Pointing with the eyes: The role of gaze in communicating danger

Nouchine Hadjikhani a,b,c,*, Rick Hoge a,b, Josh Snyder a, Beatrice de Gelder a,d,*

a Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, 149 13th Street, Charlestown, MA 02129, USA
b Division of Health Sciences and Technology, Harvard-MIT, Cambridge, MA, USA
c Brain Mind Institute, EPFL, Switzerland
d Cognitive and Affective Neurosciences Laboratory, Tilburg University, P.O. Box 90153, Tilburg, The Netherlands

Accepted 24 January 2008
Available online 30 June 2008

Abstract

Facial expression and direction of gaze are two important sources of social information, and what message each conveys may ultimately depend on how the respective information interacts in the eye of the perceiver. Direct gaze signals an interaction with the observer but averted gaze amounts to “pointing with the eyes”, and in combination with a fearful facial expression may signal the presence of environmental danger. We used fMRI to examine how gaze direction influences brain processing of facial expression of fear. The combination of fearful faces and averted gazes activated areas related to gaze shifting (STS, IPS) and fear-processing (amygdala, hypothalamus, pallidum). Additional modulation of activation was observed in motion detection areas, in premotor areas and in the somatosensory cortex, bilaterally.

Our results indicate that the direction of gaze prompts a process whereby the brain combines the meaning of the facial expression with the information provided by gaze direction, and in the process computes the behavioral implications for the observer.

Keywords: Face processing; Gaze; Emotion; Fear recognition; Action representation

1. Introduction

Facial expression and direction of gaze are significant components of the information provided by a face. In the course of development, detection of gaze orientation plays an important role. Human and nonhuman primates interact with their caregivers and learn from observing the direction of gaze what objects to avoid. In primates, fear can be communicated through the mechanism of joint attention: young rhesus monkeys, initially unafraid of snakes, show fear after witnessing the fearful reaction of their parents to a snake (toy or real), suggesting that they are able to couple their parents’ fearful expression and their direction of gaze to learn that snakes are dangerous (Mineka, Davidson, Cook, & Keir, 1984).

Until recently, studies of gaze processing have used neutral faces, and investigated the attentional cues provided by gaze direction. Direction of gaze influences the level of activity in areas involved in processing faces, including the fusiform gyrus, the superior temporal sulcus and the intraparietal sulcus (George, Driver, & Dolan, 2001; Hoffman & Haxby, 2000; Pelphrey, Singerman, Allison, & McCarthy, 2003). But gaze direction may play a more complex role when it belongs to a face expressing a specific emotion (Adams, Gordon, Baird, Ambady, & Kleck, 2003; Adams & Kleck, 2003; Klucharev & Sams, 2004; Mathews, Fox, Yiend, & Calder, 2003; Sato, Yoshikawa, Kochiyama, & Matsumura, 2004); but see (Hietanen & Leppanen, 2003). A fearful facial expression with an observer-averted gaze is automatically recognized (Anderson,
Christoff, Panitz, De Rosa, & Gabrieli, 2003) as possibly indicating environmental threat (“there is a danger located where I am looking at”) and as requiring an adaptive response (“you need to avoid it”).

Several recent studies indicate gaze direction influences other brain areas besides the ones specifically associated with face recognition. To account for this, several groups have developed a distributed representation model of face perception (Bruce & Young, 1986; de Gelder, Frissen, Barton, & Hadjikhani, 2003; Haxby, Hoffman, & Gobbini, 2000; Haxby et al., 1994; Haxby et al., 1996; Hoffman & Haxby, 2000), in which different areas of the brain respond to different attributes of a face, such as identity ( fusiform gyrus, inferior occipital gyrus), gaze direction and recognition of action (superior temporal sulcus), and expression and/or emotion (orbitofrontal cortex, amygdala, anterior cingulate cortex, premotor cortex).

It is well known that the sight of a fearful face expression provides a very strong signal, however, a fearful face can be either empathy-evoking to the observer or threat-related, depending on its gaze direction. Facial expression of fear can be ambiguous because it may be unclear whether the emphasis is on communicating an experienced emotion to the observer and possibly provoking empathy, or on providing a danger signal to an observer with the goal of preparing him to act. To date there have only been very few studies that have manipulated separately the facial expression and the direction of gaze (Adams & Kleck, 2005; Ganel, Goshen-Gottstein, & Goodale, 2005). Two previous studies used behavioral measures (Adams & Kleck, 2003) and event-related brain imaging (Adams et al., 2003) to investigate the combined effect of gaze and facial expression. The behavioral study revealed a shorter reaction time for fearful faces looking away compared with fearful faces looking at the observer. The brain imaging study revealed increased amygdala activation for stimuli consisting of angry faces with an averted gaze and fearful faces with a direct gaze, that were ambiguous in terms of their significance as threat to the observer. No other areas were reported.

It is reasonable to expect that different brain networks are involved as a function of the direction of gaze a facial expression. To address this issue, we manipulated direction of gaze in fearful faces. Our goal was to test a single very specific prediction related to the combined perception of fear and averted gaze and its impact on action readiness. Our hypothesis was that a fearful face expression with averted gaze signals a danger in the environment and may trigger activity in brain areas involved in characteristic adaptive action associated with fear such as preparing to flight. If fearful faces with averted gaze do indeed signal danger in the environment (threat-related), then they should prompt more premotor and motor activity than faces with directed gaze where the emphasis is more on communicating the emotion and triggering empathy in the observer (empathy-evoking). We tested the hypothesis that threat-related fearful face would modulate activation in areas involved in stimulus detection, in fear processing, and in preparation for action. We did not use neutral faces in this study, because in the context of an emotional expression study, neutral faces suffer from carry-over effects and acquire an unintended emotional significance. To specifically address our hypothesis related to fear processing we limited our paradigm to fearful expression of emotion.

2. Materials and methods

Stimuli were taken from the NimStim Emotional Face Stimuli database (http://www.macbrain.org/faces/index.htm#faces), a set of over 600 face images, consisting of 16 expressions posed by 45 professional actors. All expressions have been validated, and only faces that received more than 90% agreement in the validation were used. Eight fearful faces (four females) were selected. We used Adobe Photoshop 8.0 to alter gaze direction towards the left and the right and downwards (Fig. 1). Grayscale stimuli were shown in an AB-blocked presentation of 8 cycles of 24 s. Blocks of averted and direct gaze alternated every 24 s. Within each block, images were presented in a random order every 1.5 s for 300 ms with a 1200 ms blank-screen interval between stimuli. A fixation cross was present on the screen between the stimuli.

Subjects viewed images passively and were instructed to observe the images attentively and maintain fixation. No other task was required, because of possible interference with recognition of the emotion (Lange et al., 2003).

3. Imaging

Structural and functional MR images of brain activity of eight participants (3 males, age 29 ± 6 years) were collected in a 3T high-speed echoplanar imaging device (Allegro, Siemens) using a phased-array head coil. Participants all had normal or corrected-to-normal vision. Informed written consent was obtained before the scanning session, and the Massachusetts General Hospital Human Studies Committee under Protocol #2002P-000228 approved all procedures. Structural images were collected with the following parameters: sagittal MPRAGE: 128 slices, 1.33 mm isotropic voxels, repetition time (TR) = 2730 ms, echo time (TE) 3.44 ms, flip angle 7°. Functional image volumes consisted of 40 contiguous 3 mm thick slices covering the entire brain (TR = 3000 ms, 3.125 mm by 3.125 mm in plane resolution, 128 images per slice, TE = 30 ms, flip angle 90°, matrix = 64 × 64).

4. Image statistics

Image analysis was conducted using the NeuroLens analysis package (Hoge & Lissot, 2004) (http://www.neurolens.org, version 1.3). All functional EPI and structural scans were first converted from DICOM to MINC format using NeuroLens. Functional image series were motion corrected to the third frame in each series within Neuro-
Lens using a hardware-accelerated module based on source code from AFNI’s 3dvolreg module (Cox & Jesmanowicz, 1999). Next, each image series was spatially smoothed in 3D with a 6 mm FWHM 3D Gaussian kernel. Intensity normalization was also applied to set the mean intra cranial signal of each EPI series to a standard value of 10,000. The signal at each voxel in the motion-corrected, smoothed, and intensity normalized image series was then fit with a linear model consisting of a regressor representing the periods of gaze diversion, plus four regressors containing the terms of a third order polynomial to represent the baseline EPI signal (in this case corresponding to direct gaze) plus low frequency signal drift. Volumes containing the estimated effect size and associated standard error for the primary contrast (gaze diversion vs. direct gaze) at each voxel were then registered to a standard space based on the MNI Talairach template (Collins, Neelin, Peters, & Evans, 1994). This spatial normalization was performed in Neuro-Lens by fitting the third frame of each individual’s EPI series to an EPI target brain and applying the resultant transformation to the computed effect size and standard error volumes for that individual. The EPI template was generated by registering whole-brain EPI scans from forty subjects (using the same pulse sequence and parameters as the present study) to the MNI standard space and averaging. This spatial normalization was performed in Neuro-Lens based on the method described by Worsley et al. (2002). This procedure combines fixed and estimated random effects variance in proportions required to achieve a user-specified number of degrees of freedom (in this case 100). The modeled group effect size and standard error were then divided to produce a volumetric map of \( T \) statistic with 100 degrees of freedom. Based on this \( T \) statistic volume, a map of \( p \) values was computed based on the \( T \) value at each voxel. The computed significance values were displayed as the negative base ten logarithm of each voxel’s \( p \) value, which produces a low background value while highlighting areas of elevated significance. The map of \(-\log(p)\) was then thresholded using an amplitude cutoff of 2.0 (corresponding to \( p = 0.01 \)), and a cluster size threshold of 0.16 ml, which requires that 20 contiguous voxels must all exceed the specified amplitude threshold to be included. This size threshold, plus restriction of the search volume to the intracranial space, reduces the effective \( p \) value for the minimal accepted cluster to less than \( 10^{-5} \). The thresholded \( p \) map was then sampled on the cortical surface of an individual subject using on the inverse coordinate transformation between this individual’s native space and the group Talairach space. Cortical surface files were generated using FreeSurfer (Dale, Fischl, & Sereno, 1999; Fischl & Dale, 2000; Fischl, Sereno, & Dale, 1999) (http://surfer.nmr.mgh.harvard.edu) and loaded in Neuro-Lens, which was then used to interpolate the values in the group \( T \) Statistic volume (transformed to the individual’s space) at the vertex locations of the cortical surface.

5. Results and discussion

Fearful faces with the gaze averted compared with the same faces directly gazing at the observer increased activation in areas belonging to face and emotion processing networks. Areas of increased BOLD signal were found in gaze processing areas (superior temporal sulcus, intraparietal sulcus); in face identification areas (fusiform gyrus, inferior occipital gyrus); in areas involved in rapid stimulus detection (left amygdala, visual area MT+); in areas involved in fear processing (left amygdala, hypothalamus, pallidum, dorsomedian nucleus of the thalamus); and in areas involved in motor preparation (premotor and motor cortices, superior parietal lobule) (Figs. 2, 3 and 5; Table 1).

Fig. 1. Example of the stimuli used in this experiment. (A) The actor looking with averted gaze and a fearful expression, and (B) The same actor looking directly at the observer. The left stimulus was generated by moving the direction of the eyes with Photoshop 8.0.
The reverse comparison (fearful direct > fearful averted) resulted in activation in the posterior occipital cortex, in the foveal representation of extrastriate areas (Fig. 4).

5.1. Gaze and face processing network

As expected, faces with diverted gaze—indicating a potential danger by “pointing with the eyes”—activated areas that have been previously involved in gaze perception, including the superior temporal sulcus, and the intraparietal sulcus (Grosbras, Laird, & Paus, 2005; Hoffman & Haxby, 2000; Pelphrey, Viola, & McCarthy, 2004; Pelphrey et al., 2003; Puce, Allison, Bentin, Gore, & McCarthy, 2004; Pelphrey et al., 2003; Puce, Allison, Bentin, Gore, & McCarthy, 2004; Pelphrey et al., 2003; Puce, Allison, Bentin, Gore, & McCarthy, 2004; Pelphrey et al., 2003; Wicker, Michel, Henaff, & Decety, 1998). Modulation by gaze was also found in areas involved in face processing (fusiform gyrus, inferior occipital gyrus) as had already been shown in studies using neutral faces with shifting gaze (George et al., 2001; Hoffman & Haxby, 2000; Pageler et al., 2003) (Table 2).

5.2. Stimulus detection and emotional network

Activity in the stimulus detection system (amygdala; area MT+) supports the model developed by several investigators (Adolphs, 2002; de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999; LeDoux, 1992; Morris, Ohman, & Dolan, 1998) according to which a rapid system based on coarse visual analysis sustains orientation towards novel stimuli and detection of potentially dangerous signals.

The amygdala is a brain structure essential to the perception of fear (Amaral, 2003; LeDoux, 2003), and activation in the amygdala is reliably produced by presentation of biologically relevant sensory stimuli, and stimuli that predict threat (Whalen, 1998). We observed left but not

Fig. 2. Regions involved in the visual analysis of facial expressions of fear with averted vs. direct gaze. Lateral and ventral views of a 3D reconstruction of a single subject brain, with a projection of group data activation. Areas that were activated in the mixed-effect group average by faces with a fearful expression and an averted gaze are displayed in red to yellow ($p < 0.05$). Activation for averted fear was found in areas involved with gaze detection (STS, IPS), with face identification (fusiform gyrus, inferior occipital gyrus), with action representation (superior parietal lobule, premotor and motor cortices) as well as in visual area MT and in the somatosensory cortex. (For interpretation to colours in this figure, the reader is referred to the web version of this paper).

Fig. 3. Sagittal (panel A) and horizontal slices (panels B) showing subcortical regions involved in the visual analysis of facial expressions of fear with averted vs. direct gaze. Areas that were activated in the mixed-effect group average by faces with a fearful expression and an averted gaze are displayed in red to white ($p < 0.05$) on one individual MP-RAGE. A signal of danger in a fearful face gazing away, the hypothalamus and the thalamus DM (panel A) as well as the amygdala (panel B). (For interpretation to colours in this figure, the reader is referred to the web version of this paper).
right amygdala activation, consistent with the findings of Morris et al. (1998) who demonstrated left amygdala activation for consciously recognized stimuli, and right amygdala involvement for unseen stimuli. Laterality effects in the amygdala have however not been consistently obtained so far (Zald, 2003). Left amygdala activation for ambiguous threat-related stimuli was observed previously (Adams et al., 2003). However, in that study (which only reported results for the amygdala), activation was more important for fearful faces looking directly at the observer and was interpreted as being a result of threat ambiguity, while we observed here clear amygdala activation for fearful faces gazing away.

The discrepancies between these two studies may reflect several differences between them. First, there is a difference in the stimuli used. Adams et al. study used several exemplars from diverse emotion libraries (the Montreal Set of Facial Display of Emotions (Beaupr, Cheung, & Hess, 2000); the classical Ekman faces (Ekman & Friesen, 1976); custom-made stimuli Young Adult Facial Display made by (Adams & Kleck, 2001); a set from (Kirouac & Doré, 1984) in which gaze was displaced laterally. In our study, we used validated facial expression from the NimStim database and selected the intense fear expressions on which there was more than 90% agreement; in addition gaze was displaced laterally and towards the ground as if indicating the presence of a danger (snake, spider) coming from the floor. It is therefore possible that our averted stimuli felt more like indicating the presence of an environmental danger. In addition, one needs to note that looking down increases the amount of sclera visible over the iris, which has been shown to trigger amygdala response to fear even subliminally (Whalen et al., 2004). The second difference between our study and that of Adams et al. (2003) is to be found in the experimental design: we used a block-design in our experiment, whereas they used an event-related design; it is possible that response to and processing demands associated with threat-related ambiguity is reduced by cue redundancy. In addition, we used rapid presentation of faces (300 ms) compared to 2.5 s in the Adams et al. study. One may speculate that longer exposure durations trigger more in depth processing demands and increase sensitivity to ambiguity, whereas perhaps initial responding is more cued only highly salient threat cues. This, exposure duration may be very relevant issue in explaining these differences. A third point is to be found in recent data from Graham and Labar (Graham & LaBar, 2007) who reported that when the emotional expression is very intense and easily recognizable it is processed more quickly than gaze direction such that all emotional expressions tend to be processed faster when coupled with direct gaze. However, they also reported that when gaze direction
is processed more quickly than expression, the two interact (i.e., the influence of gaze on emotion varies as a function of the type of emotion displayed), which explains conditions under which emotions like fear and sadness are processed more efficiently when combined with averted gaze. In our study, because we used very intense, extremely recognizable expressions (ones with 90% agreement) it is possible that fear combined with direct gaze might be actually more readily and efficiently processed than fear combined with averted gaze. If this is the case, then the findings of the current study are not necessarily in conflict with the previous findings. In other words, both studies then would be reporting more activation for those pairings that require longer to process. We note also that in their behavioral study, Adams and Kleck (2003) reported data that are compatible with our imaging data, namely a decreased reaction time to stimuli of fearful faces with averted gaze compared with fearful faces with direct gaze, that can be interpreted as an amygdala-mediated response. Overall though, it remains difficult to postulate a direct link between behavioral and brain activation data also when the influence of task and stimulus presentation is not fully controlled.

Activations in the other areas involved in fear detection are consistent with previous studies. Concomitant activation of the left amygdala and the left pallidum for fearful faces has been reported previously. The hypothalamus receives direct connections from the central nucleus of the amygdala (LeDoux, 2000) and is responsible for the behavioral manifestations of fear such as tachycardia, increased galvanic skin response, paleness, pupil dilation and blood pressure elevation. The dorsomedial thalamus receives input from the amygdala and is a key structure for visual input into the prefrontal cortex (Jones, 1981; Krettek & Price, 1977a, 1977b). It belongs to the circuitry activated by fear-inducing stimuli (Sewards & Sewards, 2002, and in the animal model activated neurons were recorded in stimuli such as predator exposure (Canteras, Chiavegatto, Valle, & Swanson, 1997).

Activity in the somatosensory cortex is consistent with the somatic marker hypothesis, whereby feedback from the autonomic, musculoskeletal and endocrine system to the somatosensory areas plays a crucial role in emotion perception (Damasio, 1996) and was shown to be active together with the left amygdala in a paradigm of visual fear acquisition in healthy controls (Birbaumer et al., 2005).

Taken together, we found that important elements of the fear network were positively modulated when a fearful
face gazing towards an unseen danger was compared to a fearful face looking directly at the observer. This implies that the brain’s perception of fear itself can be modulated by gaze direction.

5.3. Premotor and motor network

Activation in the premotor cortices together with the caudate nucleus may be related to a preparatory defense response (Fischer, Andersson, Furmark, Wik, & Fredrikson, 2002; Knight, Cheng, Smith, Stein, & Helmstetter, 2004), and are consistent with the results of our previous study, where we found premotor and motor activation when comparing bodies expressing fear with bodies performing neutral actions (de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004).

5.4. Conclusion

The network of areas observed in the present study of combined fear and gaze direction resembles closely those found in our previous studies looking at brain activation elicited by viewing fearful bodies (de Gelder et al., 2004; Hadjikhani et al., 2004). Observing fearful faces “pointing to a danger” with their eyes compared with fearful faces looking at the observer modulates cortical and subcortical areas sustaining detection and rapid orientation, as well as fear perception and preparation for action. A fearful expression in somebody looking away may signal imminent danger for both the actor and the observer, and consequently it will activate a circuit leading to an adaptive fear response to danger.

Our findings indicate that viewing facial expressions may trigger an active process of gaze interpretation allowing the brain to elaborate not only the meaning of the facial expression, but also what the implications are for the observer.

Acknowledgments

This research was supported by NIH Grant RO1 NS44824-01 and Swiss National Foundation Grant PPOOB—110741 to Nouchine Hadjikhani. We thank Reginald Adams and an anonymous reviewer for their comments on our manuscript.

References


