

Dissociation Between Local Field Potentials and Spiking Activity in Macaque Inferior Temporal Cortex Reveals Diagnosticity-Based Encoding of Complex Objects

Kristina J. Nielsen, Nikos K. Logothetis, and Gregor Rainer

Max Planck Institute for Biological Cybernetics, D-72076 Tübingen, Germany

Neurons in the inferior temporal (IT) cortex respond selectively to complex objects, and maintain their selectivity despite partial occlusion. However, relatively little is known about how the occlusion of different shape parts influences responses in the IT cortex. Here, we determine experimentally which parts of complex objects monkeys are relying on in a discrimination task. We then study the effect of occlusion of parts with different behavioral relevance on neural responses in the IT cortex at the level of spiking activity and local field potentials (LFPs). For both spiking activity and LFPs, we found that the diagnostic object parts, which were important for behavioral judgments, were preferentially represented in the IT cortex. Our data show that the effects of diagnosticity grew systematically stronger along a posterior–anterior axis for LFPs, but were evenly distributed for single units, suggesting that diagnosticity is first encoded in the posterior IT cortex. Our findings highlight the power of combined analysis of field potentials and spiking activity for mapping structure to computational function in the brain.

Key words: monkey; visual cognition; object recognition; electrophysiology; psychophysics; behavior

Introduction

Because we live in a three-dimensional world, distant objects are often only partially visible, and in part covered by closer objects. Under most circumstances, partially occluded objects are recognizable despite the lack of information about the occluded shape regions. However, it has been demonstrated previously that occlusion of specific, behaviorally relevant shape regions renders both humans and monkeys unable to perform tasks on partially occluded shapes (Biederman, 1987; Gosselin and Schyns, 2001; Nielsen et al., 2006). Occlusion of other shape regions leads to no behavioral impairments. In this study, our goal is to systematically examine how occlusion of visual shape regions of differing behavioral relevance impacts the neural representation of these shapes in the inferior temporal (IT) cortex of the macaque monkey.

The IT cortex is thought to play a major role in object recognition processes and contains many neurons that respond to ethologically relevant objects such as faces (Perrett et al., 1982; Desimone et al., 1984), but also to arbitrary shapes after the monkey has learned to identify them (Logothetis and Sheinberg, 1996; Tanaka, 1996). IT neurons have been shown to retain their shape selectivity despite occlusion of randomly selected shape portions (Kovács et al., 1995). However, it has not been tested whether

occlusion effects depend on which parts of a shape are occluded, taking the behavioral relevance of the occluded shape parts into account. Yet, several studies have provided evidence that parts of objects can be sufficient to evoke responses from IT neurons (Tanaka et al., 1991; Tsunoda et al., 2001; Baker et al., 2002). Furthermore, it has been shown that learning modifies neural responses in the IT cortex. Learning of associations between different shapes (Sakai and Miyashita, 1991; Messinger et al., 2001) and learning of task-relevance of shape features (Baker et al., 2002; Sigala and Logothetis, 2002; Sigala, 2004) are both reflected in IT cortical neural activity. As different shape regions acquire behavioral relevance because of training on a task, and are thus the outcome of a learning process, it is likely that the effects of occlusion will depend on the behavioral relevance of the occluded shape parts.

Materials and Methods

Behavioral and electrophysiological methods. Two adult male monkeys (*Macaca mulatta*) participated in the experiments. All studies were approved by the local authorities (Regierungspräsidium, Tübingen, Germany) and were in full compliance with the guidelines of the European Community (European Union directive 86/609/EEC) for the care and use of laboratory animals. Stimuli were presented on a γ -corrected 21 inch monitor, placed at a distance of 97 cm from the monkeys. Each image subtended 6 by 6° of visual angle. Stimuli were generated as described in a previous study (Nielsen et al., 2006). Of the six natural scenes used in the previous study, we chose four scenes for each monkey. The average gray-scale value of each stimulus was set to the same value to control overall luminance. Furthermore, all modifications of an image had the same overall contrast as the original image (measured as the SD of the gray-scale values). For occluded images, only unoccluded image

Received May 29, 2006; revised July 21, 2006; accepted Aug. 10, 2006.

This work was supported by the Max Planck Society. G.R. is a Deutsche Forschungsgemeinschaft Heisenberg investigator (RA 1025/1-1). We thank N. Sigala and A. Tollas for comments on this manuscript.

Correspondence should be addressed to Dr. Gregor Rainer, Max Planck Institute for Biological Cybernetics, Speemannstrasse 38, D-72076 Tübingen, Germany. E-mail: gregor.rainer@tuebingen.mpg.de.

DOI:10.1523/JNEUROSCI.2273-06.2006

Copyright © 2006 Society for Neuroscience 0270-6474/06/269639-07\$15.00/0

parts were considered when computing the mean and SD of the gray-scale values.

During the recording sessions, the monkeys performed a fixation task. Each trial began when the monkeys acquired fixation on a central fixation point. After a variable baseline duration of at least 100 ms, a stimulus was presented for 500 ms. The monkeys were required to maintain fixation within 1° of the center of the screen for the whole trial. Fixation was monitored with a scleral search coil and sampled at 200 Hz (CNC Engineering, Enfield, CT). Successful fixation was rewarded with a drop of juice delivered 1 s after stimulus offset. The monkeys completed at least 10 repetitions for each condition during a recording session.

Single-cell activity and the local field potential (LFP) were recorded from a recording chamber consisting of a ball-and-socket joint with an 18-gauge stainless-steel tube passing through its center (Schiller and Koenner, 1971). Horsley–Clark coordinates for the chambers were antero-posterior (AP), 18.1, mediolateral (ML), 17.7 for monkey 1, and AP, 15.4, ML, 16.8 for monkey 2. Neural signals were recorded using a five-channel electrode drive (Thomas Recording, Giessen, Germany), and platinum/tungsten electrodes coated with quartz glass with an impedance between 1 and 2 M Ω (ESI2ec; Thomas Recording). The recorded signal was divided into multiunit activity (band-passed signal between 500 Hz and 10 kHz) and LFPs (band-passed signal between 1 Hz and 100 Hz). From the multiunit activity, the activity of single neurons was extracted using standard spike-sorting techniques (Offline Sorter; Plexon, Dallas, TX). To ensure an unbiased estimate of neural activity, we made no attempt to select neurons based on task selectivity. Instead, we advanced each electrode until the activity of one or more neurons was well isolated and then began collecting data. The position of each electrode in terms of AP and ML coordinates and distance from the superior temporal sulcus was noted. We sampled different AP positions in a systematic manner in both monkeys. In monkey 1, initial recording positions were anterior; over the course of the experiments, the recording positions were moved more and more posterior to minimize structural brain damage caused by guide tube movement. In monkey 2, we proceeded in the opposite way, and recording locations were moved from posterior to anterior locations.

Data analysis. Single-unit activity was analyzed in a 300 ms time window beginning 100 ms after stimulus onset to account for neural latency. Baseline activity was determined in a 100 ms time window preceding stimulus onset. Spike density functions were computed by convolution of the spike trains with a Gaussian kernel ($\sigma = 10$ ms), using a resolution of 1 ms. Spike density functions of different modifications of the same scene were normalized by the maximal value observed across all modifications. The diagnostic variance was computed as $V_{\text{Group}}/V_{\text{total}} \times 100\%$, where V_{total} is the total firing rate variance,

$$V_{\text{group}} = 1/2 \cdot [(\bar{f}_{\text{diag}} - \bar{f})^2 + (\bar{f}_{\text{ndiag}} - \bar{f})^2],$$

where \bar{f}_{diag} , \bar{f}_{ndiag} , and \bar{f} represent the diagnostic, nondiagnostic, and overall firing rate mean, respectively (Bortz, 1993).

The LFP, which was originally sampled at a rate of 4.46 kHz, was first downsampled to 1 kHz. A bandpass filter (first order Butterworth filter, bandpass between 5 and 80 Hz) was applied to remove slow drifts. Finally, each LFP channel was z-transformed using the mean LFP amplitude and SD of the channel in the 100 ms baseline period preceding stimulus onset. Visual evoked potentials (VEPs) were computed by stimulus-locked averaging of the LFP data. Individual sites were identified as responsive to a particular stimulus if the absolute VEP amplitude was larger than 1.5 SD at three consecutive time bins during the stimulus presentation. Computation of the variance explained by diagnosticity was based on the mean LFP amplitude in an interval of 20 ms duration centered on the maximum of a positive VEP peak at ~ 140 ms (P140). The same formula was used as for the single units, but replacing mean firing rate with mean LFP amplitude. The P140 latency depended on the visible stimulus size; it also differed between monkeys. We therefore used a different interval for each condition and monkey. Because the visible stimulus size seemed to be the major determinant of the peak latency, the same interval was used for diagnostic and nondiagnostic conditions of the same visible stimulus size. Intervals were always 20 ms long. Their placement was determined by computing the grand average VEP over all

responsive LFP cases from one monkey for one particular stimulus size (either full, 10, 30, or 50%). The latency of the peak of the P140 was determined, and used as the center for the 20 ms interval.

Results

In a previous study, two Rhesus monkeys (*Macaca mulatta*) were trained to discriminate between the members of small sets of natural scenes. We used natural scenes because they are good examples of complex visual stimuli, and contain information at many different spatial scales. After training, we systematically determined the parts of each scene that the monkeys relied on to perform the discrimination task (Nielsen et al., 2006). To investigate how occlusion of different scene parts influenced neural responses in the IT cortex, we used these results to split each scene into parts with and without behavioral relevance. By constructing appropriate masks, we generated three occluded versions of each scene which revealed only the behaviorally relevant parts of the scene (diagnostic conditions). We similarly constructed three occluded versions in which only behaviorally irrelevant scene parts were visible (nondiagnostic conditions). Across the three diagnostic versions of each scene, and similarly across the three nondiagnostic versions, we varied how much of the original scene remained visible (visible stimulus size: 10, 30, or 50%). Exemplar stimuli are shown in Figure 1. To avoid low-level differences between conditions, all stimuli were adjusted to have the same mean luminance, as well as the same overall contrast. Because each of the monkeys relied on different image regions to perform the discrimination task, each monkey had its own stimulus set. We verified that the monkeys could correctly identify a scene when presented with any of the diagnostic, but not when presented with any of the nondiagnostic conditions (Nielsen et al., 2006).

Using these behaviorally defined stimuli, we recorded the activity of well isolated single neurons in area TE in the two monkeys. During the recording sessions, the monkeys viewed a set of 28 stimuli, consisting of four scenes and the corresponding diagnostic and nondiagnostic conditions. Activity of 423 neurons was recorded from both monkeys. Neural responses to the different scenes were treated independently. For each neuron, the responses to all versions of a scene were included in the analysis if at least one version (either the original scene or one of the modifications) evoked significant excitatory responses from the neuron (t test vs baseline activity, $p < 0.05$ corrected for the 28 comparisons). Thus, each neuron could contribute between one and four “cases” (the responses to all versions of one scene) to the group analysis. By these criteria, 220 cases generated by 135 neurons were selected for additional analysis.

The activity of an exemplar neuron from this group is shown in Figure 1. Presentation of the full natural scene elicited a visual response from the neuron (t test vs baseline activity, $p < 10^{-6}$), as did all diagnostic conditions ($p = 0.0004$, $p < 10^{-4}$, and $p < 10^{-4}$ for the three conditions). In contrast, only the largest nondiagnostic condition triggered a significant response from the neuron ($p = 0.54$, $p = 1.0$, and $p = 0.003$ for these conditions). The visible stimulus size also influenced the neural firing rates, with larger responses to stimuli that revealed more of the original natural scene. However, responses to the diagnostic conditions were always larger than to the nondiagnostic conditions (ANOVA with factors diagnosticity and size: main effect size, $p < 0.001$; main effect diagnosticity, $p < 0.001$; interaction, $p = 0.29$).

Similar effects were seen across the whole population of neurons. The population spike density function (Fig. 2a) showed a response to the full stimulus that began with the typical latency of

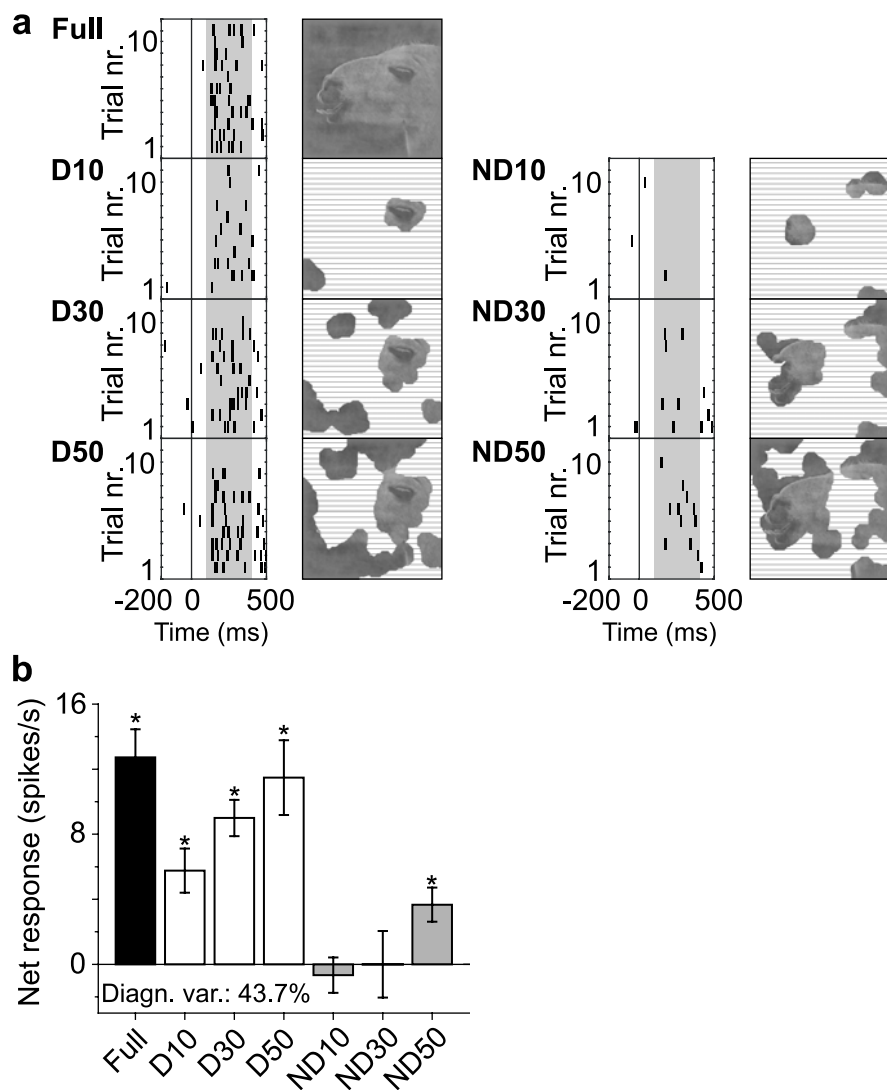


Figure 1. Responses of an exemplar single unit. *a*, Raster plots. In these plots, each line denotes the occurrence of an action potential generated by the selected neuron (stimulus onset at 0 ms). Each plot summarizes the responses in one of the seven conditions, using the stimuli shown next to each raster plot (occluded image parts are shown as hatched regions; they were gray in the actual stimuli). The labels next to each raster plot indicate the type of condition (D, diagnostic; ND, nondiagnostic; numbers correspond to the visible stimulus size). The gray region in each raster plot corresponds to the time window used for computing the stimulus evoked firing rate. *b*, Average net firing rate for the selected neuron. Error bars denote SEM, asterisks indicate conditions for which a *t* test between stimulus and baseline firing rate yielded $p < 0.05$. As a reference, the diagnostic variance for this case is given in the plot.

TE neurons of ~ 100 ms (Baylis et al., 1987; Tamura and Tanaka, 2001) and lasted throughout the stimulus presentation period. As shown in Figure 2*b*, the corresponding diagnostic and nondiagnostic conditions evoked significantly less activity than the full stimulus (paired *t* tests between the full and other conditions, $p < 0.001$ in all cases). Thus, occlusion of parts of the original scenes in general reduced the response rate of TE neurons. Similar effects of occlusion on responses of TE neurons have been reported previously (Kovács et al., 1995). In addition, we found that responses to diagnostic stimulus parts were greater than responses to nondiagnostic parts. Diagnostic conditions resulted in larger mean firing rates than nondiagnostic conditions at all visible stimulus sizes (paired *t* tests between conditions of the same visible stimulus size, $p \leq 0.02$ in all three cases). Interestingly, responses in the diagnostic conditions were on average independent of the visible stimulus size (one-way repeated measures

ANOVA, $p = 0.9$). Revealing 10% of the image with high behavioral relevance triggered the same responses as revealing 50% with high behavioral relevance. The differences between diagnostic and nondiagnostic conditions were seen in many individual neurons. For 90 cases, firing rates in at least one diagnostic condition were significantly different from the matching nondiagnostic condition (*t* test, $p < 0.05$ adjusted for the three comparisons). Furthermore, we plotted the net firing rate of each case to a diagnostic stimulus condition against the net firing rate for the matching nondiagnostic condition (for the visible stimulus size of 10%, see Fig. 2*c*). At all visible stimulus sizes, more cases had higher firing rates in the diagnostic than in the nondiagnostic condition (χ^2 test, visible stimulus size of 10%, 147 vs 70 cases, $p < 0.001$; 30%, 129 vs 87, $p = 0.004$; 50%, 127 vs 90, $p = 0.01$). Diagnostic regions were determined in experiments in which the monkeys performed a discrimination task. In contrast, we used a passive fixation paradigm for the neurophysiological recordings. We verified in a separate control experiment that our findings were not influenced by the different tasks (supplemental experiment 1, Fig. 1, available at www.jneurosci.org as supplemental material). In conclusion, our results indicate that the behavioral relevance of a scene part is a major determinant for the influences of occlusion in area TE. Note that many neurons also responded to nondiagnostic parts, suggesting that learning the visual discrimination task led to a relative reweighting of neural representations of parts according to their diagnosticity, but did not completely abolish responses to nondiagnostic regions.

Given the robust influences of diagnosticity on occlusion effects on the population level, we investigated how different subregions of TE were influenced by diagnosticity. A subdivision of TE into smaller subregions has been suggested based on anatomical data (Seltzer and Pandya, 1978; Iwai and Yukie, 1987; Yukie et al., 1990), but also because a functional specialization of neurons has been observed in different parts of TE (Hasselmo et al., 1989; Janssen et al., 2000; Perrett et al., 1991, 1992; Tamura and Tanaka, 2001). To map the influences of diagnosticity across TE, we quantified the effect of stimulus diagnosticity on each case by computing how much of the total trial-by-trial variance in firing rate could be explained by diagnosticity (the “diagnostic variance”). If the firing rate of a neuron for the occluded conditions was solely determined by the diagnosticity of the visible parts, then the diagnostic variance equals 100%, whereas a diagnostic variance of 0% would indicate equal responses to diagnostic and nondiagnostic conditions.

Cases with high diagnostic variance values responded preferentially to diagnostic scene parts. To show this, we selected the

cases for which the diagnostic variance value was above the 90th percentile of all diagnostic variance values. The average spike density function for these 22 “high-diagnosticity cases” is plotted in Figure 3*a*, showing that the neurons with high diagnostic variance values indeed responded more strongly to diagnostic than to nondiagnostic conditions (repeated measures ANOVA on the mean firing rates with factors diagnosticity and size; significant main effect of diagnosticity, $p = 0.001$, with no interaction with size, $p = 0.92$). Similarly, across all cases, higher diagnostic variance values were associated with increasingly larger responses to diagnostic than to nondiagnostic conditions (supplemental Fig. 2, available at www.jneurosci.org as supplemental material).

To study the influences of diagnosticity across TE, we plotted the recording locations of the high diagnosticity cases. Figure 4*a* shows their distribution along the AP dimension of the recording region for one of the monkeys. The results for the other monkey were similar (supplemental Fig. 3, available at www.jneurosci.org as supplemental material). For each monkey, we divided the recording region into two halves with an equal extent along the AP axis. We found that high-diagnosticity cases were evenly distributed across the posterior and anterior half of the recording region (monkey 1, χ^2 test, $p = 1$; monkey 2, χ^2 test, $p = 1$). As a second step, we plotted the diagnostic variance of each case as a function of its AP recording location (Fig. 4*c*). There was no influence of the recording location on the diagnostic variance. This was the case for each monkey individually, as well as for the data of both monkeys combined (Pearson correlation coefficients not significantly different from zero; monkey 1, $p = 0.5$; monkey 2, $p = 0.9$; combined data, $p = 0.6$). We similarly tested for differences along the ML axis (data not shown). Again, no consistent influence of recording location on the diagnostic variance could be observed (Pearson correlation coefficients not significantly different from zero; monkey 1, $p = 0.1$; monkey 2, $p = 0.8$; combined data, $p = 0.8$). Furthermore, we investigated whether cases in the lower bank of the superior temporal sulcus (STS) and ventral TE were differently influenced by stimulus diagnosticity. These TE regions have been shown previously to be differently involved in the encoding of three-dimensional objects; they also differ in their connection pattern with other brain regions (Janssen et al., 2000). Here, no differences were found between cases located in the lower bank of the STS or ventral TE. Critically, the dependency of the diagnostic variance on AP position remained the same in both regions. The correlation coefficients computed between diagnostic variance and AP position were not significantly different between the lower bank of the STS and ventral TE, both for the combined data as well as both monkeys individually (monkey 1, $p = 0.7$; monkey 2, $p = 0.6$; combined data, $p = 0.3$). Thus, across the tested region, TE neurons were homogeneously influenced by stimulus diagnosticity.

Spike counts capture local processing as well as long range outputs of neurons in a brain region. However, the LFP is a mass signal that is influenced by currents originating from axons, somata, and dendrites around the electrode (Mitzdorf, 1987; Logothetis, 2002; Logothetis and Wandell, 2004), and thus reflects local neural processes as well as the inputs from other brain regions to the region under study. It has been shown previously that

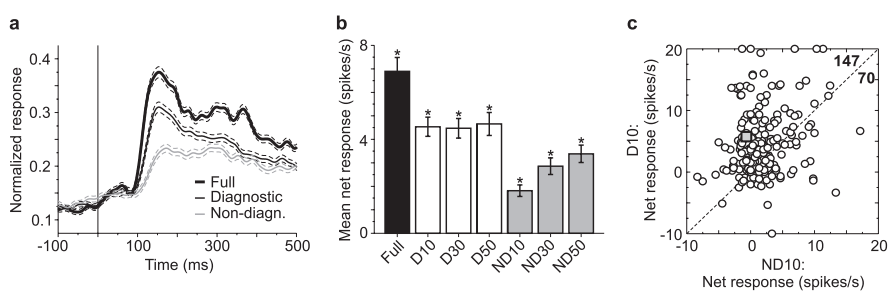


Figure 2. Population response. *a*, Average normalized spike density function for the 220 visually responsive cases. Spike density functions are averaged across the three diagnostic and nondiagnostic conditions. Dashed lines correspond to the SEM. The stimulus onset occurs at time 0 ms. *b*, Mean net firing rate for the complete population. Errors bars show the SEM and asterisks indicate conditions with a mean significantly different from zero (t test, $p < 0.05$). Labels for conditions are as in Figure 1. *c*, Net firing rate in a diagnostic condition (visible stimulus size 10%) versus the net firing rate in the matching nondiagnostic condition. Each point represents one case. A minority of cases had firing rates higher than 20 spikes/s or lower than -10 spikes/s. These cases are plotted overlying the corresponding axis. The square represents the example neuron depicted in Figure 1. The dashed line indicates equal responses in the diagnostic and nondiagnostic condition. The numbers list the cases above and below this line.

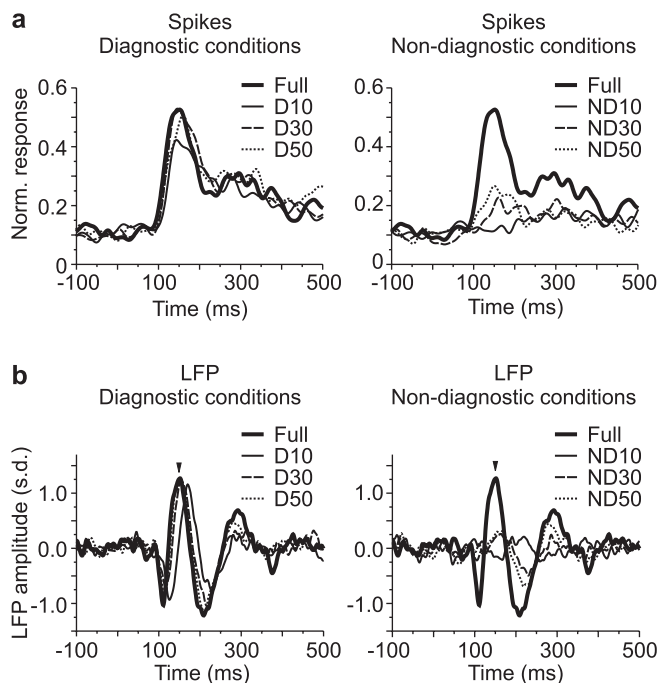


Figure 3. Population average for the high diagnosticity cases. *a*, Average normalized spike density function, computed across the 22 single-unit cases for which the diagnostic variance exceeded the 90th percentile of all cases. *b*, Average VEP for the 23 LFP cases for which the diagnostic variance exceeds the 95th percentile. The VEP component labeled with an arrow is the P140. For spiking activity and LFP, the diagnostic and nondiagnostic conditions are plotted separately; the response to the full condition is repeated in the two plots. Labels indicate the condition (D, diagnostic; ND, nondiagnostic; numbers correspond to the visible stimulus size).

the LFP recorded from individual sites in the IT cortex carries object-selective information (Kreiman et al., 2006). Here, we study the influence of diagnosticity on the LFP. If task-related neural signals can be observed at the level of spiking activity but not LFP, it suggests that these signals are locally computed rather than relayed to the region under study from other brain areas. The relation between spiking activity and LFP can thus provide useful information about the localization of particular computations. We subjected LFP signals recorded concurrently with the spiking activity discussed above to an analysis that was similar to the previous analysis, but took into account the continuous and time-varying nature of the LFP. We first selected LFP sites exhib-

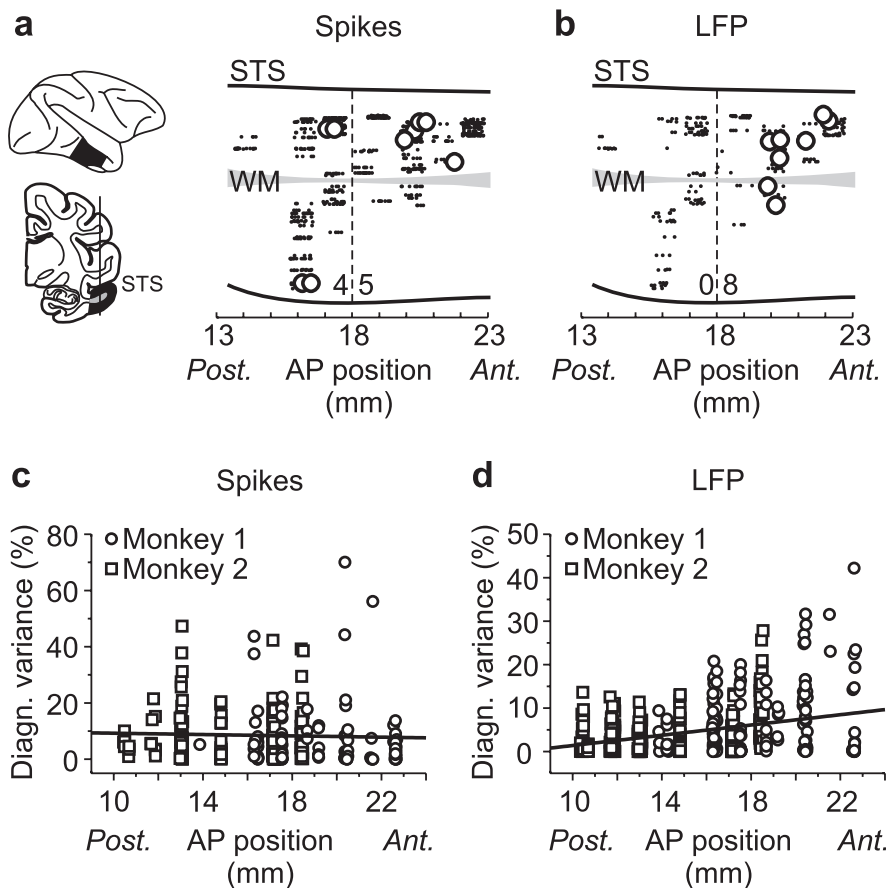


Figure 4. Influence of recording position on the properties of single units and the LFP. *a–b*, Location of high diagnosticity cases in monkey 1, shown on a sagittal view of parts of the temporal lobe. In *a*, the two small brain pictures on the left indicate the location of the selected brain region. This region is indicated in black in the upper image; it is generated by slicing along the line depicted in the lower image. The right side in *a* shows the distribution of single-unit high-diagnosticity cases. *b*, Distribution of the high-diagnosticity LFP cases. In these plots, each dot corresponds to one case recorded from this monkey. Large dots show the location of the diagnosticity cases; small dots show the locations of the rest of the cases. To allow a better separation of different cases, the AP position of each case was randomly jittered by a small amount for display purposes only. The dashed line divides the recording region into a posterior and anterior half with equal extent along the AP axis; numbers list the diagnosticity cases in each half. Thick black lines indicate the location of the STS and the ventral end of the brain in the selected slice, which is shown by the gray region. The position of these landmarks is plotted as estimated during recordings. *c–d*, Diagnostic variance as a function of recording location. In these plots, the diagnostic variance of each case is plotted as a function of its AP position. Symbols indicate the monkey in which a case was recorded; the thick line plots the regression computed between diagnostic variance and AP position. *c*, Single unit data. *d*, LFP data. In all plots, Post and Ant label the posterior and anterior end of the recording region, respectively.

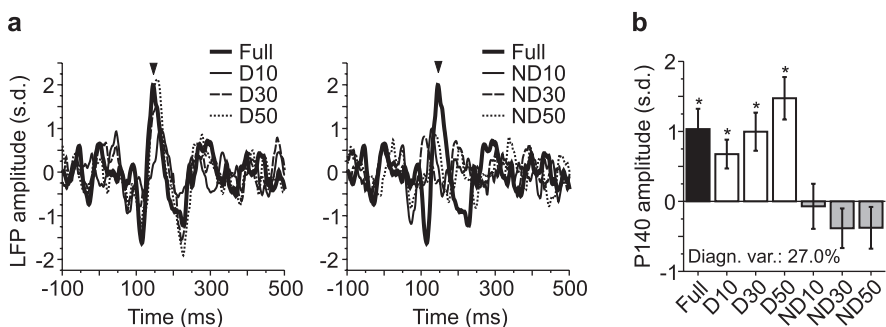


Figure 5. Responses of an exemplar LFP site. *a*, Visual evoked potentials. The P140 is labeled with an arrow. Diagnostic and nondiagnostic conditions are plotted separately; the response to the full condition is repeated in both plots. Conditions are labeled as in Figure 3. *b*, Amplitude of the P140, averaged across trials. Error bars denote the SEM; asterisks indicate the conditions for which a *t* test of the peak amplitude against 0 yielded $p < 0.05$. As a reference, the diagnostic variance for this site is also given.

iting a visual response to at least one stimulus, as indicated by the VEPs (Materials and Methods). Responses to different scenes were again treated as separate cases, and a total of 458 cases from 214 LFP sites were analyzed further.

The grand average VEP of these cases showed three prominent peaks in the time interval from 100 to 200 ms after stimulus onset (data not shown). A negative deflection ~ 100 ms after stimulus onset (N100) was followed by a positive peak at ~ 140 ms (P140), and finally, a second negative deflection at 200 ms (N200). To characterize activity at each LFP site, we analyzed the LFP amplitude in a 20 ms bin around the peak time of the most prominent component, which was the P140 (Materials and Methods). An analysis of the N200 amplitudes yielded similar results, whereas the N100 amplitude exhibited no systematic effects (supplemental Fig. 4, available at www.jneurosci.org as supplemental material). To illustrate the behavior of individual LFP sites, the VEP and the P140 amplitude of an exemplar site are plotted in Figure 5.

The P140 amplitude of this LFP site clearly distinguished between diagnostic and nondiagnostic conditions. All diagnostic conditions generated a P140 with an amplitude significantly larger than zero (*t* test, $p = 0.006$, $p = 0.003$, and $p < 0.001$ for the three conditions), as did the full condition ($p = 0.003$). However, none of the nondiagnostic conditions evoked a P140 with an amplitude larger than the baseline level (*t* test against 0, $p = 0.83$, $p = 0.20$, and $p = 0.23$ for the three conditions). The visible stimulus size had no influence on the P140 amplitudes for this site (ANOVA with factors diagnosticity and size: main effect size, $p = 0.61$; main effect diagnosticity, $p < 0.001$; interaction, $p = 0.15$).

Based on the P140 amplitudes, we computed the percentage of trial-by-trial variance in the LFP, which could be explained by influences of diagnosticity. As was the case for spiking activity, LFP cases with high diagnostic variance values responded preferentially to diagnostic conditions. Figure 3*b* plots the average VEP for the 23 LFP cases with high diagnostic variance values (above the 95th percentile). For these cases, the amplitude of the P140 was significantly greater in diagnostic than in nondiagnostic conditions (repeated measures ANOVA with factors diagnosticity and size; significant main effect of diagnosticity, $p < 0.001$, with no interaction with size, $p = 0.39$). In addition to the influences of diagnosticity, the visible stim-

ulus size also had an effect on the P140, as its latency depended on the visible stimulus size. In all diagnostic conditions, the latency of the P140 (computed as the latency of the positive peak between 100 and 200 ms after stimulus onset) was significantly longer than in the full condition (paired *t* tests, $p < 0.001$, $p = 0.001$, and $p = 0.005$ for the three conditions). We did not determine the latencies of the P140 in the nondiagnostic conditions, as the peak amplitudes in these conditions were too small to allow a reliable measurement of latencies. A high diagnostic variance was not only linked to larger P140 amplitudes in the diagnostic conditions for the high diagnosticity cases; a positive correlation between the diagnostic variance and stronger responses to diagnostic conditions was also obtained at the population level, suggesting that in general higher explained variance values were linked to stronger responses in diagnostic conditions (supplemental Fig. 1, available at www.jneurosci.org as supplemental material).

Using the diagnostic variance values, we then mapped across TE how occlusion effects observed in the LFP depended on stimulus diagnosticity. Figure 4*b* shows the locations of the LFP cases strongly influenced by diagnosticity in one monkey. It can be seen that high-diagnosticity LFP cases clearly clustered in the anterior half of the recording location (χ^2 test, $p = 0.008$). The same was the case for the other monkey (χ^2 test, $p = 0.005$). Figure 4*d* plots the diagnostic variance as a function of AP position for the entire population of LFP cases. Across all LFP cases, there was a significant correlation between the AP location of a LFP case and its diagnostic variance. This was the case for each monkey individually, as well as for the combined data from both monkeys (Pearson correlation coefficient, monkey 1, $r = 0.16$, $p = 0.04$; monkey 2, $r = 0.31$, $p < 0.001$; combined data, $r = 0.33$, $p < 0.001$). This indicates that the influence of diagnosticity on LFP responses grew systematically stronger the more anterior in TE the LFP responses were recorded.

Learning effects can modify the responses of TE neurons during individual recording sessions (Messinger et al., 2001). The observed gradient in the LFP responses could thus have been generated because of a systematic sampling of the recording locations. Recording from posterior locations in the initial sessions of the experiment, and from anterior locations in the final sessions could have made a learning effect appear as a spatial gradient. Whereas the recording location was slowly moved from posterior to anterior locations across the different sessions for one monkey, we used the opposite direction for the other monkey. Because we find the same gradients for both monkeys, learning-dependent changes occurring during the recording sessions cannot account for the spatial gradient in LFP responses.

As for the single units, we tested whether this relationship was similarly present in the lower bank of the STS and ventral TE, and computed correlation coefficients separately for cases located in these two regions. The correlation coefficients were not significantly different between the two regions (monkey 1, $p = 0.7$; monkey 2, $p = 0.4$; combined data, $p = 1$). We also tested whether the position of an LFP case along the ML axis had an influence on the diagnostic variance of the case. Across the whole population of LFP cases, we observed a significant correlation between ML location and diagnostic variance ($r = 0.26$; $p < 0.001$). However, this effect was caused by a strong correlation observed for monkey 2 ($r = 0.27$; $p < 0.001$). In monkey 1, there was no significant correlation between ML position and diagnostic variance ($r = -0.04$; $p = 0.6$).

Discussion

We have experimentally determined which parts of natural scenes monkeys are using in a visual task. Occlusion of these diagnostic parts had a larger influence on neural responses in area TE of macaque cortex than occlusion of nondiagnostic parts, at the level of individual single neurons as well as at the level of local populations as measured by the LFP. This suggests that not all aspects of learned stimuli are encoded equally, but instead that those parts are preferentially represented which are diagnostic for the behaviors associated with these stimuli. Thus, we find signatures of a scene encoding in IT which is based on the diagnosticity of scene parts. In our case, monkeys learned to perform a particular saccadic eye movement associated with each member of a small set of natural images. In learning this task, each of the monkeys came to rely on particular features in each image and these behaviorally relevant or diagnostic features were preferentially encoded in the IT cortex. This lends support to the notion that the neural representation of objects in IT may be not be fixed but instead strongly influenced by the visual experience and viewing history of each observer.

Extensive studies of how single-cell responses in area TE to whole objects can be understood in terms of the responses to object parts have been performed in anesthetized monkeys (Tanaka et al., 1991; Tanaka, 2003). In these studies, experimenters used a reductive determination procedure to identify optimal features for each neuron under study. Our approach is different in that we rely on the monkeys' performance to systematically determine for each stimulus the parts that allow correct recognition behavior. Our results show that this behavioral relevance or diagnosticity is a major determining factor of how learned stimuli are encoded in memory. There has also been published work with behaving monkeys that more indirectly speaks to the issue of parts based representation. For example, TE neurons tended to show systematic tuning for dimensions in parametrically defined line-drawing stimuli that were important for the performance in a categorization task, but not for unimportant features (Sigala and Logothetis, 2002; Sigala et al., 2002). Similarly, TE neurons selectively represented feature conjunctions in visual stimuli composed of two parts, when these were relevant for correct task performance (Baker et al., 2002). In these studies, behaviorally relevant parts exerted a greater influence on neural responses consistent with our findings, but responses to whole stimuli were never directly compared with the responses to the parts alone.

We have observed effects of diagnostic parts-based encoding not only at the level of spiking activity, but also at the level of local populations as measured by the LFP. The LFP is a mass signal that originates from current flow in dendrites and somata in neural populations near the tip of the electrode. It is estimated that between 60 and 70% of excitatory connections of a given pyramidal cell remain local and only between 30 and 40% project to other cortical area (Braitenberg and Schüz, 1998; Binzegger et al., 2004). The LFP thus provides a combined measure of local processing, as well as the inputs from other brain regions. Spiking activity, however, can be considered to provide a combined measure of local processing and outputs to connected target areas. Our results reveal that diagnosticity-related spiking activity was found evenly along the posterior to anterior progression of area TE. Critically, diagnostic LFP activity was only observed in the anterior part of the recording area. To our knowledge this is the first such dissociation between spiking and LFP activity as a function of anatomical recording area. This finding has two major consequences. In posterior TE, diagnosticity was represented

only at single cell but not LFP level. Thus, absence of task-relevant signals in the LFP of a given brain region does not necessarily imply that no single neuron in that region shows such task-relevant signals. Posterior TE regions project strongly to the anterior TE, where we did find evidence of diagnostic parts-based encoding in both LFP and single cell activity. Our findings are consistent with the idea that diagnosticity is first represented by select populations of neurons in the posterior TE, and then transmitted to the anterior TE. We suggest that observation of task-relevant effects at the LFP level does not necessarily imply that the associated functions are performed in the region under study, but rather that they may be computed in brain areas that project to the region under study. Because the LFP is closely related to EEG signals recorded in human subjects, this has profound effects on the interpretation of related findings in humans. For example, EEG studies in humans have been used to link activity in the human lateral occipital complex to the perception of coherent objects from their isolated parts (Doniger et al., 2000; Murray et al., 2004). Our work suggests that this may underestimate the size of the computational network underlying this function. Thus, in general, brain areas where correlates of cognitive functions are actually computed should show effects at the single unit but not LFP level, whereas regions where this information is dynamically routed should show effects at both the single cell and LFP levels. Combined single-cell and LFP recordings thus provide more information than each kind of signal alone, and analysis of the relationship between these signals can provide a novel and powerful method for mapping structure to function in the brain.

References

- Baker CI, Behrmann M, Olson CR (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat Neurosci* 5:1210–1216.
- Baylis GC, Rolls ET, Leonard CM (1987) Functional subdivisions of the temporal lobe neocortex. *J Neurosci* 7:330–342.
- Biederman I (1987) Recognition-by-components: a theory of human image understanding. *Psychol Rev* 94:115–147.
- Binzegger T, Douglas RJ, Martin KA (2004) A quantitative map of the circuit of cat primary visual cortex. *J Neurosci* 24:8441–8453.
- Bortz J (1993) Statistik für sozialwissenschaftler. Berlin: Springer.
- Braitenberg V, Schüz A (1998) Cortex: statistics and geometry of neuronal connectivity. Heidelberg: Springer.
- Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4:2051–2062.
- Doniger GM, Foxe JJ, Murray MM, Higgins BA, Snodgrass JG, Schroeder CE, Javitt DC (2000) Activation timecourse of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. *J Cog Neurosci* 12:615–621.
- Gosselin F, Schyns PG (2001) Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Res* 41:2261–2271.
- Hasselmo ME, Rolls ET, Baylis GC (1989) The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behav Brain Res* 32:203–218.
- Iwai E, Yukie M (1987) Amygdalofugal and amygdalopetal connections with modality-specific visual cortical areas in macaques (*Macaca fuscata*, *M. mulatta*, and *M. fascicularis*). *J Comp Neurol* 261:362–387.
- Janssen P, Vogels R, Orban GA (2000) Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science* 288:2054–2056.
- Kovács G, Vogels R, Orban GA (1995) Selectivity of macaque inferior temporal neurons for partially occluded shapes. *J Neurosci* 15:1984–1997.
- Kreiman G, Hung CP, Kraskov A, Quiroga RQ, Poggio T, DiCarlo JJ (2006) Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex. *Neuron* 49:433–445.
- Logothetis NK (2002) The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos Trans R Soc Lond B Biol Sci* 357:1003–1037.
- Logothetis NK, Sheinberg DL (1996) Visual object recognition. *Annu Rev Neurosci* 19:577–621.
- Logothetis NK, Wandell BA (2004) Interpreting the BOLD signal. *Annual Rev Physiol* 66:735–769.
- Messinger A, Squire LR, Zola SM, Albright TD (2001) Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proc Natl Acad Sci USA* 98:12239–12244.
- Mitzdorf U (1987) Properties of the evoked potential generators: Current source-density analysis of visually evoked potentials in the cat cortex. *Int J Neurosci* 33:33–59.
- Murray MM, Foxe DM, Javitt DC, Foxe JJ (2004) Setting boundaries: brain dynamics of modal and amodal illusory shape completion in humans. *J Neurosci* 24:6898–6903.
- Nielsen KJ, Logothetis NK, Rainer G (2006) Discrimination strategies of humans and rhesus monkeys for complex visual displays. *Curr Biol* 16:814–820.
- Perrett DI, Rolls ET, Caan W (1982) Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res* 47:329–342.
- Perrett DI, Oram MW, Harries MH, Bevan R, Hietanen JK, Benson PJ, Thomas S (1991) Viewer-centered and object-centered coding of heads in the macaque temporal cortex. *Exp Brain Res* 86:159–173.
- Perrett DI, Hietanen JK, Oram MW, Benson PJ (1992) Organization and functions of cells responsive to faces in the temporal cortex. *Philos Trans R Soc Lond B Biol Sci* 335:23–30.
- Sakai K, Miyashita Y (1991) Neural organization for the long-term memory of paired associates. *Nature* 354:152–155.
- Schiller PH, Koerner F (1971) Discharge characteristics of single units in superior colliculus of the alert rhesus monkey. *J Neurophys* 34:920–936.
- Seltzer B, Pandya DN (1978) Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res* 149:1–24.
- Sigala N (2004) Visual categorization and the inferior temporal cortex. *Behav Brain Res* 149:1–7.
- Sigala N, Logothetis NK (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415:318–320.
- Sigala N, Gabbiani F, Logothetis NK (2002) Visual categorization and object representation in monkeys and humans. *J Cog Neurosci* 14:187–198.
- Tamura H, Tanaka K (2001) Visual response properties of cells in the ventral and dorsal parts of the macaque inferotemporal cortex. *Cereb Cortex* 11:384–399.
- Tanaka K (1996) Inferotemporal cortex and object vision. *Annu Rev Neurosci* 19:109–139.
- Tanaka K (2003) Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. *Cereb Cortex* 13:90–99.
- Tanaka K, Saito HA, Fukuda Y, Moriya M (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J Neurophys* 66:170–189.
- Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M (2001) Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nat Neurosci* 4:832–838.
- Yukie M, Takeuchi H, Hasegawa Y, Iwai E (1990) Differential connectivity of inferotemporal area TE with the amygdala and the hippocampus in the monkey. In: *Vision, memory, and the temporal lobe* (Iwai E, Mishkin M, eds), pp 129–135. New York: Elsevier.