Temporal dynamics of face repetition suppression

Alumit Ishai a,*, Philip C. Bikle b, Leslie G. Ungerleider b

a Institute of Neuroradiology, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
b Laboratory of Brain and Cognition, National Institute of Mental Health, NIH, Bethesda MD 20892, USA

Received 31 March 2006; received in revised form 21 May 2006; accepted 1 June 2006
Available online 27 June 2006

Abstract

Single-unit recordings and functional brain imaging studies have shown reduced neural responses to repeated stimuli in the visual cortex. Using MEG, we compared responses evoked by repetitions of neutral faces to those evoked by fearful faces, which were either task relevant (targets) or irrelevant (distracters). Faces evoked a bi-phasic response in extrastriate cortex, peaking at 160–185 ms and at 220–250 ms, with stronger responses to neutral faces at the earlier interval and stronger responses to fearful faces at the later interval. At both latencies, repetitions of neutral and fearful targets resulted in reduced amplitude of the MEG signal. Additionally, we found that the context in which targets were presented affected their processing: fearful distracters increased the responses evoked by both neutral and fearful targets. Our data indicate that valence enhancement and context effects can be detected in extrastriate visual cortex within 250 ms and that these processes likely reflect feedback from other regions.

Keywords: MEG; Face perception; Emotion; Valence enhancement; Context effects

1. Introduction

Single-unit recordings in non-human primates [3,5,27] and functional brain imaging studies in humans [2,7,8,20,23,30,33,37] have shown reduced neural responses to repeated stimuli (e.g., words, objects and faces), a phenomenon known as repetition suppression. To account for the reduced neural responses associated with stimulus repetition, three models have recently been suggested: the Fatigue model, according to which reduced responses reflect firing rate adaptation; the Sharpening model, according to which repetition suppression reflects sparser neural representation of stimuli; and the Facilitation model, according to which decreased activation reflects shorter latencies of neural firing and reduced prediction error [17].

Using event-related fMRI, we have recently compared activation evoked by repetitions of neutral and fearful faces, which were either task relevant (targets) or irrelevant (distracters). We found that within a distributed cortical network that included visual, limbic and prefrontal face-selective regions, targets evoked stronger responses than distracters and their repetition was associated with significantly reduced responses. Repetition suppression, as manifested by the difference in response amplitude between the first and third repetitions of a target, was stronger for fearful than for neutral faces. Moreover, distracter faces, regardless of their repetition or valence, evoked negligible activation, indicating top–down attenuation of behaviorally irrelevant stimuli [22]. Our fMRI study, which provided neural evidence for a three-way interaction between valence, repetition and task relevance in the human brain, supports the predictive coding model [13]. We interpreted these results to indicate sharp tuning of the cortical representations of behaviorally relevant emotional faces. Due to the limited temporal resolution of the fMRI technique, we were unable to investigate the temporal dynamics of detecting neutral as compared with fearful targets. Moreover, we did not have the resolution to study the context in which targets were presented, namely whether there were differences between the detection of neutral targets embedded in a stream of neutral distracters and the detection of neutral targets presented with fearful distracters. We therefore adapted the fMRI paradigm to the magnetoencephalography (MEG) environment.

Previous event-related potential (ERP) and MEG studies have identified a likely neural correlate of face processing, namely a negative, posterior potential peaking at 170 ms post-stimulus (N170/M170), which is characterized by a
selective response to faces as compared with assorted common objects and thus reflects the structural encoding of face stimuli [4,6,11,14,15,21,25,26]. We postulated that valence enhancement and context effects could be detected with MEG and hypothesized that the type of distracters presented in each trial would affect the processing of the targets. Specifically, we expected differential modulation of the M170 response to neutral and fearful targets as a function of the context in which they were presented, such that in the context of fearful distracters, the evoked responses to neutral targets would be stronger than in the context of neutral distracters.

We report here a bi-phasic response: neutral and fearful faces evoked significant responses that peaked at 170 ms and at 230 ms, with stronger responses to neutral faces at the earlier latency and stronger responses to fearful faces at the later latency. Consistent with our previous fMRI study, we found that target repetition resulted in reduced amplitude of the MEG response and that behaviorally relevant targets evoked stronger responses than the irrelevant distracter faces. Finally, we found that the context in which targets were presented affected their processing, namely, that fearful distracters increased the responses evoked by both neutral and fearful targets.

2. Experimental procedures

2.1. Subjects

Twelve normal, right-handed subjects (six males, age 25 ± 2 year) with normal vision participated in this study. All subjects gave written informed consent for the procedure in accordance with protocols approved by the NIMH Institutional Review Board.

2.2. Stimuli and task

Gray-scale photographs of neutral and fearful faces from our database [22] and scrambled pictures of these faces were displayed using Presentation (www.neurobs.com). The phased scrambled pictures were generated using an in-house Matlab script, by randomizing the phase of each spatial frequency in the face images while maintaining a constant spatial frequency power density. In each trial, a target face was presented for 4 s, followed by 13 faces, each presented for 2 s. The target and one of the distracters were presented three times, mixed with seven neutral distracters. Targets were presented in two contexts: either with same-valence repeated distracters or with different-valence repeated distracters, thus five trial types were included: repeated neutral targets + repeated neutral distracters; repeated neutral targets + repeated fearful distracters; repeated fearful targets + repeated neutral distracters; repeated fearful targets + repeated fearful distracters; and repeated scrambled targets + repeated scrambled distracters. Subjects were instructed to memorize the target face and press a button when detecting it, thereby making the target the behaviorally relevant stimulus. Each session included 80 trials, with 16 randomized epochs with 5 consecutive repetitions of each trial type. Stimuli were projected with a magnetically shielded LCD video projector while subjects assumed a supine position.

2.3. Data acquisition

Neuromagnetic data were recorded at 600 Hz with a bandwidth of 0–150 Hz using a CTF 275 MEG system (VSM MedTech Ltd., Canada) composed of a whole-head array of 275 radial 1st order gradiometer/SQUID channels housed in a magnetically shielded room (VacuumSchmelz, Germany). Synthetic third gradient balancing was used to remove background noise on-line. For each subject we performed two localizations, in the beginning and at the end of the session, to determine the position of the head.

2.4. Data analysis

Noise reduction was implemented during recording using third gradient balancing and the DC offset was removed based on the whole trial. Data were high pass filtered at 0.61 Hz. Stimuli and responses were flagged within each dataset for analysis. Additional datasets were created to reflect the average MEG response for specific stimuli across all trials. Root mean squares were applied to all waveforms to convert them to a magnitude scale. A low-pass filter of 15 Hz was applied to the data.

2.4.1. Waveform analysis

For each subject, six sensors were selected in the right ventral stream region that showed significant evoked responses to faces (faces > scrambled faces). Data were averaged across all sensors from each subject. For both latencies (160–185 ms and 220–250 ms) and for each individual, MEG responses were averaged across each stimulus type. This averaged response was then converted to RMS magnitude. Due to the sampling rate of the data, namely 600 Hz, there were 16 and 19 sampling points, for the early and late intervals, respectively. A recent study has shown that log- or sqrt-transform of EEG/MEG data render the data normal, thus inferences can be made using a parametric approach [24]. We therefore used the sqrt-transform data and report the results of a two-tailed t-test for all comparisons.

2.4.2. Synthetic aperture magnetometry (SAM) analysis

Event-related desynchronizations (ERDs) and event-related synchronizations (ERSs), respectively, were estimated using SAM statistical analysis [35]. SAM measures the source magnitude by forming a linear combination of sensors that can suppress magnetic field generated by sources other than the target voxel. To determine which frequency bands would be used in the SAM analysis, the magnitude distribution in the time and frequency domains was examined. These plots revealed an evoked field in the range of 7–13 Hz from around 150 to 250 ms. The SAM analysis was then set up to determine source magnitude distributions within the 7–13 Hz range from 150 to 250 ms. Changes in the source magnitude were assessed at each voxel for the active state (all face stimuli; 150–250 ms post stimulus) and for the control state (all scrambled faces; 150–250 ms post stimulus) using Student t-values. Finally, a random permutation t-test was performed to create a group map of ERDs and ERSs. The MEG data were co-registered with each subject’s anatomical MR images using a head-shape surface match.

2.5. Acquisition of MR images

A 3 Tesla General Electric Signa scanner with a whole head coil (IGC-Medical Advances, Milwaukee, WI) was used to collect high-resolution spoiled gradient recalled echo structural images for each subject (124, 1.3-mm thick sagittal slices, TR = 15 ms, TE = 5.4 ms, FOV = 24 cm, flip angle = 45°, 256 × 256 matrix). These T1-weighted images were 3D normalized to the brain atlas space [34] using AFNI [9] and the source data were converted into z-scores. Finally, a random permutation t-test was performed to create a group map of ERDs and ERSs. The MEG data were co-registered with each subject’s anatomical MR images using a head-shape surface match.

3. Results

3.1. Behavioral data

The mean accuracies for detecting the neutral and fearful targets were 95 and 97%, respectively. Fig. 1 shows the reaction times for detecting neutral and fearful targets during their first, second and third repetitions. Consistent with our prior fMRI study [22], detection of fearful target faces was faster than detection of neutral target faces (t11 = 2.2, p < 0.05). Furthermore, we found shorter latencies with repetition: the difference in reaction
times between the first and third repetitions of neutral targets was statistically significant \((t_{11} = 2.42, p < 0.03)\). A similar trend was observed for fearful targets \((t_{11} = 1.9, p = 0.08)\).

3.2. MEG data

3.2.1. Activation evoked by visual perception of faces

Fig. 2 shows the location of sensors in the right extrastriate visual cortex (A) and a coronal section indicating the location of activation in the right fusiform gyrus, as determined by the SAM analysis (B). Perception of faces, as compared with scrambled faces, evoked a bi-phasic response with two significant latencies: 160–185 ms (peaking at 170 ms) and 220–250 ms (peaking at 230 ms). At both latencies, the difference between the response to faces and the response to scrambled faces was statistically significant \((t_{11} = 3.14, p < 0.009\) and \(t_{11} = 4.57, p < 0.001\), respectively) (Fig. 2C). We then compared the responses evoked by neutral faces to the responses evoked by fearful faces. Interestingly, stronger responses to neutral than fearful faces were observed at 160–185 ms \((t_{11} = 2.2, p < 0.05)\), whereas stronger responses to fearful than neutral faces were observed at 220–250 ms \((t_{11} = 4.34, p < 0.001)\) (Fig. 2D).

3.2.2. Valence enhancement and context effects

The comparison between responses evoked by the behaviorally relevant targets and responses evoked by the behaviorally
irrelevant repeated distracters revealed that both neutral and fearful targets evoked stronger responses than neutral and fearful distracters (Fig. 3A and B). The difference between the response to neutral targets and neutral distracters was significant at the earlier latency of 160–185 ms ($t_{11} = 3.11, p < 0.01$). A similar trend was observed for the difference between the response to fearful targets and fearful distracters at the later interval of 220–250 ms ($t_{11} = 1.82, p = 0.09$). At both latencies, fearful repeated distracters evoked stronger responses than neutral ones ($t_{11} = 2.2, p < 0.05$ and $t_{11} = 2.7, p < 0.02$, respectively).

To test whether the context in which the targets were presented had any influence on their processing, we compared responses on trials in which neutral targets were presented with repeated neutral distracters with responses on trials in which the repeated distracters were fearful (Fig. 3C). Similarly, we compared responses on trials in which fearful targets were presented among neutral repeated distracters with responses on trials in which fearful targets were presented among fearful distracters (Fig. 3D). Interestingly, at the later interval (220–250 ms), stronger responses were observed during trials in which neutral targets were presented with fearful repeated distracters than with neutral ones ($t_{11} = 2.2, p < 0.05$). Similarly, fearful targets presented with fearful distracters evoked stronger responses than fearful targets presented with neutral distracters ($t_{11} = 2.2, p < 0.05$).

### 3.2.3. Repetition suppression

To test whether the amplitude of the MEG signal decreased with target repetitions, we compared responses evoked by the first, second and third repetitions of the neutral and fearful targets (Fig. 4A). The repetition of both neutral and fearful targets resulted in reduced responses; however, we found differences between the early and the late response components. At the earlier latency (160–185 ms), neutral targets evoked stronger responses than fearful ones during the first, second and third repetitions ($t_{11} = 2.2, p < 0.05$). At the later interval (220–250 ms), fearful targets evoked stronger responses than neutral targets (first repetition: $t_{11} = 2.2, p < 0.05$; second repetition: $t_{11} = 3.6, p < 0.004$; third repetition: $t_{11} = 2.8, p < 0.02$). At both latencies, repetition suppression, as manifested by the difference in response amplitude between the first and the third repetitions, was significant for neutral and fearful targets ($t_{11} = 2.2, p < 0.05$).

We then compared the repetition of neutral distracters with the repetition of fearful distracters. At the early latency, the repetition of both neutral and fearful distracters was associated with reduced responses (Fig. 4B). The difference between the first and the third repetition of neutral distracters was significant ($t_{11} = 2.66, p < 0.02$), as was the difference between first and third repetition of fearful distracters ($t_{11} = 2.2, p < 0.05$). At the later interval, the difference between the first and the third repetition of neutral distracters was significant ($t_{11} = 2.4, p < 0.03$).
but the repetition of fearful distracters was not associated with reduced responses.

4. Discussion

We investigated, using MEG, the neural responses evoked by repetition of neutral and fearful faces, which were either task relevant (targets) or irrelevant (distracters). Behaviorally, we found facilitation with target repetition, as manifested by shorter reaction times, consistent with previous reports [32]. When compared with scrambled faces, face stimuli evoked significant responses at two latencies: 160–185 and 220–250 ms. Neutral faces evoked stronger responses that peaked at 170 ms, whereas fearful faces evoked stronger responses that peaked at 230 ms. The behaviorally relevant target faces evoked stronger responses than the behaviorally irrelevant distracters. The repetition of both neutral and fearful targets resulted in reduced amplitudes of the MEG signal. Finally, we found that in the context of repeated fearful distracters, the processing of both neutral and fearful targets was enhanced.

4.1. Evoked responses to faces

The visual response evoked by faces revealed a bi-phasic waveform with early and late components. These responses likely originated in extrastriate visual cortex, as determined by the source analysis. Consistent with numerous ERP and MEG studies, we found a significant response to faces that peaked at 170 ms [4,6,11,14,15,21,25,26]. It has been suggested that the N170/M170 reflects a face-selective response component, as the amplitude of the response to assorted face stimuli (e.g., natural faces, painted portraits, sketches of faces, or schematic faces) can be distinguished from the response evoked by any other control stimuli [18,31].

Interestingly, we found stronger responses to fearful faces that peaked at 230 ms. To our knowledge, this is the first study to show that fearful faces evoke stronger responses at this latency. A previous ERP study reported that affective judgments of faces modulated the response at 160 ms, with stronger amplitudes for faces that subjects “liked” as opposed to “neutral” or faces they “disliked” [29]. These subjective ratings, however, depended on facial attractiveness and not facial expressions, thus a comparison of neutral with emotional faces was not provided. Our findings of a late component that was associated with fearful faces suggest that the M230 response is modulated by feedback processes, presumably originating in the amygdala or prefrontal cortex [1,38]. Consistent with this interpretation, a recent ERP study has shown that fearful faces elicit a positive response in prefrontal cortex 120 ms after stimulus presentation [12].

4.2. Valence enhancement and context effects

When compared with neutral faces, fearful faces evoked stronger responses at a later latency, peaking at 230 ms. It is of interest that this valence enhancement was observed not only for the behaviorally relevant targets, but also for the
irrelevant distracter faces. These delayed, enhanced responses to emotional faces are consistent with previous fMRI studies that showed stronger responses to emotional faces [22,28,36].

Although the responses evoked by targets were significantly stronger than the responses evoked by distracters, we found that the processing of the targets was affected by the valence of the distractors. Our subjects detected neutral or fearful targets, which were presented in a stream of either neutral or fearful distracters. In our prior fMRI study, we did not find any effect of the context in which the targets were presented and therefore reported the mean responses to neutral and fearful targets, regardless of the valence of the distractors among which they were presented. Here, however, we were able to dissociate the MEG responses and revealed differential context effects. The presence of repeated fearful distracters enhanced the responses evoked by neutral targets at the later interval of 220–250 ms and the response evoked by fearful targets at both latencies. Our findings suggest that these two neural manifestations, namely, valence enhancement and context effect, may reflect the privileged status of emotional stimuli in the brain [10].

4.3. Targets and distracters

Consistent with our fMRI study [22], we found that the behaviorally relevant targets, regardless of their valence, evoked stronger responses than the behaviorally irrelevant distracters in extrastriate cortex. The task relevance of targets was presumably accompanied by enhanced attention to them. Recent studies have also shown that top–down effects, such as stimulus familiarity, task demands and attention modulate the responses associated with stimulus repetition [19,20,39].

4.4. Repetition suppression

We found that repetition of target faces resulted in decreased amplitudes of the MEG signal, consistent with many functional brain imaging studies of repetition suppression [7,8,16,20,23,33,37]. At both latencies, the repetition of both neutral and fearful targets resulted in reduced MEG signals, although at the earlier latency (160–185 ms), neutral targets evoked stronger responses than fearful targets, whereas at the later interval (220–250 ms), fearful targets evoked stronger responses than neutral ones. In our prior fMRI study, we reported that repetitions of both neutral and fearful targets resulted in reduced hemodynamic responses, yet within all face-selective regions, the repetition suppression associated with fearful faces was stronger [22]. As the fMRI signal was averaged across several seconds from stimulus onset, we could not detect the differential responses at the early (170 ms) and late (230 ms) components. Thus, the high temporal resolution provided by the MEG technique enabled us to identify the dynamics of valence-dependent repetition suppression in extrastriate cortex. Interestingly, the repetition of distracters was also associated with reduced MEG responses. In our fMRI study, we found that only repetition of the behaviorally relevant items was associated with suppression. Due to the low temporal resolution of the fMRI technique and because the distracters, regardless of their valence, evoked negligible activation, we were unable to detect the repetition effects. The reduced MEG responses associated with the repetition of both targets and distracters provide empirical evidence for the predictive coding model, according to which learning is reflected by reduced prediction error [13].

Taken together, our data indicate that the processing of neutral and fearful faces in the visual cortex is not a mere bottom-up feature analysis, but, rather, a dynamic process that is modulated by cognitive factors such as task relevance, attention, valence and context.

Acknowledgments

We thank Richard Coppola, Tom Holroyd and Fred Carver from the NIMH MEG Core Facility for assistance with data collection and data analysis, Talma Hendler for assistance with the waveform analysis and Stefan Kiebel and Micah Murray for their valuable suggestions. This study was supported by the NIMH Intramural Research Program.

References