

The Neurophysiology of Backward Visual Masking: Information Analysis

Edmund T. Rolls, Martin J. Tovée, and Stefano Panzeri

University of Oxford

Abstract

■ Backward masking can potentially provide evidence of the time needed for visual processing, a fundamental constraint that must be incorporated into computational models of vision. Although backward masking has been extensively used psychophysically, there is little direct evidence for the effects of visual masking on neuronal responses. To investigate the effects of a backward masking paradigm on the responses of neurons in the temporal visual cortex, we have shown that the response of the neurons is interrupted by the mask. Under conditions when humans can just identify the stimulus, with stimulus onset asynchronies (SOA) of 20 msec, neurons in macaques respond to their best stimulus for approximately 30 msec. We now quantify the information that is available from the responses of single neurons under backward masking conditions when two to six faces were shown. We show that the information available is greatly decreased as the mask is brought closer

to the stimulus. The decrease is more marked than the decrease in firing rate because it is the selective part of the firing that is especially attenuated by the mask, not the spontaneous firing, and also because the neuronal response is more variable at short SOAs. However, even at the shortest SOA of 20 msec, the information available is on average 0.1 bits. This compares to 0.3 bits with only the 16-msec target stimulus shown and a typical value for such neurons of 0.4 to 0.5 bits with a 500-msec stimulus. The results thus show that considerable information is available from neuronal responses even under backward masking conditions that allow the neurons to have their main response in 30 msec. This provides evidence for how rapid the processing of visual information is in a cortical area and provides a fundamental constraint for understanding how cortical information processing operates. ■

INTRODUCTION

Our visual environment is constantly changing. To interact with it in real time we need to rapidly process and interpret visual stimuli. How fast our visual system can do this, and the amount of time needed at each synapse for the computations it performs, is of fundamental importance for understanding cortical function (Golledge, Hilgetag, & Tovée, 1996; Wallis & Rolls, 1997). We have previously investigated this vital question using a visual backward masking paradigm. In this paradigm there is a brief presentation of a test stimulus, which after a brief interval (the stimulus onset asynchrony, or SOA) is rapidly followed by the presentation of a second stimulus (the mask), which impairs or masks the perception of the test stimulus. Although backward masking has been extensively used psychophysically (Humphreys & Bruce, 1989), there is little direct evidence of the effects of visual masking on neuronal responses in cortical areas (cf. Schiller, 1968 for lateral geniculate nucleus). In two of the studies that have been performed (Rolls & Tovée, 1994; Rolls, Tovée, Purcell, Stewart, & Azzopardi, 1994) it was found that under short SOA conditions, visual neurons in the primate temporal cortex fire for a very short time. In particular, with SOAs of 20 msec it was found

that the face selective neurons responded for only 20 to 30 msec. This suggests that presentation of the mask interrupts the processing of the test stimulus. The neurophysiological data can be compared directly with the effects of backward masking in human observers studied in the same apparatus with the same stimuli. For the human observers, identification of which face from a set of six had been seen was 50% correct with an SOA of 20 msec, and 97% correct with an SOA of 40 msec (corrected for guessing) (Rolls et al., 1994). Comparing the human performance and the macaque neuronal responses under the same stimulus conditions leads to the conclusion that, when it is just possible to identify which face has been seen, neurons in a given cortical area may be responding for only 20 to 30 msec. A subsequent study by Kovacs, Vogels, and Orban (1995) using a similar backward masking paradigm combined with primate electrophysiology confirmed our results.

The results of the masking experiments were consistent with previous work that suggests that very little time is required at each processing area for object recognition (see Tovée, 1994). The techniques of information theory have been used to analyze the responses of visual neurons in the temporal visual cortex of awake, behaving macaques (Tovée, Rolls, Treves, & Bellis, 1993;

Tovée & Rolls, 1995). Up to 64.9% of the information available in a 400-msec period is available in a 20-msec sample near the start of the spike train, and up to 87% is available in a 50-msec sample. The response latencies in different cortical areas also suggest a processing duration of 10 to 20 msec at each area both in primates (Raiguel, Lagae, Gulyas, & Orban, 1989; Perrett, Hietanen, Oram, & Benson, 1992; Vogel & Orban, 1994) and in cats (Dinse & Kruger, 1994), a figure consistent with a recent visual evoked potential study in humans (Thorpe, Fize, & Marlot, 1996). Moreover, we have also recently shown that visual learning can occur very rapidly. The simple prepresentation of a stimulus can significantly alter the response of a cell to the stimulus when it is shown in an ambiguous situation, just seconds later, suggesting the cell has “learned” to recognize the stimulus under the ambiguous conditions (Tovée, Rolls, & Ramachandran, 1996).

The masking experiment we described (Rolls & Tovée 1994; Rolls et al., 1994) measured the firing rate to an effective stimulus, or the average firing rate to all the stimuli, as a function of the SOA. However, that analysis did not take into account how the SOA affected the responses of the cell not just to the best stimulus, or the average, but also to the less effective stimuli for a cell. It is possible that at short SOAs neurons do not actually discriminate between effective and less effective stimuli as well as they do with long and unmasked stimulus presentation. One way to address this is to measure how SOA influences the difference in firing rate to the best and the least effective stimulus in the set of stimuli. Kovacs et al. (1995) showed that this difference measure was greatly decreased when short SOAs are used. What is still to be answered is if, at short SOAs, the neuronal responses can actually quantitatively discriminate between different stimuli and if the differences between neuronal responses to different stimuli are still significant with respect to the variability of the responses and the spontaneous activity. The proper way to analyze this is to use an information theoretic measure, which will reflect at the same time how different the firing rates are to different stimuli, and how significant the differences in the responses are taking into account the noise (see Optican & Richmond, 1987; Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1996; Rolls & Treves, 1998; Tovée et al., 1993). In this paper we use an information theoretic approach to measure how much information is actually provided by the neurons at different SOAs. The results show that at short SOAs rather less information is available than might be predicted from the average firing to all stimuli, because at short SOAs the responses to the different stimuli become much smaller, but some spontaneous firing remains.

RESULTS

Using the masking protocol we recorded from a total of 34 cells in the inferior temporal cortex and the cortex in the banks and floor of the superior temporal sulcus. It was possible to complete the information analyses on the subset of 15 of these cells for which sufficient trials were available for this type of information-theoretic analysis (see Methods for details). None of the cells responded to the mask stimulus itself.

Figure 1 provides examples of the responses of one cell to the most effective and to the least effective stimulus without a mask, and with a mask at an SOA of 20 msec. It is notable that in the unmasked condition (indicated by an “SOA” of 1000 msec) the cell fired for approximately 120 msec to the effective stimulus, even though the stimulus lasted only for 20 msec. (A more typical value is 200 to 300 msec, as will be shown in Figure 3.) It is also shown that the mask effectively interrupts the firing of the neuron. It is data of this type that were the subject of the information analysis described here. The latency of the neuronal response for this and the other neurons described here was in the region of 75 to 85 msec.

As a preliminary to the analysis, we show in Figure 2 the effect of the SOA on the neuronal responses, averaged across the population of 15 neurons. The responses for the most (max) and the least (min) effective stimuli are shown for the period 0 to 200 msec with respect to stimulus onset. Although the main effect of the mask on the number of spikes emitted was evident at shorter SOAs, and was smaller than for the larger set of neurons analyzed previously (Rolls & Tovée, 1994), there was a statistically significant reduction of the most effective firing rate due to the mask (as tested by a one-way analysis of variance, or ANOVA) at $p < 0.05$. There was little effect (not significant) of the mask on the responses to the least effective stimulus in the set, for which the number of spikes was close to the spontaneous activity. (For Figure 2 and the ANOVA, the responses of different cells were combined by scaling them so that the average response, across SOAs, of each cell to the most effective stimulus was that of the average response to the most effective stimulus of the population of cells across the different SOAs. We used this normalization because it removes the variability that is just due to the inhomogeneity of the cell sample, without changing the dependence of the responses on the SOAs.)

Also as a preliminary to the analysis, we show in Figure 3 the mean firing rate computed in 50-msec epochs for the 15 cells to the most effective stimulus, and to the least effective stimulus, as a function of the SOA. We also show the average response to all stimuli. It can be seen that the average firing rate across stimuli does reduce to some extent as the SOA becomes shorter but that there is a much larger decrease in the difference

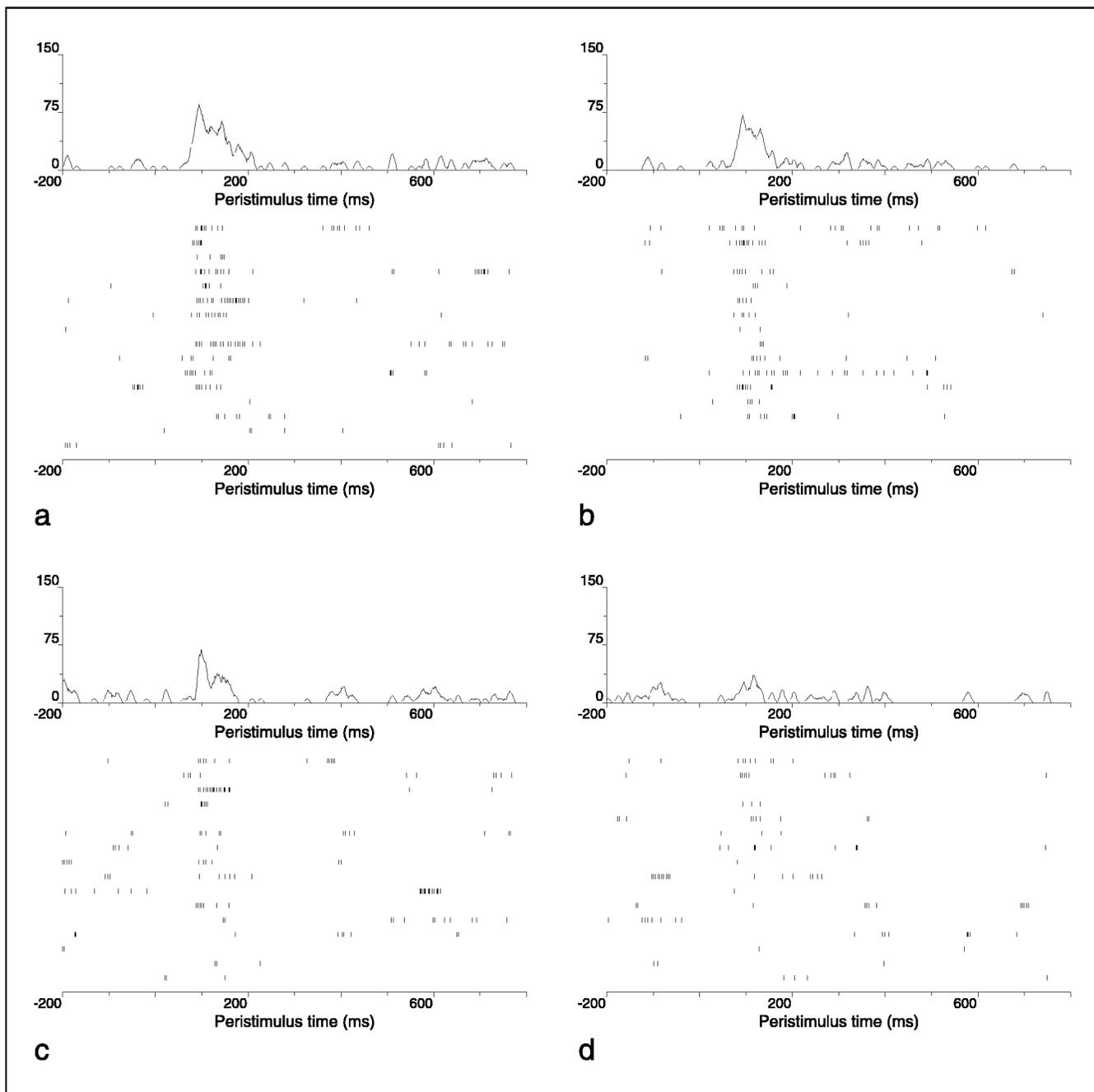


Figure 1. Examples of the responses of one cell to the most effective (a) and to the least effective stimulus (b) without a mask and with a mask at an SOA of 20 msec (c vs. d). The responses are shown in rastergram and peristimulus time histogram form. The stimulus appeared at time 0.

in the responses to the best and the least effective stimuli. It is also clear that the responses, which are all to a 16-msec stimulus, last for longer in the unmasked condition (“SOA” = 1000 msec) than at an SOA of 20 msec.

As a further preliminary to the information analysis, for which the information shown will be that available by different poststimulus times, we show in Figure 4 the average across the cells of the number of spikes cumulated by different poststimulus times for the most effective

and the least effective stimuli for each cell, and also the average across stimuli. It can be seen that the number of spikes continues to increase a little more for the unmasked than for the masked conditions, but that at short SOAs (e.g., 20 msec), the cumulated number of spikes that separates the most effective from the least effective stimulus becomes quite small. This is an indication that at short SOAs the information available about which stimulus was shown might be quite small relative to that with longer SOAs or no masking stimulus.

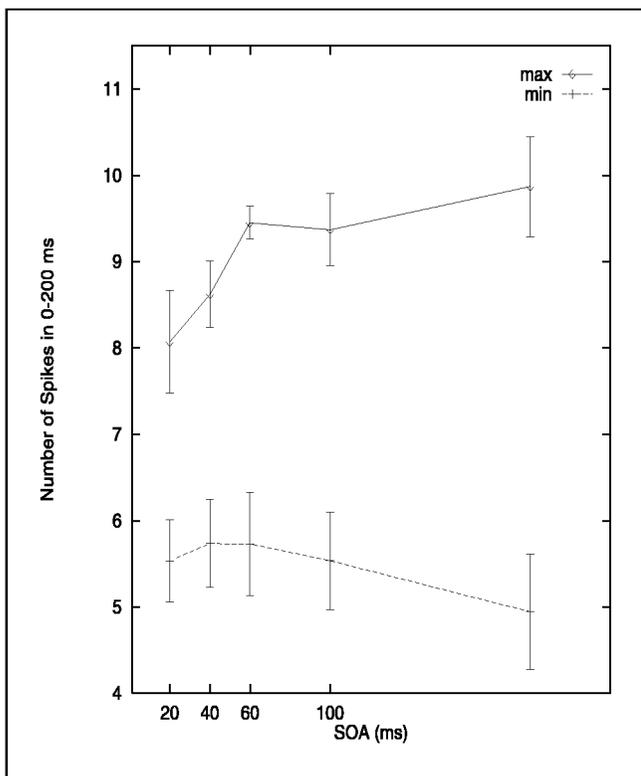


Figure 2. The mean (\pm s.e.m.) across cells of the number of spikes produced by the most effective stimulus (max) and the least effective stimulus (min) as a function of SOA.

Figure 5 shows the cumulated information (averaged across stimuli) at different poststimulus times for the different SOAs and for the unmasked condition. It is clear that there is considerable information in the unmasked condition, with an average value reached typically by 150 msec poststimulus of 0.30 bits. In the masked condition, the information is progressively reduced as the SOA is decreased to 20 msec. At an SOA of 40 msec, the average information was 0.16 bits (53%), and at an SOA of 20 msec, the average information was 0.10 bits (33%). Thus the mask produces a very considerable reduction in the information available about which stimulus was shown. The information tends to decrease after approximately 250 msec, indicating that after this time the responses tend to introduce noise and no net further information about which stimulus was shown. This issue is addressed specifically in Figure 6, which shows the information available in 50-msec epochs taken at different poststimulus times with different SOAs. It is clear that most information is available in 50-msec epochs taken up to 200 to 250 msec. At longer poststimulus times, there is still some information in the unmasked condition, but little in the masked conditions. The reason for the small decrease in information shown in Figure 5 at poststimulus times longer than approximately 250 msec can overall be accounted for by the fact that the large difference in the responses to the different stimuli evident by 250 msec tend to become proportionally

smaller as more spikes in the longer poststimulus times are included. Any firing after 250 msec poststimulus is effectively spontaneous activity of the neuron and there contributes only noise about the stimulus.

To show that the results in Figures 5 and 6 are not due to averaging across cells, we show in Figure 7 the information available from the responses of one cell. Information is encoded in the cell's response for several hundred milliseconds after the cell starts responding to the stimulus, rising to a peak of 0.85 bits at 100 to 150 msec after the onset of stimulus presentation. The cell continues to encode a significant amount of information about the stimulus in 50-msec epochs taken up to 400 msec.

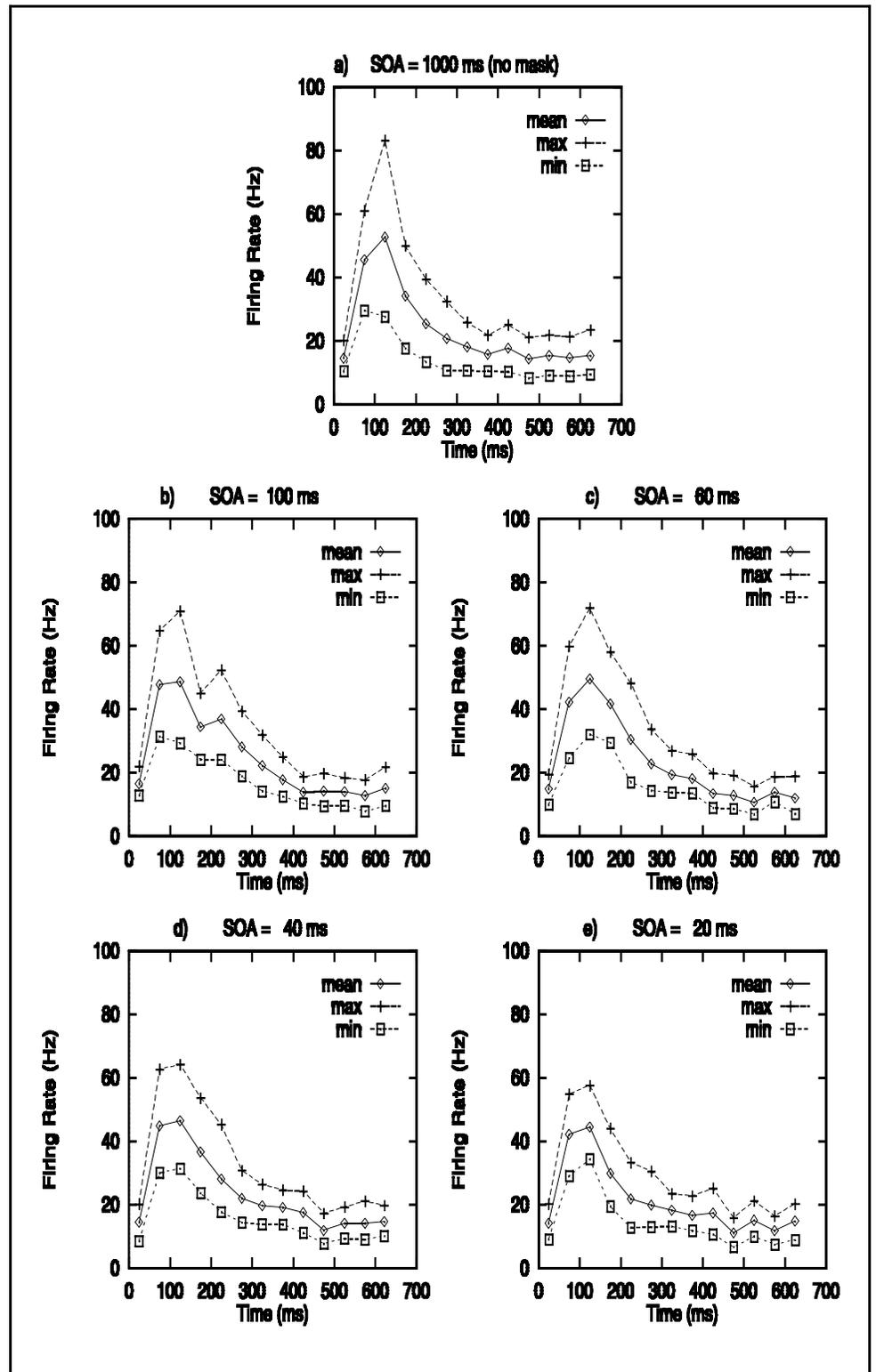
To summarize the results, we show in Figure 8a the average across the cells of the cumulated information available in a 200-msec period from stimulus onset from the responses of the cells as a function of the SOA. This emphasizes how as the SOA is reduced toward 20 msec the information does reduce rapidly, but that nevertheless at an SOA of 20 msec there is still considerable information about which stimulus was shown. The reduction of the information at different SOAs was highly significant (one-way ANOVA) at $p < 0.001$. (The ANOVA was performed after rescaling the data so that each cell had the same average information and so that variance due to differences in the magnitude of the information carried by each cell would not interfere with the test of whether the SOA affected the information.) For comparison, we show in Figure 8b the difference in the number of spikes to the most effective and the least effective stimulus as a function of the SOA (normalized as described for Figure 2). This response difference is closely related to the information available to the different stimuli. Indeed, the difference in the number of spikes is much more closely related to the information than is the mean firing rate of the neurons (shown in Figure 4), emphasizing that it is *differences* of neuronal responses to different stimuli that convey information. The reduction of the difference in the number of spikes at different SOAs was again significant ($p < 0.05$), although less significant than the reduction of the information.

DISCUSSION

The advance described in this paper is that by combining neurophysiology and information theory we have been able to quantitatively measure how backward visual masking affects the information available from the responses of neurons in the visual system. The effect of the mask is to reduce the total information available from the cell, the size of the information peak, and the length of time it is signaling information (Figures 5 to 8). The results clearly show that there is a systematic and statistically significant reduction in the amount of information with decreasing SOA (see, e.g., Figure 8).

The results emphasize that very considerable informa-

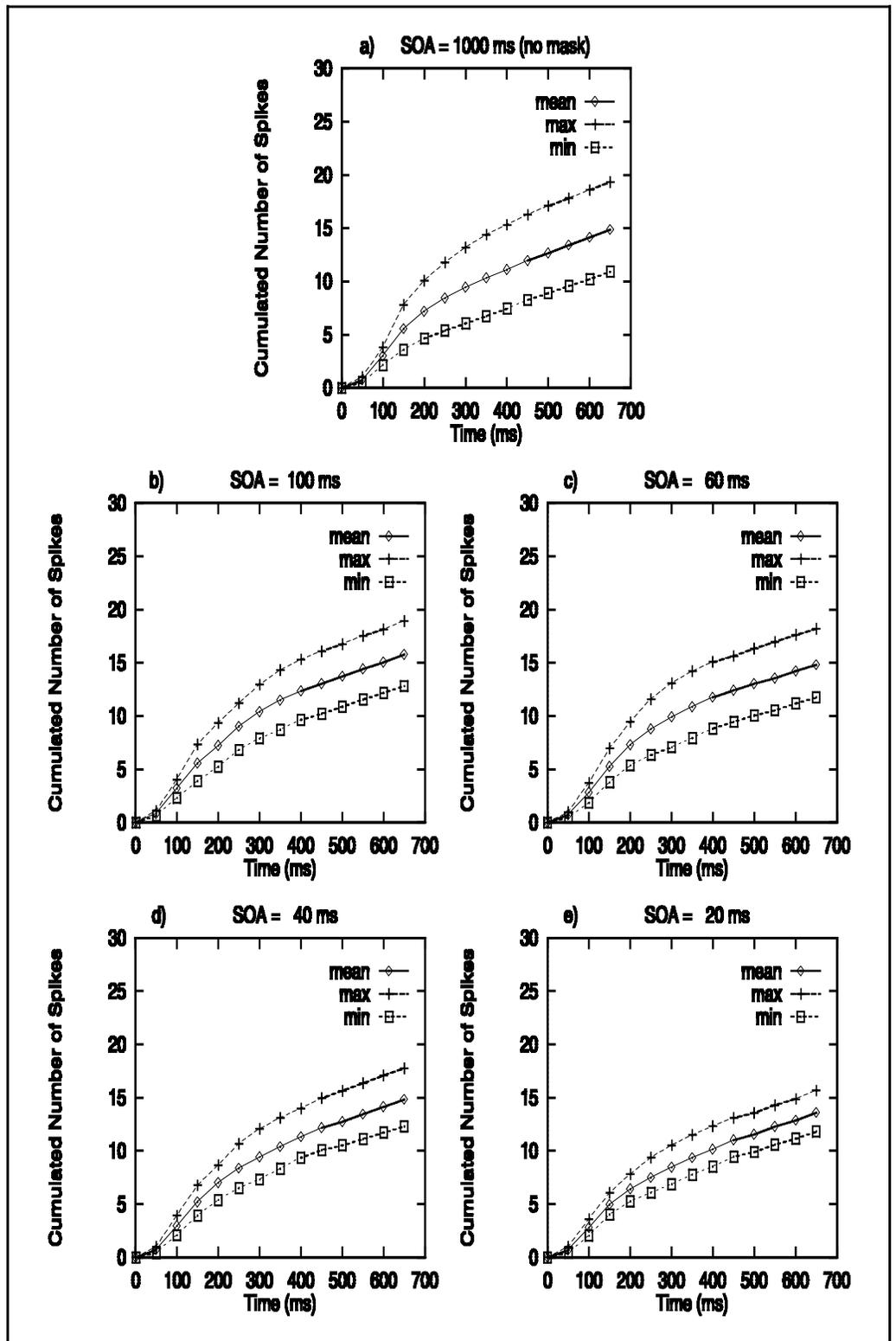
Figure 3. The mean firing rate computed in 50-msec epochs for the 15 cells to the most effective stimulus, and to the least effective stimulus as a function of the SOA. The average response to all stimuli is shown as the middle line of the three.



tion about which stimulus was shown is available in a short epoch (e.g., 50 msec, see Figure 6) of the firing rate. This information is available even when the epoch is taken near the start of the neuronal response. This confirms the findings of Heller, Hertz, Kjaer, and Richmond (1995), Tovée et al. (1993), and Tovée and Rolls (1995), also made with recordings from inferior tempo-

ral cortex neurons. In those studies in which the stimulus lasted for several hundred milliseconds, the epoch could be taken at a wide range of poststimulus times (in the range 100 to 500 msec). In the present study the stimulus itself lasted for only 20 msec, and correspondingly in the no mask condition the information available did drop over the next few hundred milliseconds. How-

Figure 4. The average across the cells of the number of spikes cumulated by different poststimulus times for the most effective and the least effective stimuli for each cell and also the average across stimuli.



ever, it was notable that the information in the no-mask condition did outlast the end of the stimulus by as much as 200 to 300 msec, indicating some short-term memory trace property of the neuronal circuitry (cf. Wallis & Rolls, 1987).

The results show also that an effect of the masking is that the decrease in information is more marked than

the decrease in firing rate, because it is the selective part of the firing that is especially attenuated by the mask, not the spontaneous firing (see Figures 2 and 8). The results also show that even at the shortest SOA of 20 msec, the information available was on average 0.1 bits. This compares to 0.3 bits with the 16-msec stimulus shown without the mask. It also compares to a typical

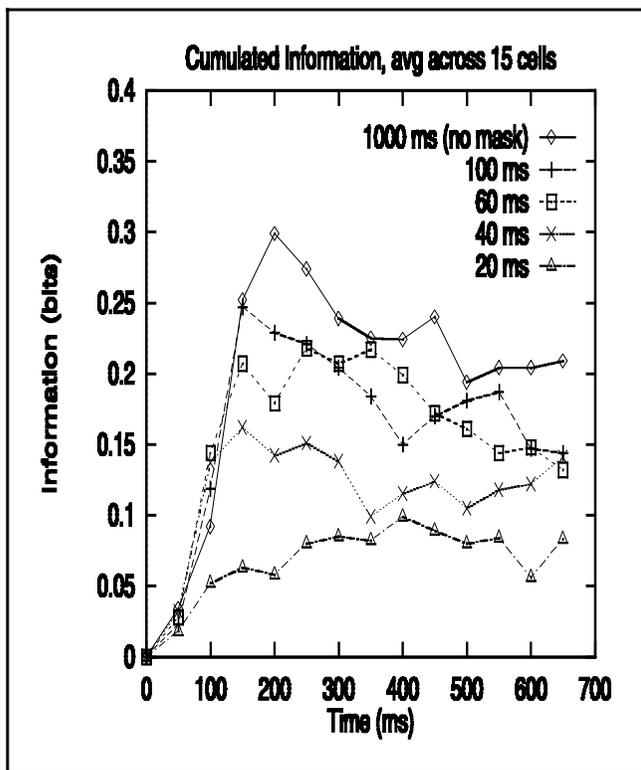


Figure 5. The average across the cells of the cumulated information (averaged across stimuli) at different poststimulus times for the different SOAs and for the unmasked condition.

value for such neurons of 0.35 to 0.5 bits with a 500-msec stimulus presentation (Rolls, Treves, Tovée, & Panzeri, 1997; Tovée & Rolls, 1995). The results thus show that considerable information (33% of that available without a mask and approximately 22% of that with a 500-msec stimulus presentation) is available from neuronal responses even under backward masking conditions that allow the neurons to have their main response in 30 msec. Also, we note that the information available from a 16-msec unmasked stimulus (0.3 bits) is a large proportion (approximately 65 to 75%) of that available from a 500-msec stimulus. These results provide evidence of how rapid the processing of visual information is in a cortical area and provide a fundamental constraint for understanding how cortical information processing operates (see Rolls & Treves, 1998).

The analysis using information theory draws out an interesting point about the relation between the firing rate measure used previously (Rolls & Tovée, 1994; Rolls et al., 1994) and the information measure (see Figure 8). It is shown in Figure 8 that the difference in the mean responses to different stimuli (not the mean response to all stimuli) is related to the information. This reflects the fact that the information is related to the differences in neuronal responses to different stimuli. However, the information also reflects the variance in the neuronal responses, and it is presumably the larger variance at short SOAs that makes the information decrease more

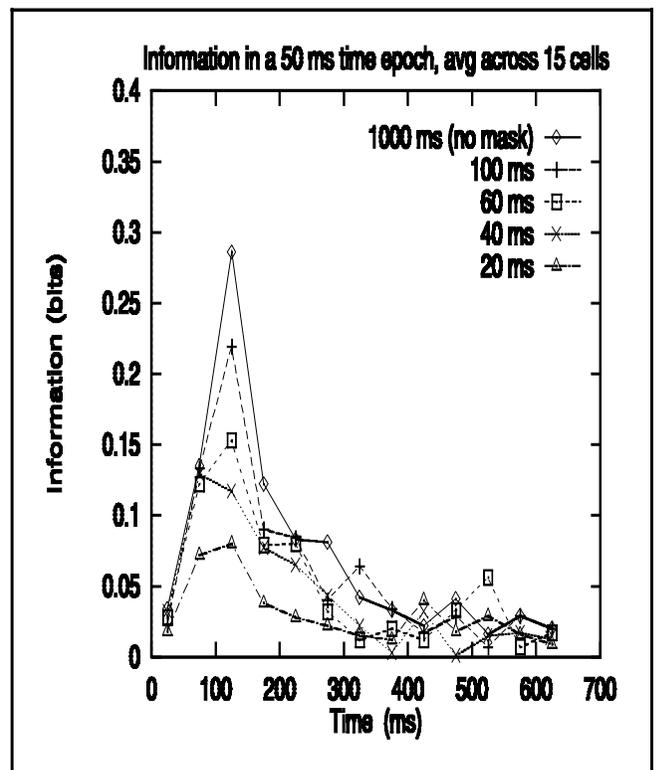


Figure 6. The average across the cells of the information available in discrete 50-msec epochs taken at different poststimulus times with different SOAs and for the unmasked condition.

than the difference in neuronal responses. In fact, the standard deviation of the neuronal responses at an SOA of 20 msec is approximately 0.45 of the mean firing rate, whereas it is 0.4 of the mean firing rate without masking.

The neurophysiological and information data described here can be compared directly with the effects of backward masking in human observers, studied in the same apparatus with the same stimuli (Rolls et al., 1994). For the human observers, identification of which face from a set of six had been seen was 50% correct with an SOA of 20 msec, and 97% correct with an SOA of 40 msec (corrected for guessing) (Rolls et al., 1994.) Comparing the human performance purely with the changes in firing rate under the same stimulus conditions suggested that when it is just possible to identify which face has been seen, neurons in a given cortical area may be responding for only approximately 30 msec (Rolls & Tovée, 1994; Rolls et al., 1994). The implication is that 30 msec is enough time for a neuron to perform sufficient computation to enable its output to be used for identification. The new results based on an analysis of the information encoded in the spike trains at different SOAs support this hypothesis by showing that a significant proportion of information is available in these few spikes (see Figures 5 to 7). Thus there is very rapid processing of visual stimuli in the visual system. We also note that the analysis using information theory adds to previous

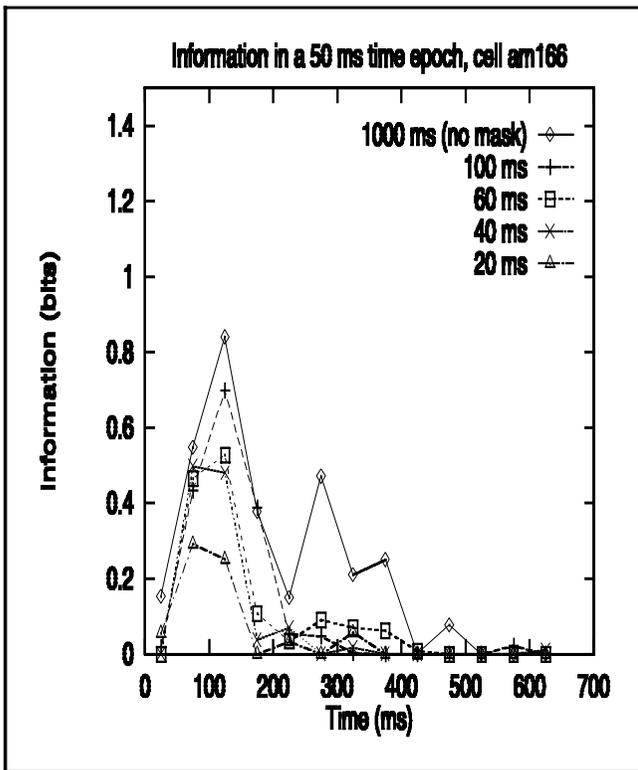


Figure 7. The information available from the responses of one of the cells as a function of the poststimulus time, for different SOAs.

analyses by providing quantitative information measures that can in principle be compared to the performance of human observers using the same information metric.

A comparison of the latencies of the activation of neurons in the different visual cortical areas V1, V2, V4, posterior inferior temporal cortex, and anterior inferior temporal cortex suggests that approximately 10 to 15 msec is added by each stage (Dinse & Kruger, 1994; Oram & Perrett, 1992; Raiguel et al., 1989; Rolls, 1992; Vogel & Orban, 1994). This lag also seems to be common to the passage of information between different subdivisions of a given area. For example, there seems to be a lag of 15 msec between neurons in layer 4C and in the more superficial layers of V1, and a lag of 11 msec between viewer-centered cells and object-centered cells in the temporal visual cortex (Maunsell & Gibson, 1992; Perrett et al., 1992). These studies show that neurons in the next stage of processing start firing soon after (15 msec) the neurons in any stage of processing have started to fire. The fact that considerable information is available in short epochs of, for example, 20 msec, of the firing of neurons provides part of the underlying basis for this rapid sequential activation of connected visual cortical areas (Tovée & Rolls, 1995). The high firing rates of neurons in the visual cortical areas to their most effective stimulus may be an important aspect of this rapid transmission of visual information (cf. Rolls & Treves 1998, section A2.3.2). Nevertheless, the firing of

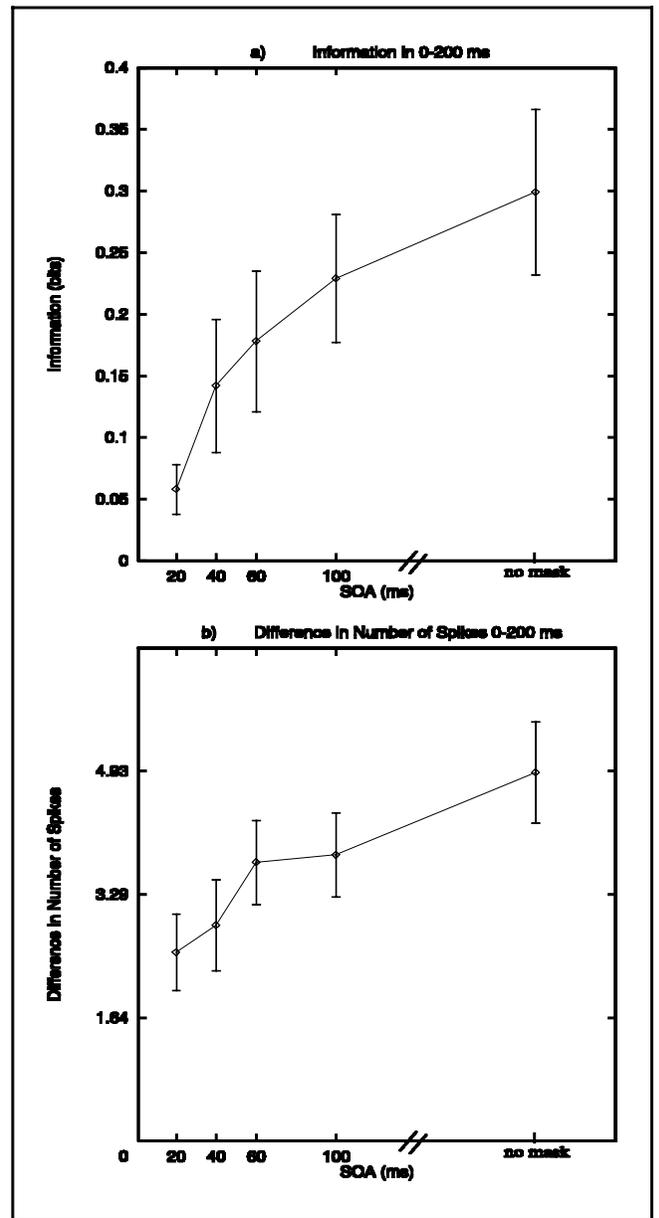


Figure 8. (a) The average (\pm s.e.m.) across the cells of the cumulated information available in a 200-msec period from stimulus onset from the responses of the cells as a function of the SOA. (b) The difference in the number of spikes to the most effective and to the least effective stimulus is also shown as a function of the SOA. The scale in (b) of the “number of spikes” axis is scaled so that the information value (a) and the difference in the number of spikes (b) coincide in the no-mask condition.

neurons within a cortical area normally continues to a visual stimulus for several hundred milliseconds (see Rolls & Treves, 1998). Over this long period of time different factors will influence the firing of neurons in a given cortical area. Initially the firing will be based mainly on incoming information from the preceding cortical area (feed-forward information), but at varying temporal intervals different feed-back mechanisms will play a modulating role. Initially, this is likely to include

lateral inhibition, followed by intracortical feedback from local excitatory recurrent collateral axons of other pyramidal cells and then feedback from higher cortical areas. However, although all these separate inputs seem to be very different conceptually in time, it is of interest that in dynamical systems using integrate-and-fire neurons to model the dynamics of the real brain, the exchange of information required to achieve rapid settling of the network into a final state can be very much faster than might be expected by taking the contribution of each stage as a separate time step (see Rolls & Treves, 1998; Treves, 1993). This is exactly consistent with the rapid speed of processing demonstrated here and emphasized by the fact that considerable information is available from inferior temporal cortex neurons when their processing is interrupted by a mask starting only 20 msec after the onset of a stimulus.

Our previous studies suggested that there is a population of neurons in the temporal cortical areas that respond to a single frame (16 msec) presentation of the target stimulus, a face, with an increased firing rate that can last for 200 to 300 msec (Rolls & Tovée, 1994; Rolls et al., 1994). This is in the unmasked condition. The information analysis confirms and extends this finding. The neuronal responses can encode significant amounts of information for up to 300 msec after the presentation of a 16-msec stimulus in the unmasked condition (see Figure 5). These results suggest that there may be a short-term visual memory implemented by the continued firing of these neurons after a stimulus has disappeared. This short-term visual memory may be implemented by the recurrent collateral connections made between nearby pyramidal cells in the cerebral cortex. These recurrent connections may function in part as an autoassociative memory, which not only enables cortical networks to show continued firing for a few hundred milliseconds after a briefly presented stimulus but also confer some of the response specificity inherent in the responses of cortical neurons (Rolls & Treves, 1998). One effect that may be facilitated by this short-term visual memory is implementation of a trace learning rule in the visual cortex that may be used to learn invariant representations (Rolls, 1992, 1994, 1995; Rolls & Tovée, 1994; Wallis & Rolls, 1997). The continuing neural activity after one stimulus would enable it to be associated with the succeeding images, which given the nature of the statistics of our visual world, would probably be images of the same object. These associations between images produced within a short time by the same object could form the basis for the construction of an invariant representation of that stimulus. For example, the neurons activated by a stimulus are still in a state of activation (and postsynaptic depolarization) perhaps 300 msec later when the same stimulus is seen translated across the retina, viewed from a different angle, or at a different size so that the active axons carrying the transformed

representation can undergo Hebbian associative synaptic modification onto just those neurons that remain in an activated state from the previous input produced by the same object. In this way, the invariant properties evident across short time epochs of the inputs produced by objects may be learned by the visual system (Rolls, 1992; Rolls & Treves, 1998; Wallis & Rolls, 1997).

We finally note that the information per spike, defined as the information calculated in short windows divided by the number of spikes emitted by the cell in that time window, is in the range 0.05 to 0.2 bits per spike for longer SOAs, slightly decreasing down to 0.03 to 0.15 bits per spike at very short SOAs. (The values of the information per spike can be easily extracted by comparing Figure 6, which reports the values of the information available in short time epochs, and Figure 3, which plots the mean firing rates of the neurons in the same time epochs.) These values are similar to the values of the information per spike found in IT neurons responding to long-lasting (e.g., 500 msec) stimuli (Heller et al., 1995; Rolls, Treves, Tovée, & Panzeri, 1997). Thus for the experiments described the information per spike is only moderate, even when there are a very few spikes in response to a stimulus, as in the mask condition with very short SOAs. This result has some implications for the nature of the neuronal code. It has been suggested by Rieke et al. (1996) that the fact that, at short SOAs, only a few spikes are emitted in response to a stimulus, and at the same time a psychophysical discrimination of the visual stimuli is still possible, may imply that even in the mammalian central nervous system, when a single stimulus is rapidly varying or is presented just for a very short time, one or two spikes may be enough to carry very high values of information (thus challenging the idea of a "rate code"). This is indeed the case for, for example, the visual system of the fly (Rieke et al., 1996), where a single spike can carry several bits of information, at least when studied with dynamically varying stimuli. However, the evidence presented here seems to indicate that several spikes from (perhaps different) neurons in the IT cortex of the monkey are needed to provide much information about a briefly presented visual pattern. This is consistent with the distributed representation of information used in many mammalian neural systems and with the advantages that computation with this type of distributed representation of information confers (Rolls & Treves, 1998; Rolls, Treves, & Tovée, 1997).

EXPERIMENTAL PROCEDURES AND DATA ANALYSIS

The activity of single neurons was recorded with glass-insulated tungsten microelectrodes in the anterior part of the superior temporal sulcus (STS) in two alert macaque monkeys (*Macaca mulatta*, mass 3.0 kg) seated in

a primate chair, using techniques that have been described elsewhere (Tovée et al., 1993; Tovée & Rolls, 1992). The preparative procedures were performed aseptically under sodium thiopentone anaesthesia by using pretreatment with ketamine and posttreatment with the analgesic buprenorphine (Temgesic and the antibiotic amoxicillin Cynulox), and all procedures were in accordance with the Policy Regarding the Care and Use of Animals approved by the Society for Neuroscience and were licensed under the U.K. Animals Scientific Procedures Act 1986. Eye position was measured to an accuracy of 0.5° with the search coil technique. A visual fixation task ensured that the monkey looked steadily at the screen throughout the presentation of each stimulus. The task was a blink version of a visual fixation task in which the fixation spot was blinked off 100 msec before the target (otherwise called test) stimulus appeared. The stimuli were static visual stimuli subtending 8 in the visual field presented on a video monitor at a distance of 1.0 m. The fixation spot position was at the center of the screen. The monitor was viewed binocularly, with the whole screen visible to both eyes.

Each trial started at -500 msec with respect to the onset of the test image with a 500-msec warning tone to allow fixation of the fixation point, which appeared at the same time. At -100 msec the fixation spot was blinked off so there was no stimulus on the screen in the 100-msec period immediately preceding the test image. The screen in this period, and at all other times including the interstimulus interval and the interval between the test image and the mask, was set at the mean luminance of the test images and the mask. At 0 msec, the tone was switched off and the test image was switched on for 16 msec. This period was the frame duration of the video framestore with which the images were presented. The image was drawn on the monitor from the top to the bottom in the first 16 msec of the frame period by the framestore, with the remaining 4 msec of the frame period being the vertical blank interval. (The PAL video system was in use.) The monitor had a persistence of less than 3 msec, so no part of the test image was present at the start of the next frame. The time between the start of the test stimulus and the start of the mask stimulus (the SOA) was either 20, 40, 60, 100 or 1000 msec (chosen in a pseudorandom sequence by the computer). The 1000-msec condition was used to measure the response to the test stimulus alone (which was possible because the mask was delayed for so long). The duration of the masking stimulus was 300 msec. At the termination of the masking stimulus the fixation spot reappeared, and then after a random interval in the range 150 to 3350 msec, it dimmed to indicate that licking responses to a tube in front of the mouth would result in the delivery of a reward. The dimming period was 1000 msec, and after this the fixation spot was switched off, and reward availability terminated 500 msec later.

The monkey was required to fixate the fixation spot, and if it licked at any time other than when the spot was dimmed, saline instead of fruit juice was delivered from the tube. If the eyes moved by more than 0.5° from time 0 until the start of the dimming period, the trial was aborted and the data for the trial were rejected. When a trial aborted, a high-frequency tone sounded for 0.5 sec, no reinforcement was available for that trial, and the intertrial interval was lengthened from 8 to 11 sec.

The criterion for the face-selective neurons analyzed in this study was that the response to the optimal stimulus should be at least twice that to the optimal nonface stimulus and that this difference should be significant (Rolls, 1984; Rolls & Tovée, 1995; Rolls, Treves, Tovée, & Panzeri, 1997). If the neuron satisfied the criterion, it was tested with two to six of the effective face stimuli for that neuron. We checked that none of the selected cells had any response to the mask presented alone.

The transmitted information carried by neuronal firing rates about the stimuli was computed with the use of techniques that have been described fully previously (e.g., Rolls, Treves, Tovée, & Panzeri 1997; Rolls & Treves, 1998) and have been used previously to analyze the responses of inferior temporal cortex neurons (Gawne & Richmond, 1993; Optican & Richmond, 1987; Rolls, Treves & Tovée, 1997; Tovée & Rolls, 1995; Tovée et al., 1993). In brief, the general procedure was as follows. The response r of a neuron to the presentation of a particular stimulus s was computed by measuring the firing rate of the neuron in a fixed time window after the stimulus presentation. (In this experiment, the information in a number of different time windows was calculated.) Measured in this way, the firing rate responses take discrete rather than continuous values that consist of the number of spikes in the time window on a particular trial and span a discrete set of responses R across all stimuli and trials. In the experiment, the number of trials available for each stimulus is limited (in the range of 6 to 30 in the present experiment). When calculating the information, the number of firing rate bins that can be used must be smaller than (or equal to) the number of trials available for each stimulus to prevent undersampling and incorrectly high values of calculated information (Rolls & Treves, 1998). We therefore quantized the measured firing rates into a smaller number of bins d . We chose here d in a way specific for every cell and every time window, according to the following: d was the number of trials per stimulus (or the number of different rates that actually occurred if this was lower). This procedure is very effective in minimizing information loss due to overregularization of responses while effectively controlling for finite sampling biases; see Golomb, Hertz, Panzeri, Treves, & Richmond (1997) and Panzeri & Treves (1996). After this response quantization, the experimental joint stimulus-response probability table $P(s, r)$ was computed from the data (where $P(r)$ and

$P(s)$ are the experimental probability of occurrence of responses and of stimuli, respectively), and the information $I(S, R)$ transmitted by the neurons averaged across the stimuli was calculated by using the Shannon formula (Cover & Thomas 1991; Shannon, 1948),

$$I(S, R) = \sum_{s,r} P(s, r) \log_2 \frac{P(s, r)}{P(s)P(r)}$$

and then subtracting the finite sampling correction of Panzeri and Treves (1996) to obtain estimates unbiased for the limited sampling. This leads to the information available in the firing rates about the stimulus. We did not calculate the additional information available from temporal firing patterns within the spike train, because the additional information is low, often only 10 to 20%, and reflects mainly the onset latency of the neuronal response (Tovée et al., 1993; Tovée & Rolls, 1995).

In the experiments described here, three cells were tested with six stimuli, one with four stimuli, five cells with three stimuli, and six cells with two stimuli. We note that the number of stimuli used here is smaller than that used in other experiments that applied information theoretic measures to the responses of inferior temporal cortex cells (Optican & Richmond 1987; Rolls, Treves, & Tovée 1997; Tovée et al. 1993; Tovée & Rolls, 1995). The reason for using fewer images in the experiments described here is that each image needed to be tested with five different masking conditions. There are two potential problems arising from a calculation of information from a limited set of stimuli. The first one is that with few stimuli, the full response space of the neuron may not be adequately sampled. We guarded against this by choosing images for the tests that elicited quite different responses from the cells and ensuring that the information measured with these images was high. The second possible problem is that the information measures may be distorted by the fact that there is a ceiling to the amount of information that can be provided about which of a set of stimuli has been seen, and this entropy is the logarithm of the number of stimuli in the set. For two stimuli, just 1 bit of information is all that could be provided by a cell. This ceiling could limit the information measurement obtained from neuronal responses if too few stimuli are used (Gawne, Kjaer, Hertz, & Richmond, 1996; Gawne & Richmond, 1993; Rolls & Treves, 1998; Rolls, Treves, & Tovée 1997). For the analyses described here, we checked that the information available from the neuronal responses was always well below the ceiling so that the information measure was not distorted. In fact, for the unmasked condition, the average information from the neuronal responses was 0.3 bits, and this is much lower than the entropy of the sets of images, which varied for different cells between 1 bit

and 2.6 bits. Finally, we note that each neuron was tested with the same stimuli across the different masking conditions, and therefore the comparison of the information values obtained under the different masking conditions is homogenous.

Acknowledgments

This research was supported by Medical Research Council Programme Grant PG8513790 to E. T. Rolls. S. Panzeri is supported by an EC Marie Curie Research Training Grant ERBFMBICT972749.

Reprint requests should be sent to E. T. Rolls, Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK, or via e-mail: Edmund.Rolls@psy.ox.ac.uk.

REFERENCES

- Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York: Wiley.
- Dinse, H. R., & Kruger, K. (1994). The timing of processing along the visual pathway in the cat. *NeuroReport*, *5*, 893–897.
- Gawne, T. J., Kjaer, T. W., Hertz, J. A., & Richmond B. J. (1996). Adjacent cortical complex cells share about 20% of their stimulus-related information. *Cerebral Cortex*, *6*, 482–489.
- Gawne, T. J., & Richmond, B. J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? *Journal of Neuroscience*, *13*, 2758–2771.
- Golledge, H. D. R., Hilgetag, C., & Tovée, M. J. (1996). Information processing: A solution to the binding problem? *Current Biology*, *6*, 1092–1095.
- Golomb, D., Hertz, J. A., Panzeri, S., Treves, A., & Richmond, B. J. (1997). How well can we estimate the information carried in neuronal responses from limited samples? *Neural Computation*, *9*, 649–665.
- Heller, J., Hertz, J. A., Kjaer, T. W., & Richmond, B. J. (1995). Information flow and temporal coding in primate pattern vision. *Journal of Computational Neuroscience*, *2*, 175–193.
- Humphreys, G. W., & Bruce, V. (1989). *Visual cognition*. Hove, Eng.: Erlbaum.
- Kovacs, G., Vogels, R., & Orban, G. A. (1995). Cortical correlates of pattern backward-masking. *Proceedings of the National Academy of Sciences, USA*, *92*, 5587–5591.
- Maunsell, J. H. R., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology*, *68*, 1332–1344.
- Optican, L., & Richmond, B. J. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretical analysis. *Journal of Neurophysiology*, *57*, 132–146.
- Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology*, *68*, 70–84.
- Panzeri, S., & Treves, A. (1996). Analytical estimates of limited sampling biases in different information measures. *Network*, *7*, 87–107.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells in the macaque temporal cortex. *Philosophical Transactions of the Royal Society, London, B*, *335*, 23–50.
- Raiguel, S. E., Lagae, L., Gulyas, B., & Orban, G. A. (1989). Re-

- sponse latencies of visual cells in macaque areas V1, V2 and V5. *Brain Research*, 493, 155-159.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R. R., & Bialek, W. (1996). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.
- Rolls, E. T. (1984). Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Human Neurobiology*, 3, 209-222.
- Rolls, E. T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philosophical Transactions of the Royal Society, London, B*, 335, 11-21.
- Rolls, E. T. (1994). Brain mechanisms for invariant visual recognition and learning. *Behavioral Processes*, 33, 113-138.
- Rolls, E. T. (1995). Learning mechanisms in the temporal lobe visual cortex. *Behavioral Brain Research*, 66, 177-185.
- Rolls, E. T., & Tovée, M. J. (1994). Processing speed in the cerebral cortex, and the neurophysiology of backward masking. *Proceedings of the Royal Society, London B*, 257, 9-15.
- Rolls, E. T., & Tovée, M. J. (1995). The sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, 73, 713-726.
- Rolls, E. T., Tovée, M. J., Purcell, D. G., Stewart, A. L., & Azzopardi, P. (1994). The responses of neurons in the temporal cortex of primates and face identification and detection. *Experimental Brain Research*, 101, 473-484.
- Rolls, E. T., & Treves, A. (1998). *Neural networks and brain function*. Oxford: Oxford University Press.
- Rolls, E. T., Treves, A., & Tovée, M. J. (1997). The representational capacity of the distributed encoding of information provided by populations of neurons in the primate temporal visual cortex. *Experimental Brain Research*, 114, 149-162.
- Rolls, E. T., Treves, A., Tovée, M., & Panzeri, S. (1997). Information in the neuronal representation of individual stimuli in the primate temporal visual cortex. *Journal of Computational Neuroscience*, 4, 309-333.
- Schiller, P. H. (1968). Single unit analysis of backward visual masking and metaccontrast in the cat lateral geniculate nucleus. *Vision Research*, 8, 855-866.
- Shannon, C. E. (1948). A mathematical theory of communication. *AT&T Bell Laboratories Technical Journal*, 27, 379-423.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520-522.
- Tovée, M. J. (1994). How fast is the speed of thought? *Current Biology*, 4, 1125-1127.
- Tovée, M. J., & Rolls, E. T. (1992). Oscillatory activity is not evident in the primate temporal visual cortex with static visual stimuli. *NeuroReport*, 3, 369-372.
- Tovée, M. J., & Rolls, E. T. (1995). Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex. *Visual Cognition*, 2, 35-58.
- Tovée, M. J., Rolls, E. T., & Ramachandran, V. S. (1996). Rapid visual learning in the neurons of the primate temporal visual cortex. *NeuroReport*, 7, 2757-2760.
- Tovée, M. J., Rolls, E. T., Treves, A., & Bellis, R. P. (1993). Information encoding and the responses of single neurons in the primate temporal visual cortex. *Journal of Neurophysiology*, 70, 640-654.
- Treves, A. (1993). Mean-field analysis of neuronal spike dynamics. *Network*, 4, 259-284.
- Vogel, R., & Orban, G. A. (1994). Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. *Journal of Neurophysiology*, 71, 1428-1451.
- Wallis, G., & Rolls, E. T. (1997). Invariant face and object recognition in the visual system. *Progress in Neurobiology*, 51, 167-194.