State-Dependent Effects of Stimulus Presentation Duration on the Temporal Dynamics of Neural Responses in the Inferotemporal Cortex of Macaque Monkeys

Koorosh Mirpour1 and Hossein Esteky1,2

1 School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Niavaran, Tehran, Iran;
2 Research Center for Brain and Cognitive Sciences, School of Medicine, Shaheed Beheshti University, Tehran, Iran

Running head:
Brain State and Stimulus Duration Affect IT Cortex Dynamics

Corresponding author with complete address, including an email address:

Hossein Esteky
School of Cognitive Sciences,
Institute for Studies in Fundamental Sciences (IPM),
Niavaran, Tehran, Iran
TEL: (98) 21 2229 4035
Email: ESTEKY@IPM.IR
www.visionlab.ir
Abstract

During natural vision, stimuli are viewed for different durations as the state of brain activity changes over time. Here, we studied the effects of stimulus presentation duration on cell responses (n=259) in three subdivisions of the inferotemporal (IT) cortex of fixating macaque monkeys as neural baseline firing rates varied over the course of recording. First, cell responses to the presentation of 120 images were tested and four images that elicited significant responses with various degrees of effectiveness were selected for further study. Then, the four selected images were presented to the monkeys for five different presentation durations (18, 70, 140, 210 and 350 ms). We found that, depending on the magnitude of neural baseline activity, stimulus presentation duration affected the response properties and efficiency of neural information processing in the IT cortex. Short stimulus presentation durations elicited phasic responses consisting of rhythmic activation and inactivation, which conveyed a lower amount of stimulus information, particularly following higher baseline firing rates. Longer presentation durations elicited a sustained pattern of response and carried a greater amount of information, particularly at lower baseline firing rates. Finally, a significantly higher proportion of cells in the posterior IT compared to the anterior IT had a tendency to have high baseline activity, recruit stronger phasic responses and convey less information. It is plausible that during natural vision, as stimuli with various exposure durations affect the visual system, top-down influence or competition within local neural networks differentially influences the function of IT cells by changing their baseline activity.
INTRODUCTION

The inferior temporal cortex (IT) is the final purely visual area along the ventral visual pathway in primates and is thought to be essential for recognition of complex visual objects (Tanaka et al. 1991; Logothetis 1998; Afraz et al. 2006; Baylis et al. 1987; Desimone et al. 1984; Gross 1994; Kiani et al. 2007). Several studies suggest that information about stimulus shape evolves dynamically over the temporal course of neural responses in the ventral visual pathway. It has been shown that IT cells convey different types and levels of information across the temporal course of their responses (Allred and Jagadeesh 2007; Kiani et al. 2005; Kovacs et al. 1995; Sugase et al. 1999; Tamura and Tanaka 2001; Tovee et al. 1993). Similar results have been reported in the lower areas of the ventral visual pathway (Gawne et al. 1991; Hegde and Van Essen 2004; Hegde and Van Essen 2006; Muller et al. 2001).

Recognition of objects depends on the duration of exposure to the stimuli. Many studies have shown that the mean duration of fixation in free visual search and object recognition tasks is about 200 ms, depending on stimulus properties and task difficulty (Enoch 1959; Ford et al. 1959; Gould 1973; Jacobs 1986; Luria and Strauss 1978; Zelinsky and Sheinberg 1997). For example, it has been shown that fixation durations <70 ms result in missing the fixated target in a visual search task (Hooge and Erkelens 1996). In addition, human visual object recognition performance declines as the stimulus presentation duration decreases (Grill-Spector and Kanwisher 2005). The steepest decline in performance occurs in presentation times <70 ms. The impact of stimulus presentation duration on cortical cell responses has been shown previously in the primary visual cortex of monkeys (Gawne et al. 1991; Muller et al. 2001). In addition, it has been observed that the reliability of the primary visual cortex (V1) and IT neural discrimination between two stimuli increases with increasing stimulus presentation duration (Allred and
Jagadeesh 2007; Zohary et al. 1990). Allred and Jagadeesh showed that the pattern of IT cell responses and the amount of information they convey is affected by the stimulus exposure duration (Allred and Jagadeesh 2007).

However, stimulus presentation duration is only one of many parameters that can affect the temporal dynamics of neural responses. It is not clear how these different parameters interact to govern the way the visual cortex processes sensory information. For example, it has been shown that the baseline firing rate of cortical cells changes spontaneously (Dehaene and Changeux 2005), but the role of fluctuations in baseline activity in cortical sensory processing is not well understood. During natural vision, in addition to these spontaneous changes in baseline activity, contextual parameters such as expectation, attention and familiarity with the viewing condition and/or the effect of preceding or simultaneously viewed stimuli can change the baseline brain activity. Emerging evidence indicates that baseline activity is controlled by top-down mechanisms and that it can affect sensory processing and the subjects’ behavior. For example, higher and more correlated baseline activity in V1 cells is shown to be correlated with higher stimulus detection performance by monkeys (Super et al. 2003). Several other studies have shown that spontaneous and evoked changes of baseline neural activity affect cortical sensory processing in rodents’ barrel cortex and cats’ V1 (Haider et al. 2007; Leger et al. 2005; Petersen et al. 2003; Sachdev et al. 2004; Timofeev et al. 1996). However, little is known about the effect of neural baseline activity on high-level sensory processing in the IT cortex of primates. In particular, there have been no studies to investigate the concurrent impact of stimulus exposure duration and the spontaneous change of the state of neural activity on sensory processing and neural response dynamics.
Here, we studied the effects of stimulus form and presentation duration on IT cell responses as their baseline firing rates fluctuated over the temporal course of the recording. We found that the amount of information conveyed by IT cells depended on presentation durations and the level of the cells’ baseline activities, particularly during the later part of responses. High baseline activity was followed by phasic responses, which consisted of rhythmic activation and inactivation and conveyed a lower amount of stimulus information. At low baseline activities, the neural responses were more sustained and carried a higher amount of information.

MATERIALS AND METHODS

Subjects
Two male adult macaque monkeys (*M. mulatta*) were used in this study. All experimental procedures were performed in accordance with the guidelines on the care and use of laboratory animals of the Iranian Society for Physiology and Pharmacology and the US National Institutes of Health.

Recording
We recorded single-unit responses from the inferior temporal cortex area (IT) in awake monkeys during a passive fixation task, using the same experimental procedures as described previously, (Afraz et al. 2006; Kiani et al. 2005) except when mentioned otherwise. In brief, head restraints and recording chambers were stereotaxically implanted under aseptic conditions on the dorsal surface of the skull of the monkeys while the animals were anesthetized with sodium pentobarbital. After recovery, single units were recorded with tungsten electrodes (FHC, USA). The electrodes were advanced with an Evarts-type manipulator (Narishige, Japan) from the
dorsal surface of the brain through a stainless steel guide tube inserted into the brain down to 10-
15 mm above the recording sites. The recording sites were localized using the MRI images
acquired before the surgery and electrophysiological criteria defining white/gray matter borders.
The recording sites were evenly distributed at the anterior 11–20 mm in one monkey and 7-19
mm in the other one over the ventral bank of the superior temporal sulcus and the ventral
convexity up to the medial bank of the anterior middle temporal sulcus. The cortical sulci were
located using the MRI images. Subdivisions of the IT cortex were defined using the location of
cortical sulci as illustrated by Tanaka and colleagues (Saleem and Tanaka 1996; Saleem et al.
2000; Tamura and Tanaka 2001). All neurons located in the cortical area between the medial
bank of the anterior middle temporal sulcus (amts) and lateral to the lower lip of the superior
temporal sulcus (sts), including the lateral bank of the amts, were considered as the anterior
inferotemporal cortex (TEa). The posterior border of the TEa was defined as the first sagittal
section in which the amts appeared in the MRI images. All neurons located in the cortical area
between the medial bank of the anterior middle temporal sulcus (amts) and lateral to the
occipitotemporal sulcus (ots) were considered as the posterior inferotemporal cortex (TEp). The
posterior and anterior borders of the TEp were defined as anterior to the pmts and posterior to the
amts. Figure 1 represents a schematic that illustrates the location of the subdivisions of the
recorded IT neurons in this study.

The action potentials from single units were isolated in real-time by a template matching
algorithm (Worgotter et al. 1986). The monkey had to fixate the eyes, with a precision of ± 2°,
on a 1° circular fixation spot presented at the center of the display. The eye position was
measured by an infrared system which allowed a precision of 1° in the measurement of eye
position (i_rec, http://staff.aist.go.jp/k.matsuda/eye/).
**Visual stimuli**

The stimuli were 120 photographs of natural and artificial objects isolated on a gray background. The stimulus set consisted of simple geometrical shapes and complex multicolored pictures presented on a computer graphics display. We presented visual stimuli in two stages. In the first stage, 120 stimuli were presented in a pseudo-random order, each repeated five times. It should be noted that we did not use any visual masks following stimulus presentation. The stimuli were colorful photographs of natural and artificial objects isolated on a gray background. They extended to $7 \times 7$ degrees of visual angle. Each block started with appearance of a $1^\circ$ fixation point in the center of the display. When the monkeys maintained their gaze at the fixation point for 200 ms, the fixation point disappeared, and after a 300 ms delay, presentation of the stimulus sequence started. Each stimulus was presented for 210 ms, with a 400 ms interstimulus interval. The sequence stopped when the monkey broke the gaze fixation and a new block started with the reappearance of a fixation point. The monkeys were rewarded with a drop of juice every 1–3 s during the fixation.

To test the effect of stimulus presentation duration, 4 stimuli were selected from the images that evoked statistically significant responses: the stimulus that evoked the largest response (called the best or optimal stimulus hereafter) and 3 stimuli (called sub-optimal stimuli hereafter) that evoked the highest response magnitudes in the following ranges: 50-75%, 25-50% and <25% of the largest response. In all analyses of the results, the effect of stimulus strength was calculated using these 4 different stimuli for each neuron.

The response magnitudes were calculated as the average firing rate within a 430 ms window, starting at 71 ms after the stimulus onset. The significance of the responses to an individual
stimulus was determined by comparing the individual response magnitude with the average firing rate of 300 ms of spontaneous activity before the stimulus onset (p<0.05, t-test).

Spike data from 685 neurons were collected for the first stage of the experiment from the right (monkey Sh) and left (monkey Jn) hemispheres of two monkeys. Among all recorded neurons, 388 cells had stimuli that evoked responses within the range of 50-75%, 25-50% and <25% of the largest response and were used for further study in the second stage of the experiment. In this stage, when the monkeys maintained their gaze at the fixation point for 200 ms, the fixation point disappeared and after a 300 ms delay, the presentation of the stimulus sequence started. Each one of the four selected stimuli was shown to the monkey 15 times for five different stimulus presentation durations: 18, 70, 140, 210 and 350 ms. Stimuli and presentation durations were pseudo-randomly interleaved. The interstimulus interval was 1300 ms from stimulus onset in all conditions. The sequence stopped when the monkey broke the gaze fixation and a new block started with the appearance of a fixation point (Fig. 2). The monkeys were rewarded with a drop of juice every 1–3 s during the fixation.

Stimulus selectivity in the second phase was defined as the significant response (compared to activity 0-50 ms post-stimulus) to the best stimulus for at least four of the five stimulus presentation durations (p<0.05, t-test). All of the cells with significant responses in this stage were used for further data analysis (n = 259).

**Data analysis**

INDICES. In each trial, the early and late phases of the responses were defined as 60-240 ms and 250-550 ms from the stimulus onset, respectively. We used three parameters for comparing the two phases of responses as follows: 1) The response peak magnitude was calculated as the
maximum response value of the normalized and smoothed spike density function during the early or late phase of responses, 2) The peak latency was defined as the time from stimulus onset to the response peak magnitude, and 3) The average firing rate was defined as the mean firing rates of the normalized spike density function during the initial (60-240 ms) and late phases (250-550 ms).

To investigate the effects of stimulus presentation duration on the oscillatory activity of neurons, we calculated the oscillation score in the theta band (4-8 Hz) as proposed by Muresan and colleagues (Muresan et al. 2008). Oscillation scores estimate the degree to which a neuron is oscillating in a given frequency band relying on autocorrelation histograms computed on individual trials of spike trains. For calculating the oscillation score in the theta band, first, autocorrelation histograms (ACH) for individual trials were calculated with a bin size of 1 ms and a time lag of ± 1024 ms. The maximum height of the central peak (lag = 0) solely indicates the firing rate of the neuron, which is a confounding factor for the analysis of the oscillation score. For removing the effect of the central peak in the ACH at lag zero, we had to calculate the left and right boundaries of the central peak. First, we calculated the slow-smoothed ACH by convolving the ACH with a Gaussian kernel with \( \sigma = 10 \) ms. Then, the slope of the slow-smoothed ACH was calculated using the first-order derivative. The left and right boundaries of the central peak were calculated on negative and positive time lags by finding the time lags around lag zero so that the slope of the slow-smoothed ACH was \( \leq 10^\circ \). Since the ACH is a symmetric histogram, the left boundaries were in negative time lags and right boundaries were in positive time lags. Slow-smoothed ACHs were only used for calculating the left and right boundaries of the central peak. For calculation of the oscillation score, we used the fast-smoothed ACH, which is the ACH smoothed by a Gaussian kernel with \( \sigma = 2 \) ms. The central
peak from the left to the right boundary of the fast-smoothed ACH was replaced by the values of
the fast-smoothed ACH at the time lag of the left (or right) margin of the central peak. Then, the
frequency spectrum was computed on the fast-smoothed ACH with the central values removed.
Then, the oscillation score was calculated as the ratio between the peak magnitude in the
frequency band of interest (4-8 Hz) and the mean magnitude of the spectrum. The theta
frequency band, 4-8 Hz, was chosen based on the qualitative observation of histograms of
frequency spectra plotted for individual trials and the population average. Consistent with our
observation, Rollenhagen and Olson have shown that the inferotemporal cortex (IT) of the
macaque monkey responds to visual stimuli by firing action potentials in a series of sharply
defined bursts at a frequency of about 5 Hz (Rollenhagen and Olson 2005). The efficiency of the
oscillation score was tested for the theta band by its creators (for more details about the
calculation and evaluation of oscillation scores see Muresan et al. 2008).

To examine the effect of baseline firing rates on oscillation scores, we calculated simulated
oscillation scores by shuffling the spike times of each trial 1000 times while keeping the firing
rate constant. Then, we calculated the oscillation score for each shuffled trial. The mean of 1000
shuffled oscillation scores was assigned to each trial as the simulated oscillation score.
Oscillation scores that are calculated with this method estimate the degree to which a single
spike train is oscillating in a given frequency band regardless of the spikes’ temporal structure.

To investigate the oscillatory properties of spike trains without having the confounding effect of
the firing rate, we defined an oscillatory threshold for each trial. To calculate the oscillatory
threshold, we shuffled the spikes of each trial 1000 times in a way that the firing rate remained
the same. Then, we calculated the oscillation score for each shuffled trial. The average of these
1000 simulated oscillation scores plus two standard deviations were assigned to each trial as the
oscillation threshold for that trial. If the oscillation score of a recorded trial was greater or equal to the oscillation threshold of that trial, it was considered as an oscillatory trial. If the oscillation score of a recorded trial was less than the oscillation threshold of that trial, it was considered as a non-oscillatory trial. Because the threshold for the oscillatory trials was calculated based on the firing rate of each individual trial, the confounding effect of firing rate on the number of the oscillatory trials was minimal.

To study the amount of information conveyed by each cell, we calculated the Response Modulation Index (RMI) as described by Hegde and Van Essen (Hegde and Van Essen 2004). The RMI was calculated within 50 ms bins of raw data (non-normalized, non-smoothed). For calculation of RMI during each bin, first we calculated the $F$-statistic of the cell’s responses to the four stimuli presented in the second stage of the experiment. The $F$-statistics were calculated with the following formula:

$$F = \frac{MS_{\text{between}}}{MS_{\text{within}}}$$

Where the $MS_{\text{between}}$ is the stimulus-to-stimulus variance across trials and the $MS_{\text{within}}$ is the average trial-to-trial variance for each stimulus. The $F$-statistic can be interpreted as a measure of the amount of information conveyed about the stimuli or the signal-to-noise ratio. We randomized the responses across the stimuli and recalculated the $F$-statistic to correct the deviations of the $F$-statistic from normality and for calculation of significance. The randomization process was repeated $10^5$ times. The RMI index was defined as the actual $F$-statistic divided by the average $F$-statistic from $10^5$ randomization rounds. For estimating the significance of the RMI in each bin of each cell response, the one-tailed $p$-value was calculated.
as the proportion of times the average $F$-statistic from randomization rounds exceeded the actual $F$-statistic (Hegde and Van Essen 2004; Hegde and Van Essen 2006).

For calculating the RMI values, the spikes in response to the presentation of each stimulus were divided into 20 consecutive time bins of 50 ms extending from 0-1000 ms after the stimulus onset. Then the indices were calculated for each bin. For calculation of cumulative RMI values, each bin was extended backward to the stimulus onset time, then the RMI indices were calculated.

NORMALIZING AND SMOOTHING. In plotting the population histograms and calculating the one-way ANOVA for each cell, the average firing rates (within 0-1000 ms post-stimulus) of 15 repetitions of each stimulus for each presentation duration were normalized to a maximum of 1.0. Then, the given responses were used as a 1 ms resolution spike density function. Qualitatively similar results were observed when any of the reported analyses were done using the raw, instead of normalized firing rates. For the calculation of the peak magnitude and peak latency, the normalized spike density functions were convolved by a Gaussian kernel ($\sigma = 5$ ms).

RESULTS

Here we studied the effect of stimulus presentation duration and baseline activity on the stimulus selectivity and the temporal pattern of neural responses in the inferior temporal cortex (IT) of two macaque monkeys during a passive fixation task. The recording sites were distributed along the anterior parts of the IT (TE) and the lower bank of the superior temporal sulcus (sts) extending over the anterior posterior 11–20 mm in monkey Jn and 7-19 mm in monkey Sh. Altogether, we examined the responses of 259 IT cells by presenting four visual stimuli for five different presentation durations (see Methods for details).
**Temporal course of firing rate**

The activity patterns of two sample neurons in response to their preferred stimuli with five different presentation durations are illustrated in Figure 3. Responses of IT cells consisted of brief excitatory activity that peaked around 140 ms (139.73 ± 32.93 ms; called the “initial burst” hereafter) (Fig. 3, thick arrows) and was sometimes followed by a period of inactivation or inhibition (i.e., activity below baseline) that reached its minimum around 250 ms (248.69 ± 42.41 ms; called the “trough” hereafter) (Fig. 3A, long arrows) and a later excitatory peak around 370 ms (371.59 ± 80.20 ms; called the “second burst” hereafter) (Fig. 3, thin arrows). The presence and magnitude of these and other response components depended on the stimulus effectiveness, presentation duration and cell’s baseline activity as described below.

Responses of most of the recorded IT cells (250 of 259) contained the initial burst component. The stimulus presentation duration had no significant effect on the peak magnitude of the initial burst, the onset peak latency and the mean firing rate measured during a window of 60-240 ms (one-way ANOVA; p>0.05). The initial burst was also observed when other suboptimal stimuli were presented but its peak magnitude and mean firing rate decreased, regardless of presentation duration, as less effective stimuli were presented (p<0.01) (Fig. 4, thick arrow). There was a weak, but significant, correlation between the trough and initial burst peak magnitudes (correlation coefficient = -0.18; p<0.001). This significant correlation suggests that stronger stimuli may invoke stronger inhibition.

The response magnitude of the second burst was also increased as more effective stimuli were presented (Fig. 4, thin arrow). However, this effect was observed only for long (140, 210 and 350 ms; one-way ANOVA; p<0.01) and not brief (18 and 70 ms) presentation durations (p>0.05) (compare the second burst in Figs 4A and B with Figs 4C, D and E). The peak magnitude, the
onset latency and the average firing rate of the second burst were significantly different between
the presentation durations (one-way ANOVA; p<0.05).

Initial observation of the data suggested a clear relationship between the stimulus presentation
duration, the presence and magnitude of the trough and the rhythmic pattern of IT cell activity.
Some cells showed rhythmic phases of sequential excitation and inactivation or inhibition (Fig.
3A), while others showed less oscillatory responses (Fig. 3B). Different degrees of rhythmic
activity were observed in each individual neuron. Brief presentation durations with preferred
stimuli elicited rhythmic activity more often than did long presentation durations. Similar
patterns were observed in the cell population (Fig. 5). The rhythmic activity was mainly
observed in the brief presentations (≤70 ms) with 3-4 excitatory peaks that had a peak-to-peak
time interval of about 200-300 ms (Fig. 5).

To investigate the relationship between the stimulus presentation duration and the rhythmic
pattern of cell activities, we calculated the “oscillation score” for each trial within the theta
frequency band (4-8 Hz). In brief, we first calculated the autocorrelation histograms (ACH) of
each individual trial. Then, a fast Gaussian kernel was used to smooth the ACH and to remove
high-frequency noise. Then, for the sole purpose of detecting the boundaries of the central peak,
a slow Gaussian kernel was applied. Using this information, the central peak was efficiently
removed from the buffer containing the fast smoothing, which was then subjected to FFT.
Eventually, the oscillation score was computed as the ratio between the highest frequency
magnitude within the band of interest and the average baseline magnitude of the spectrum.
Oscillation scores that were calculated with this method estimate the degree to which a neuron
oscillates in a given frequency band (for more detail see Methods and Muresan et al. 2008).
Figure 6 shows the distribution of oscillation scores from trials in three sample neurons with different oscillation score values. Increases in the stimulus presentation duration shifted the distribution of the oscillation score toward lower values (Fig. 6A-C). The average oscillation score of all neurons at each condition is plotted in Figure 6D. In general, we observed that presenting stimuli for shorter durations resulted in higher neural oscillatory activity (Figs 4 and 6).

To test the significance of this effect, we calculated the average oscillation scores of all trials for each presentation duration. Then we performed a two-way ANOVA analysis on the oscillation scores with stimulus strength and stimulus presentation duration as the factors. Results of the two-way ANOVA showed that both stimulus strength and presentation duration had a significant effect on the average oscillation scores of the neurons, but there was no interaction between the two factors (p=0.0002, p<<0.0001 and p=0.9 for stimulus strength, stimulus presentation duration and their interaction, respectively).

To investigate the effect of baseline activity on the oscillatory activity of neurons, we calculated the average firing rate of a 200 ms window before stimulus presentation as the baseline activity of each trial. Figure 7 illustrates the relationship between the oscillation score and the baseline activity. In Figure 7A the average oscillation scores are plotted against the average baseline activity of cells and show a positive correlation (r=0.7, p<0.0001). The histograms presented in Figure 7A represent the number of cells that contributed to the scatter plot. Figure 7B plots the average oscillation score of the neurons with similar average baseline activity. The correlation coefficient between the average oscillation score and the average baseline activity of neurons was smaller for the brief stimulus presentation durations (18 and 70 ms) compared with the long presentation durations (210 and 350 ms) (correlation coefficient of 18, 70, 140, 210, 350
presentation duration conditions were 0.75, 0.76, 0.63, 0.67, 0.56, respectively with p<<0.0001 in all cases).

Figure 7C shows the correlation between the average oscillation score and the baseline activity of neurons for 18 ms and 350 ms stimulus presentation durations. To examine the simple effect of the baseline firing rate on the oscillation score without the impact of the temporal structure of the spikes, we calculated simulated oscillation scores by shuffling the spike times of each trial 1000 times while keeping the firing rate constant. Then, we calculated the oscillation score for each shuffled trial. The mean of 1000 shuffled oscillation scores was assigned to each trial as a simulated oscillation score. There was a negative correlation between the baseline activity of neurons and the simulated oscillation scores (r=-0.1; p<<0.0001; see Methods). The same trend was obtained when trials from each individual condition (each stimulus presentation duration and stimulus strength, separately) were used for this analysis.

To investigate the oscillatory properties of different subdivisions of the IT cortex without the potential contamination of response rates, we defined an oscillatory threshold for each trial independent from spike rate (see Methods). We found that the number of oscillatory neurons in three different areas were significantly different (one-way ANOVA; p=0.01). Tukey-Kramer post-hoc analysis revealed that the number of oscillatory neurons was significantly lower in the TEa compared to the TEp. Figure 8 illustrates the average number of oscillatory trials across stimulus presentation durations and subdivisions of the IT cortex.

In addition, the average baseline activity of oscillatory and non-oscillatory trials was significantly different based on a t-test with a p-value<<0.0001 (mean and SEM of baseline
activity in oscillatory and non-oscillatory conditions were 6.74 ± 0.15 and 7.65 ± 0.06 (sp/s), respectively).

**Response Modulation Index**

To compare the effect of the stimulus presentation durations on the amount of information conveyed by each cell over the temporal course of its responses, we calculated the Response Modulation Index (RMI). The RMI is based on the \( F \) ratio and provides an explicit measure of the signal-to-noise ratio. In short, the RMI was defined as the actual \( F \)-statistic divided by the average \( F \)-statistic calculated from \( 10^5 \) randomization times. For estimating the significance of the RMI, the one-tailed \( p \)-value was calculated as the proportion of times the average \( F \)-statistic from randomization rounds exceeded the actual \( F \)-statistic in a given time window (see Methods, and Hegde and Van Essen 2004 and 2006).

We calculated the mean RMI values for different stimulus presentation durations. There was no significant difference in the mean RMI during the early phase of responses (from stimulus onset to 225 ms post-stimulus) among five different stimulus presentation durations according to a one-way ANOVA (\( p>0.05 \)). Interestingly, there was a systematic and statistically significant increase in the RMI values during the later part of responses as stimuli were presented for longer durations (Fig. 9A). The RMI values of different stimulus durations were significantly different within a time window of 225-675 ms (ANOVA; \( p<0.05 \)). The bold line in Figure 9A represents the time bins in which the RMI values showed a significant difference (\( p<0.05 \)) among stimulus presentation durations. Note that the RMI values declined with a slower pace at the early stage of the responses to long stimulus presentation durations compared with the brief presentation durations (Fig. 9A). The same trend was observed in cumulative RMI values. The significant
difference of the cumulative RMI values among stimulus presentation durations was started at
the 325 ms bin (Fig. 9B).

To test the relationship between the pattern of activity and the amount of information conveyed
by the neural code, we calculated the cumulative RMI value of 400 ms after stimulus onset.
There was a significant negative correlation between the cumulative RMI value and the number
of oscillatory trials in each neuron (p<<0.0001 and r=-0.11). This suggests that neurons with a
higher capacity for oscillation convey less information about stimuli. To confirm the relationship
between oscillatory activity and information conveyed about stimuli, we calculated the average
oscillation score for each neuron. Then, we divided neurons into two groups, the first group with
average oscillation scores greater than the median oscillation score of the population and the
second group with average oscillation scores less than the median oscillation scores of the
population. The average RMI values during a 70-550 ms window of first group was significantly
smaller than second group according to the t-test (p=0.005).

To examine the role of different IT subdivisions in visual information processing, we calculated
the RMI values of cells in each area separately using spiking activity within the time window of
0-550 ms (post-stimulus) for each stimulus presentation condition. We then calculated the mean
values in each IT subdivision and plotted them against stimulus presentation durations (Fig. 10).
As the stimulus presentation duration increased, there was a systematic increase in the RMI
values of the TEa and the STS, but minimal or no such increase was observed in the TEp
according to a one-way ANOVA (significant p<0.05). Note that the RMI values were highest in
the TEa for all of the presentation durations except for 18 ms. Application of a two-way
ANOVA showed that both factors, presentation duration and cortical area, had a significant
effect on the RMI values during the first 550 ms of the response (p<0.0001 in both cases) while
there was not any significant interaction between factors (p=0.21). Post-hoc analysis of the
results revealed that the “long presentation duration” and the “TEa” conditions had significantly
higher RMI values compared to other conditions. Consistent with this result, we found that the
number of oscillatory neurons in three different areas were significantly different (one-way
ANOVA; p=0.01). Tukey-Kramer post-hoc analysis revealed that the number of oscillatory
neurons was significantly lower in the TEa compared to the TEp (Fig. 8). To examine the
temporal dynamics of cell responses in the different IT subdivisions as the stimulus presentation
duration was changed, we performed a two-way ANOVA using the RMI values of variable time
windows starting from stimulus onset with a length of 50-1000 ms increasing with 5 ms steps.
We found that the effect of area appeared from the time window length of 100 ms and longer,
but the stimulus presentation duration effect appeared much later from the time window length of
225 ms and longer.

The monkeys were required to maintain fixation within a 4° window during stimulus
presentation and the following gray interval. It is possible that more frequent eye movements
during or following the brief stimulus presentation cause larger degrees of neural oscillations. To
address this issue we calculated the number of saccades inside fixation window (i.e., 1-4
degrees) during the first 1000 ms after stimulus onset. Saccades were identified as eye
movements with velocity of more than 25°/s for at least 15 consecutive milliseconds. From 388
recording sessions, in only one session (0.25%), the average number of saccades was
significantly different among different presentation durations (ANOVA p-value<0.05).
DISCUSSION

Overview

Our results demonstrate that, depending on the magnitude of the baseline firing rate, stimulus presentation duration affected the response properties and efficiency of neural information processing in the IT cortex. Short stimulus presentation durations and high baseline activity were associated with stronger oscillatory responses, which conveyed less information about stimulus. At long stimulus presentations and low baseline activities, neural responses were less oscillatory and carried a higher amount of information. Finally, there were significantly higher oscillatory responses and a lower amount of information in the posterior IT (TEp) compared to the anterior IT (TEa).

Impact of stimulus presentation duration on cell responses in the ventral visual pathway and on object recognition

Neurons in the IT cortex have been shown to respond to oscillatory activity during visual processing (Nakamura et al. 1991, 1992; Eskandar et al. 1992; Sheinberg and Logothetis 1997). Also, the impact of stimulus presentation duration on cell responses has been shown previously in the primary visual cortex of monkeys (Gawne et al. 1991; Muller et al. 2001). In addition, it has been shown that the reliability of discrimination between two stimuli increases with increasing the stimulus presentation duration in V1 (Zohary et al. 1990) and the IT (Allred and Jagadeesh 2007). In general, our findings are consistent with this previous evidence, which together provide insight into the neural basis of the effect of stimulus presentation duration on object recognition.
Behavioral studies have shown that short fixation durations (<70 ms) result in missing the fixated target in a visual search task (Hooge and Erkelens 1996). Also, object categorization and identification performance of humans rapidly declines as the stimulus presentation duration decreases (Grill-Spector and Kanwisher 2005; Kovacs et al. 1995; Rieger et al. 2005; Rolls et al. 1999). The steepest decline in performance occurs in presentation times <70 ms. Our result that with brief stimulus presentations (≤70 ms) there was no information in the IT cell responses following the initial peak (>250 ms post-stimulus) suggests an important role for these late responses in object recognition. However, the lack of object discrimination tasks in our paradigm should be considered in the interpretation of the presented neural data and its relevance to the behavioral data obtained during visual search. A comparison of neural responses and behavioral performance related to the stimulus presentation duration can be found in a recent study by Allred and Jagadeesh, 2007.

Impact of baseline firing rate on cortical sensory processing and object recognition

Even in the absence of sensory inputs, cortical cells show structured patterns of spontaneous activity (Dehaene and Changeux 2005). The functional significance of these fluctuations is not well understood. Emerging evidence suggests that changes in the baseline activity of neurons are controlled by top-down mechanisms and that they can affect sensory processing. For example, the examination of monkey V1 cell responses during a figure-ground detection task shows that baseline neural activity influences both cell responses and the monkey’s behavior. The baseline activity of V1 cells prior to stimulus presentation is stronger and the amount of synchrony between cells is larger in trials when the monkey detects the stimuli compared to trials in which it does not detect the stimuli (Super et al. 2003). However, other conflicting results have been reported for the effect of the state of baseline neural activity on
cortical sensory processing (Haider et al. 2007; Leger et al. 2005; Petersen et al. 2003; Sachdev et al. 2004; Timofeev et al. 1996). While studies in the rodent somatosensory system have shown strongly diminished responsiveness of cortical neurons when baseline activity is high (Petersen et al. 2003; Sachdev et al. 2004), a recent study shows that higher neural activity preceding the presentation of stimuli enhances responsiveness and scaling of the contrast response function in cat V1 cells (Haider et al. 2007). Consistent with the former reports, our results show that when there was a higher level of baseline activity, rhythmic responses were evoked that contained less information compared to when the baseline activity was low. Also, a similar relationship between high baseline firing rates and oscillatory activity in the IT is evident in Figure 12 of Rollenhagen et al. 2005. These data show that presentation of a stimulus before a flanker increases the baseline firing rate of IT cells and results in higher oscillatory activity following the flanker presentation compared with when the flanker is presented alone.

A better understanding of the interaction between baseline activity fluctuations at both the network and single-cell level is needed to explain how global or regional brain activity can affect the information processing properties of single cortical cells by changing their baseline activity. It has been suggested that the fluctuation in baseline activity and the resulting change in the response properties of cortical cells could be related to rhythmic changes in the global baseline brain activity that are thought to be related to top-down phenomena such as attention (Destexhe et al. 2003; Hasenstaub et al. 2005; Ho and Destexhe 2000; McCormick et al. 2003; Shu et al. 2003). Indeed, increases in baseline neural activity have been shown in primate cortical visual areas during selective attention (Luck et al. 1997; McMains et al. 2007; Reynolds et al. 2000; Williford and Maunsell 2006).

Functional significance of baseline activity and neural oscillation
Visual memory (Nakamura et al. 1991 and 1992) and interaction between competing stimuli (Rollenhagen et al. 2005) affect oscillatory activity in the IT cortex. Modeling studies suggest that as the strength of the inhibition between competing neuronal pools is increased, the pattern of network behavior shifts to an oscillatory mode (Moldakarimov et al. 2005). In our study, changing the strength of the stimulus (by using less effective stimuli and decreasing the stimulus presentation duration) caused higher levels of neural oscillation. We suggest that the strength of inhibition between competing neuronal assemblies might be affected by the state of baseline activity and the resulting differential recruitment of inhibition following stimulus presentation.

Responses of individual IT cells contained a larger amount of sensory information about an individual stimulus and carried a relatively sparser neural code when stimuli were presented for longer durations as well as when more effective stimuli were presented. The sustained activity evoked by more effective stimuli that were presented for longer durations can be conveyed to downstream sites more easily due to a decrease in the inhibitory rhythmic volleys, allowing the stimulus to recruit higher-order neurons more easily. On the other hand, brief stimulus presentations could recruit higher levels of IT cell-response coherence by setting off synchronous activity within the IT cell population, making efficient use of the distributed information and thus improving recognition performance as has been shown in the lower visual areas (Super et al. 2003).

The functional difference of IT subdivisions

The number of oscillatory trials was lower and the amount of information was higher in the TEa compared with the TEp, particularly for longer presentation durations. The pattern of cortico-cortical connections is largely different between the TEp and the Tea, prompting investigators to
suggest a differential functional role for these IT subdivisions (Morel and Bullier 1990). But little is known about the functional differences between these areas (Tamura and Tanaka 2001). Our results clearly show functional differences between the subdivisions of the IT, suggesting that each cortical area may play a unique role in visual object recognition.

The potential impact of saccades on neural oscillation

The monkeys in our study were required to maintain fixation within a 4° window during stimulus presentation and the following gray interval. Our finding that the number of small amplitude saccades (in range of 1° to 4°) was not significantly different among different stimulus presentation durations indicates that such eye movements are not causing the oscillatory neural activities reported in this study. But, it is still possible that microsaccade eye movements within the fixation window affected the reported response properties of IT neurons in this study. Due to the lack of fine eye movement data, we cannot directly rule out the possibility that more frequent microsaccades occur following the brief stimulus presentation causing larger degrees of neural oscillations. However, this seems unlikely as it has been shown that microsaccades have no effect on IT neural activity (Leopold and Logothetis 1998) and that they do not affect the low-frequency oscillation of IT cell responses (Rollenhagen et al. 2005). In addition, our finding that the oscillation is correlated with higher baseline activity (which is not related to, or affected by, stimulus presentation duration) suggests that microsaccades do not play a role in the reported results.

In conclusion, our results demonstrate that depending on the magnitude of baseline firing rate, stimulus presentation duration affected the response properties and efficiency of neural information processing in IT cortex. Given that, it is plausible that during natural vision top-
down influence or competition within local neural networks differentially influence encoding properties of IT cells by changing their baseline activity.

ACKNOWLEDGEMENTS

We would like to thank Dr. James Bisley for his useful comments on the manuscript.

References:


Figure captions:

**Fig. 1** Sagittal sections of the MRI images at the level of the posterior (left figures) and anterior (right figures) inferotemporal cortex. The posterior sections (AP 9 and AP 11) show the boundaries of the T Ep and the lower lip of the STS (blue lines). Anterior sections (both AP 19) show the boundaries of the T Ea and the lower bank of the STS (blue lines). The abbreviations stand for: superior temporal sulcus (sts), anterior middle temporal sulcus (amts), posterior middle temporal sulcus (pmts), occipitotemporal sulcus (ots) and rhinal sulcus (rhs).

**Fig. 2** The experimental paradigm of the second stage of the experiment. The monkeys had to maintain their gaze at the fixation point for 200 ms until the fixation point disappeared and after 300 ms delay, the presentation of stimulus sequence started. The selected stimuli were randomly shown with one of the five different stimulus presentation durations. The sequence stopped when the monkey broke the gaze fixation and a new block started with appearance of a fixation point.

**Fig. 3** Peristimulus time histograms of two exemplar neurons. A) A neuron with oscillatory responses particularly at shorter presentation durations. B) A neuron with the tendency to produce more sustained responses. Each histogram is an averaged PSTH response of one exemplar neuron to 15 repetitions of the best stimulus. The values indicated above the top row of PSTHs depict the stimulus presentation duration. PSTHs bin size is 1 ms. The initial burst, second burst and trough are depicted by a thick arrow, thin arrow and long arrow, respectively. Troughs are not depicted in sample B because troughs were not prominently seen in neurons without strong oscillatory activity.

**Fig. 4** Average normalized firing rates of all recorded cells (n=259) to four different stimuli for A) brief presentation duration (18 ms), B) 70 ms, C) 140 ms, D) 210 ms and E) long presentation duration (350 ms). In each case, the corresponding population average was smoothed with a Gaussian kernel, SD=5 ms. The error ranges denote ± 1 SEM. The initial burst, second burst and trough are depicted by a thick arrow, thin arrow and long arrow, respectively.

**Fig. 5** Average normalized firing rates of all recorded cells (n=259) to A) preferred (best) stimuli B) second suboptimal stimuli, C) third suboptimal stimuli and D) worst stimuli at five presentation durations. Population averages were smoothed with a Gaussian kernel, SD=5 ms. The error ranges denote ± 1 SEM.

**Fig. 6** Distribution of oscillation scores in individual neurons. The distribution of the oscillation score of all trials in three sample neurons. Three neurons are selected among A) 25% of neurons with the lowest average oscillation score, B) 50% of neurons with the middle range average oscillation score and C) 25% of neurons with the highest average oscillation score. The average oscillation scores of the sample neurons are mentioned in each subplot (OS stands for the average oscillation score of the sample). D) The average oscillation scores of all neurons are plotted at each presentation condition. The error ranges denote ± 1 SEM.
**Fig. 7** The average oscillation scores plotted against the average baseline activity. A) Each point in the plot represents one neuron. The average baseline activity (horizontal axis) was plotted against the average oscillation score (vertical axis). The dashed line shows the linear fit to the data ($r=0.7$, $p<0.0001$). The distribution of the average oscillation score and baseline activity of neurons are plotted on the vertical and horizontal axes of the plot, respectively. B) The average oscillation score of neurons are plotted at distinct levels of baseline activity with resolution of 1 spike/sec. The baseline levels containing less than three neurons are omitted from the plot. In total, 21 neurons (from 259) are omitted from the plot. The dashed line shows the linear fit to the data ($r=0.8$, $p<0.0001$). SEMs represent the variability of oscillation scores for each baseline rate. C) Each symbol in the plot represents one neuron. The average baseline activity (horizontal axis) was plotted against average oscillation score (vertical axis) for brief (18 ms) and long (350 ms) stimulus presentation duration by circles and squares, respectively. The dashed line shows the linear fit to the circles ($r=0.75$, $p<0.0001$). The solid line shows the linear fit to the squares ($r=0.56$, $p<0.0001$).

**Fig. 8** The average number of oscillatory trials across neurons. The average number of oscillatory trials of 259 neurons is plotted as the function of stimulus presentation duration for each subregion of the IT cortex. The error bars denote ± 1 SEM.

**Fig. 9** The amount of information conveyed by IT cells over the temporal course of their responses. A) The average Response Modulation Index (RMI) was calculated by averaging the RMI values of individual IT cells across different stimulus presentation durations. B) The average cumulative RMI was calculated by averaging the cumulative RMI values of individual IT cells across different stimulus presentation durations. The bold line in each graph illustrates time bins in which the RMI or cumulative RMI values showed significant differences between five different stimulus presentation durations ($p<0.05$). The error bars indicate ± 1 SEM.

**Fig. 10** The average cumulative RMI values of cells in different IT subdivisions for different stimulus presentation durations. The cumulative RMI values were calculated during first 550 ms from the stimulus presentation onset. The error bars denote ± 1 SEM.
Fixation (200 ms)
Delay (300 ms)
Stimulus 1 (18 ms)
ISI (1300 ms)
Stimulus 2 (210 ms)
ISI (1300 ms)
Fixation break
Response (sp/s)

Time from stimulus onset (ms)

18 ms 70 ms 140 ms 210 ms 350 ms
A: 18 ms

B: 70 ms

C: 140 ms

D: 210 ms

E: 350 ms
Average normalized response

Time from stimulus onset (ms)

A: Best stimuli
- 18 ms
- 70 ms
- 140 ms
- 210 ms
- 350 ms

B: Second stimuli

C: Third stimuli

D: Worst stimuli
A

Worst   OS=8.5   Best

18 ms   70 ms   140 ms   210 ms   350 ms

Frequency

B

OS=14.4

18 ms   70 ms   140 ms   210 ms   350 ms

Frequency

C

Worst   OS=34.3   Best

18 ms   70 ms   140 ms   210 ms   350 ms

Frequency

D

Average oscillation score

Presentation duration
Presentation duration

Average number of oscillatory trials ± SEM

- TEp
- TEa
- STS
Average cumulative RMI ± SEM

Stimulus presentation duration