components to collinear standard stimuli (C), orthogonal stimuli with same physical contrast (O), and same perceived contrast (OM). Response amplitudes are different for physically identical stimuli (C and O) and similar for perceptually matching stimuli (C and OM). Both components closely parallel perception, but the P80/M80 shows a slight but significant deviance for high contrasts. The dotted line shows the location of the mean for the collinear stimulus. Statistics were computed to test for an overall effect of stimulus category, and single tests were performed for the pairs C-O and C-OM. Low contrast, early: ANOVA \( F(1,7) = 33.6, P < 0.001; \) paired-samples \( t \)-test for comparison O-C \( t_{(7)} = 5.8, P < 0.001; \) comparison O-OM \( t_{(7)} = -3.8, P = 0.045. \) High contrast, early: \( F(1,7) = 88.1, P < 0.001, \) comparison O-C \( t_{(7)} = -9.4, P < 0.001, \) comparison O-OM \( t_{(7)} = 2.8, P = 0.028. \) Low contrast, late: \( F(1,7) = 13.8, P = 0.007, \) comparison O-C \( t_{(7)} = -3.7, P = 0.008; \) comparison O-OM \( t_{(7)} = 0.743, P = 0.482. \) High contrast, late: \( F(1,7) = 12.7, P = 0.009; \) comparison O-C \( t_{(7)} = -5.3, P = 0.010, \) comparison O-OM \( t_{(7)} = 2.8, P = 0.028. \) Error bars show SE.
matching of peaks and troughs (Bryngdahl 1966; Fiorentini and Maffei 1973). These disagreements make it difficult to link physiology and psychophysics based on the general shape of functions. An approach more closely related to our study was used by Goodyear et al. (2000). Using BOLD-fMRI, they demonstrated for one clinical subject with a monocular reduction of perceived contrast due to amblyopia that stimuli demonstrated for one clinical subject with a monocular reduction of perceived contrast generate similar responses in early visual cortex (presumably V1/V2). With our lateral masking paradigm, we have been able to confirm this finding in normal subjects and investigate the temporal properties of the striate response.

Previous results showing a lack of binocular transfer in lateral masking (Chubb et al. 1989), despite its orientation selectivity, have been used to argue for a strong role of primary visual cortex in the representation of perceived contrast. On the one hand, V1 is the last processing stage with substantial physiological and psychophysical ramifications for low-contrast processing. On the other hand, it is the first stage of orientation-selective processing, which is necessary to account for the orientation dependency of lateral masking. Subcortical visual neurons in the lateral geniculate nucleus (LGN) are not orientation selective, so targets with the same physical contrast should be masked in a similar manner by collinear and orthogonal flanks, at least in the feedforward sweep of processing. Likewise, a lower contrast orthogonal stimulus matched in perceived contrast to the collinear stimulus should evoke less LGN activation. If the stimulus representation in V1 were similarly based on physical contrast responses, evoked V1 responses should also follow this pattern. Our results, however, show similar responses for the contrast metamer and different responses for physically identical stimuli.

The data demonstrate a temporal development in the rescaling process. The first striate deflection already predicts perceived better than physical contrast but still has a slight deviation at high contrasts. The second deflection predicts the perceptual data even better and provides a close match for both contrast levels studied. It also extrapolates beyond the range of contrasts measured to predict the dip in the contrast discrimination data for low contrasts shown in previous studies (Boynton et al. 1999; Legge 1981; Legge and Foley 1980). Both components also allow good prediction of individual differences in perceived contrast reduction.

Numerous studies at the level of single cells and populations have demonstrated that primary visual cortex exhibits surround-effects that can account for the current data (e.g., Blake et al. 2000; Levitt and Lund 1997; Nelson and Frost 1978; Polat et al. 1998; Sengpiel et al. 1997; Walker et al. 1999). Specifically, several studies have directly shown the influence of surround effects on contrast transfer functions (Polat et al. 1998) and the dependency of surround effects on the relative contrast between center and surround (Levitt and Lund 1997; Polat et al. 1998; Somers et al. 1998; Toth et al. 1996). It is possible that the anatomical substrate of this surround modulation is feedback from higher visual areas. Mutual feedforward and feedback connections are known to exist between V1 and many extrastriate visual areas (Bullier 2001; Lamme et al. 1998; Salin and Bullier 1995). However, the most detailed study so far showed that inactivation of V2 has no effect on the surround modulation of responses in V1 (Hupé et al. 2001).

A second candidate is the rich plexus of horizontal connections in primary visual cortex. These connections have a range of =8 mm and tend to preferentially link iso-oriented orientation columns (Gilbert 1992; Gilbert and Wiesel 1979; Malach et al. 1993; Martin and Whitteridge 1984; Mitzchick and Crick 1982; Rockland and Lund 1982; Schmidt et al. 1997). This orientational anisotropy is of special interest because it may be able to account for the orientation dependency of lateral masking. The temporal dynamics observed in our study may provide a further clue as to the mechanisms. Horizontal connections are slow (approximately 0.1–0.3 m/s) (Bringuel et al. 1999; Girard et al. 2001; Grinvald et al. 1994), whereas feedforward and feedback connections are fast (approximately 3.5 m/s) (Girard et al. 2001). Based on estimates by Bullier (2001), a feedforward-feedback cycle between V1 and V2 could be completed within 4 ms. Horizontal propagation across a distance of one-half the size of our targets (0.55°) should take approximately 55 ms. The fact that our second component predicts perception better, thus fits in with a slow horizontal integration process. Interestingly, we observed no change of response latency by lateral masking. Thus although the response amplitudes for collinear compared with orthogonal stimuli were reduced, the response phase was the same. Changes in physical contrast normally lead to changes in response amplitude as well as latency, which can be accounted for by automatic gain control mechanisms (Carandini and Heeger 1994). In our case, a lack of a latency effect may indicate that the reduction of response amplitude occurs after this stage of automatic gain control.

It should be noted that some other masking paradigms can produce an opposite effect of cross-orientation inhibition (Burr and Morrone 1987; Morrone et al. 1982) when using superimposed targets and masks. These effects can be explained by models of local divisive inhibition (Carandini et al. 1997). Our results also differ from those of Polat and Norcia (1996), who observed facilitation for collinear suppression for orthogonal target-flank combinations. However, these authors used steady-state visually evoked potentials, rendering it difficult to judge whether they recorded predominantly striate activity. They also found their collinear facilitation effects at far lower contrasts than we used. A subsequent study of the same authors revealed a biphasic dependency of surround interactions on contrast, with the interaction being facilitatory for low- and inhibitory for high-target contrasts (Polat et al. 1998). In their earlier studies, Polat and Norcia (1996) presumably recorded from the low-contrast end, while we recorded from the high-contrast end of this biphasic function.

The fact that primary visual cortex activity correlates with perceived contrast can also be discussed within the framework of the “neural correlates of visual awareness” (Block 1996; Crick and Koch 1995, 1998; Lamme et al. 2000; Rees et al. 2002; Roth 2000). Previous studies have produced controversy...
sial results regarding the role of V1 in visual awareness. Studies showing that V1 can respond to stimulus features that do not enter consciousness seem to rule out V1 as a place where any dimension of conscious perception could be directly represented (Cumming and Parker 1997; Gur and Snodderly 1997; Herrmann 2001; Maier et al. 1987). On the other hand, an intact V1 may be a necessary condition for the visual awareness of spatial patterns (Stoerig and Cowey 1997), and several studies have shown a close correlation between V1 activity and perceptual phenomena such as metacognition and assimilation (Bridgeman 1980; Macknik and Livingstone 1998), binocular rivalry (Polonsky et al. 2000), and the percepts elicited by electrical cortical stimulation (Lee et al. 2000). In accord with our data, several studies have also shown that V1 activity predicts the perception of brightness when the percept is manipulated by contextual manipulations (Kinoshita and Komatsu 2001; MacEvoy and Paradiso 2001; Rossi et al. 1996, 1999). Our observation of the temporal unfolding of the percept-based contrast response might help explain why several studies failed to find a correlation between V1 responses and perceptual experience. It suggests that in some cases this correlation may only be present for the late temporal stages of V1 processing. Supporting this idea, Kinoshita and Komatsu (2001) showed that representation of the luminance of large homogenous fields is present at late (sustained) but not at early (transient) phases of striate processing. Likewise, Supér et al. (2001) demonstrated that conscious perception of texture-defined figures critically depends on late rather than early striate responses. They believe this is a consequence of feedback processes from extrastriate visual areas. Several other authors have stressed the role of reentrant processing for visual awareness (Bridgeman 1980; Di Lollo et al. 2000; Enns and Di Lollo 2000; Lamme and Roelfsema 2000). Following the studies on perceived brightness, the present study shows that perceived contrast is a further dimension for which a close correspondence exists between perception and response amplitudes in primary visual cortex.

The authors thank M. Greenlee for technical advice, J. Braun, R. Fendrich, E. Freeman, and U. Ernst for valuable comments on the manuscript and our student C. Morawetz for assisting in the experiments.

REFERENCES


