

Quantitative prediction of perceptual decisions during near-threshold fear detection

Luiz Pessoa* and Srikanth Padmala

Department of Psychology, Brown University, Providence, RI 02912

Edited by Leslie G. Ungerleider, National Institutes of Health, Bethesda, MD, and approved February 25, 2005 (received for review January 24, 2005)

A fundamental goal of cognitive neuroscience is to explain how mental decisions originate from basic neural mechanisms. The goal of the present study was to investigate the neural correlates of perceptual decisions in the context of emotional perception. To probe this question, we investigated how fluctuations in functional MRI (fMRI) signals were correlated with behavioral choice during a near-threshold fear detection task. fMRI signals predicted behavioral choice independently of stimulus properties and task accuracy in a network of brain regions linked to emotional processing: posterior cingulate cortex, medial prefrontal cortex, right inferior frontal gyrus, and left insula. We quantified the link between fMRI signals and behavioral choice in a whole-brain analysis by determining choice probabilities by means of signal-detection theory methods. Our results demonstrate that voxel-wise fMRI signals can reliably predict behavioral choice in a quantitative fashion (choice probabilities ranged from 0.63 to 0.78) at levels comparable to neuronal data. We suggest that the conscious decision that a fearful face has been seen is represented across a network of interconnected brain regions that prepare the organism to appropriately handle emotionally challenging stimuli and that regulate the associated emotional response.

decision making | emotion | functional MRI

The brain uses sensory information to make decisions that guide behavior. Currently, there is considerable knowledge, for example, about how contour, depth, and motion information is processed by early visual areas. In contrast, relatively little is known about how perceptual decisions are made. Even simple perceptual decisions, such as determining the presence or absence of a sensory stimulus, are not well understood. A fundamental goal of cognitive neuroscience is to explain how mental decisions originate from basic neural mechanisms.

In non-human primates, Newsome, Movshon, and colleagues (1, 2) have pioneered the investigation of the neural correlates of perceptual decisions. Their experiments have probed the mechanisms associated with deciding the direction of motion of an array of moving dots, only a fraction of which move coherently at a time. Newsome *et al.* (1, 2) showed that the activity of directionally selective neurons in middle temporal visual area (MT) could account for the psychophysical performance in a forced-choice direction discrimination task (1, 2). Related studies have quantified the link between performance and cell activity during depth perception (3) and the perception of structure from motion (4). Little is known, however, about the involvement of nonsensory regions in perceptual decisions. Previous neurophysiology studies typically have investigated one, and in some cases two [e.g., visual area 1 (V1) and MT], sensory processing regions (but see ref. 5).

Decision-making processes have been investigated with neuroimaging, too. Binder *et al.* (6) operationally defined decisions in an auditory object identification task in terms of reaction time (RT) and correlated RT with blood-oxygenation-level-dependent signals to reveal brain regions linked to decision processes. A more direct approach was taken by Heekeren *et al.* (7), who showed that activity in the dorsolateral prefrontal cortex covaried with the difference signal between face- and

house-responsive regions in ventral temporal cortex in a face/house categorization task. Their results are important, because they suggest the existence of a general mechanism for perceptual decision making in the human brain based on the comparison of the outputs of different pools of selectively tuned neurons, as initially suggested in single-cell recording studies (8, 9).

The goal of the present study was to investigate the neural correlates of decision making in the context of emotional perception. Although several brain regions have been implicated in the processing of emotion-laden visual stimuli (10), it is currently unknown whether they are also involved in perceptual decisions. To probe this question, we investigated how fluctuations in functional MRI (fMRI) signals were correlated with behavioral choice during a near-threshold fear detection task. Such correlation was quantified by adapting methods from single-cell physiology that allowed us to compute so-called choice probabilities (CPs) (8) in a voxel-wise manner. During our task, an initial neutral, happy, or fearful face was presented for 67 ms and immediately followed by a neutral face, which served as a mask (also shown for 67 ms; Fig. 1). In each trial, subjects indicated whether they had seen a fearful face. Like tasks studied in monkeys involving two-choice outcomes (e.g., left/right), the present study can also be viewed as a two-choice task (fear present/fear absent).

Methods

For additional methods, please consult see *Supporting Text*, which is published as supporting information on the PNAS web site.

Stimuli and Procedure. Each trial began with a white fixation cross shown for 1,000 ms on a black background, followed by a green fixation shown for 300 ms on a black background, followed by the presentation of a fearful, happy, or neutral target face, immediately followed by a neutral face, which served as a mask (Fig. 1). The identities of the target and mask stimuli were always different. Faces subtended 4° of visual angle. The duration of the target face was 67 ms, and the mask face was 67 ms. Target presentation durations were confirmed by using a photodiode and an oscilloscope. Subjects were instructed that the stimulus would always comprise two faces and to respond “fear” if they perceived fear, however briefly. After the presentation of each face pair, subjects indicated fear or no fear with a button press. On each trial, subjects also rated the confidence in their response on a scale of 1 to 4 (low to high confidence). Happy faces were included to more closely match fearful faces in terms of low-level features, such as brightness around the mouth and eye regions, because both fearful and happy faces tend to be brighter than neutral ones in these regions. Thus, the inclusion of happy faces precluded subjects from using a strategy of detecting fearful

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: fMRI, functional MRI; ROC, receiver operating characteristic; PCC, posterior cingulate cortex; MPFC, medial prefrontal cortex; IFG, inferior frontal gyrus; ACC, anterior cingulate cortex; CP, choice probability; MT, middle temporal visual area.

*To whom correspondence should be addressed. E-mail: pessoa@brown.edu.

© 2005 by The National Academy of Sciences of the USA

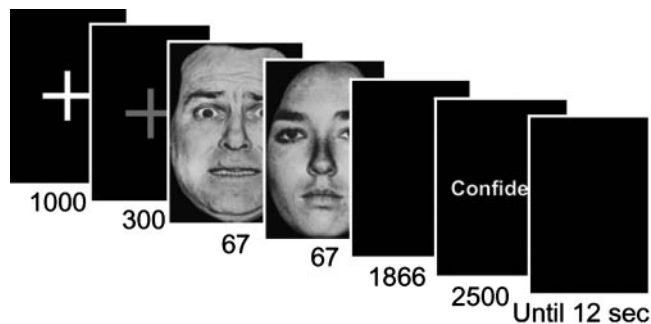


Fig. 1. Experimental paradigm. For every trial, after the target-mask face pair, subjects indicated whether they saw a fearful face and then indicated their confidence in the response. The initial target face was a fearful, happy, or neutral face, and the mask was always a neutral face. Trials occurred every 12 s in a slow event-related design. The 300-ms fixation cross was actually green in the experiment.

faces by simply using low-level cues. The inclusion of happy faces also precluded subjects from adopting a strategy of indicating fear whenever facial features deviated from those of a neutral face. The total trial duration was ≈ 12 s (11,920 ms). Each subject performed 160 trials.

fMRI Data Analysis: CP Maps. Because we were interested in the neural correlates of the perceptual decision of fear, in choice-related analyses, we considered only high-confidence trials (values of 3 and 4); these comprised 75% of the total number of trials. In this manner, we largely eliminated guess trials.

To quantify the link between behavioral choice and fMRI signals, we computed individual CP maps. Thus, we sorted trials according to the two behavioral responses, namely, trials in which the subject reported seeing fear (hits plus false alarms) and trials during which the subject reported not seeing fear (correct rejects plus misses); these may be thought of as comprising the signal and noise distributions in signal detection theory. For high-confidence trials, on average, there were 51 reported fear trials and 60 reported no-fear trials per participant. Response strength was indexed by the average of the raw fMRI signal (after linear detrending) at times 3, 6, and 9 s relative to stimulus onset. Receiver operating characteristic (ROC) curves were constructed by using response strength and trial type in a manner analogous to that done in monkey physiology (8, 11). To create the ROC curve, we first determined the distribution of fMRI signal amplitudes as a function of trial type. The entire ROC curve was calculated by sliding a criterion threshold T value over the entire range of fMRI amplitudes, i.e., by determining $P(a > T)$, where a is the fMRI signal amplitude, both for reported fear-present and -absent trials. Thus, to create the ROC curve, for a range of criterion response levels, we plotted on the x axis the proportion of reported fear-absent trials that exceeded the criterion, and on the y axis the proportion of reported fear-present trials that exceeded the same criterion.

The area under the ROC curve (A') gave the CP value for the voxel. To determine the reliability of a CP, we used a permutation test (12) to compare it with the chance expectation of 0.5. The critical value of the CP was calculated after we randomly reassigned each fMRI response to either a reported fear or a reported no-fear behavioral choice, thus disrupting any correlation between fMRI signal and choices, while leaving the distributions of fMRI response and behavioral choice unchanged. In this manner, for each voxel, we generated the distribution of CPs expected in the absence of any association from 2,000 permutations (i.e., reassignments). Observed CP values were deemed statistically significant if they lay outside the central 95% of the distribution. In this manner, we created

individual CP maps that displayed those voxels that reached significance based on the critical values computed above. Only choice values significantly >0.5 were considered, i.e., fear-present choices (see *Supporting Text* for a discussion of values significantly <0.5).

CPs were first computed by using only fear-neutral and neutral-neutral target-mask pairs. Given these trials types and behavioral responses, signal and noise distributions were determined as defined above and CP computed. Because different physical stimuli were used to compute CP, it is conceivable that stimulus differences could have contaminated the measure (see also *Choice responses for constant physical stimuli* in *Supporting Text*). Thus, we computed CPs in a separate manner by using happy targets rather than neutral ones, i.e., by using only fear- and happy-neutral target-mask pairs. To ensure that CPs reflected fear detection and not some low-level feature of the display, we determined the conjunction of the CP maps obtained in these two ways. By considering voxels that exhibited significant choice values in both contexts, we targeted voxels that were linked to deciding fear, irrespective of the type of nonfear target stimulus (neutral or happy). To assess the brain regions of greatest overlap in CP values and to display group results, we generated group CP maps. To create such maps, we smoothed each individual map with an 8-mm Gaussian filter (full width at half maximum) and averaged them. We then determined the conjunction of the choice maps obtained by separately considering neutral and happy-face trials as indicated above; the conjunction was obtained by determining the intersection of the two associated masks.

Because it was essential to formally assess the reliability of the individual CP maps in terms of inferences regarding the population (and not just in terms of the specific sample tested), we performed a standard two-stage random-effects analysis. The first stage corresponded to the individual CP maps, which were then submitted to a one-sample t test that evaluated whether the observed choice values significantly differed from 0.5 on a per-voxel basis. Because random-effects analysis may be fairly conservative (13), we used a threshold of $P < 0.001$ (uncorrected).

To illustrate the reliability of the individual participant's CP maps, for each brain region exhibiting strong fMRI signals predictive of behavioral choice independently of stimulus-driven signals, we selected the voxel coordinate corresponding to the peak CP value as the center of an 8-mm-radius spherical region. The values shown in Fig. 3 and Table 1 are the averaged-across-subject peak choice values obtained in the context of neutral and happy trials.

There are multiple ways to compare neuron- and fMRI-based CP values. In monkey studies, mean choice values are computed for all neurons that are carefully studied, which typically depends on meeting several criteria (e.g., exhibiting disparity tuning). In fMRI, because signal-to-noise ratios vary considerably across the brain, we did not average across all voxels; instead, we considered only those voxels with statistically significant CPs to determine the range of values.

The overlap between stimulus-driven, performance-related, and choice-related activation was formally assessed by computing their voxel-wise intersection by using masks. When any overlap with choice-related voxels was observed, they were marked with a "No" in Table 1. When no overlap was observed, the regions were deemed choice related (Table 1, "Yes").

Results

For both the behavioral and fMRI results that follow, we considered only high-confidence trials. Analysis of the behavioral data showed that average accuracy across subjects for masked fear detection was 86% correct. Analysis of reaction times revealed no significant differences between reported fear-

Table 1. Brain regions exhibiting significant CP values at the group level

Brain region	Montreal Neurological Institute coordinates			Peak CP*	Choice related†
	x	y	z		
Right fusiform gyrus	44	−50	−23		No
Right superior temporal sulcus	56	−45	23		No
PCC	0	−32	32	0.68	Yes
ACC	2	26	42		No
Right IFG	44	19	−10	0.67	Yes
Left insula	−60	2	4	0.70	Yes
MPFC	−2	50	12	0.69	Yes
Right Anterior IFG	44	46	−4	0.68	Yes

*CP values were the average of values obtained in separate analyses that considered neutral-neutral and happy-neutral trials as catch trials.

†Only regions whose responses were not related to the physical characteristics of the stimulus (fear containing vs. neutral) or behavioral accuracy (correct vs. incorrect) were deemed to encode behavioral choice per se.

present [676 (mean) \pm 148 (SD) ms] and fear-absent trials (726 ± 177 ms); $P > 0.05$, paired t test.

To investigate how fMRI signals were related to behavioral choice, we probed how the trial-by-trial variability in fMRI signal amplitude was correlated with the choices the subject made, independently of the effects of visual stimulation. In other words, in the fear-detection task used, could fMRI amplitude predict whether the subject responded fear or no fear on a trial-by-trial basis? We used a method based on signal detection theory, which is analogous to ROC analysis and which has been used in monkey physiology to link cell responses to behavioral choice (4, 11, 14). This method gives the probability that a so-called ideal observer, given access only to the fMRI amplitude in a trial, would be able to identify accurately which behavioral response was made in that particular trial (fear present or fear absent). Although in monkey physiology research, spike data are used as a measure of response, in the present case, fMRI amplitude was used as an index of response strength (see *Methods*). Because the task was to indicate whether fearful faces were shown, we discuss CPs linked to fear-present reports (see *Supporting Text* for further analyses). A value of 0.5 represents chance performance, and a value of 1.0 represents a perfect association between fMRI signals and behavioral response, i.e., fMRI amplitude would perfectly predict choice. Following previous work (8), we term this value CP.

To probe regions involved in deciding whether fearful faces were perceived, we computed CPs in a voxel-wise manner. CP maps were computed for each individual and combined into a group map. To ensure that CPs truly reflected fear detection and not some low-level feature of the display, we determined ROCs in two separate contexts, by using only neutral-neutral trials as nontarget catch trials and separately only happy-neutral trials as

catch trials. The conjunction of such maps was then determined and comprised our final results. In this manner, only voxels that predicted choice independently of the type of nontarget trial were considered to represent behavioral choice (see *Methods*). Moreover, we considered only high-confidence trials, which more directly reflect perceptual decision (thus largely eliminating guess trials). The resulting group CP maps revealed a set of brain regions that reliably predicted the subjects' responses (Fig. 2). The most strongly predictive regions included posterior cingulate cortex (PCC), medial prefrontal cortex (MPFC), right inferior frontal gyrus (IFG; this site encroached into the anterior insula), a more anterior site in the right IFG, and the left insula (see Table 1 for all regions and *Supporting Text* for discussion of motor-related activations). To formally assess the reliability of the group results, we performed a standard two-stage random-effects test of the CP values. The first stage corresponded to the individual CP values, which were then submitted to a one-sample t test that evaluated whether the observed values significantly differed from chance (0.5) on a per-voxel basis. With the exception of the right superior temporal sulcus, all regions were statistically significant.

To investigate the consistency of choice-related signals, we inspected individual choice maps (Table 1). To do so, individual maps were thresholded at a critical value obtained by a permutation test (12) to compare it with the chance expectation of 0.5. The averaged-across-subjects peak CPs in a set of regions (Table 1) ranged from 0.67 to 0.70. Fig. 3 displays CP maps and ROC curves for two representative individuals for the MPFC and the right IFG. The averaged-across-subjects peak CP was 0.69 for the MPFC (range, 0.64–0.78) and 0.67 for the right IFG (range, 0.63–0.73). Thus, the above results illustrate that results ob-

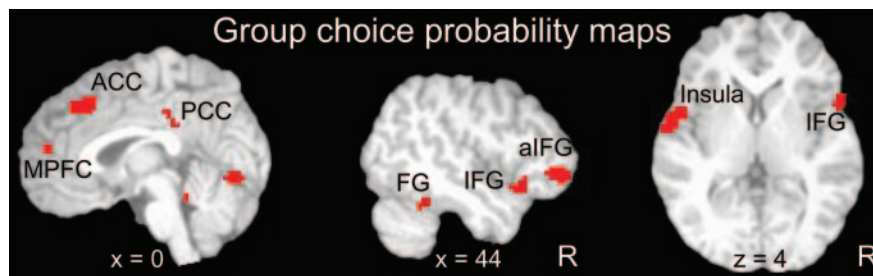


Fig. 2. Group CP maps illustrate areas that significantly predicted behavioral choices associated with reporting that fearful faces were seen according to trial-by-trial fluctuations in the fMRI signal. Note that the IFG site (on the right) abuts the insula. Group maps ($n = 8$) were obtained by averaging individual maps (see *Methods*) and were thresholded at a level to illustrate regions that predicted performance at the group level as indicated by random-effects analysis (Table 1). Slice coordinates are according to the Montreal Neurological Institute brain template. a, anterior; FG, fusiform gyrus.

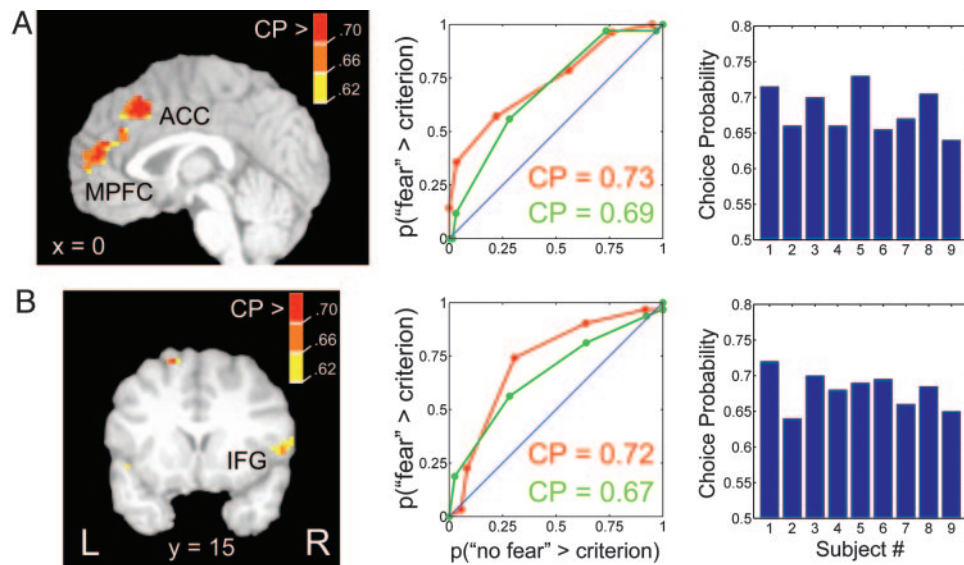


Fig. 3. Individual CP data for the (A) MPFC and (B) right IFG. (Left) CP maps from two representative individuals. Maps were thresholded according to a per-voxel critical value for statistical significance ($P < 0.05$) obtained by a permutation test. The color bar indicates CP of individual voxels, which were the average value from the separate analyses with neutral and happy trials (see *Methods*). (Center) Choice-related ROC curve for the peak voxel of the activation (Left) when considering neutral trials only (red) and happy trials only (green). The x axis corresponds to the probability that fMRI magnitude in no-fear-response trials exceeds a criterion response magnitude; the y axis corresponds to the probability that fMRI amplitude in fear-response trials exceeds the same criterion. The ROC curve is determined by sliding the criterion value. The CP corresponds to the area under the ROC curve. The blue line ($y = x$) represents chance performance (area = 0.5), and departures from it indicate the possibility of distinguishing between the two trial types. In particular, ROC curves above the identity line indicate that responses during fear-response trials were consistently greater than responses during no-fear-response trials. Peak MPFC voxel coordinates: red and green curves, $x = 0$, $y = 24$, and $z = 42$; right IFG: red curve, $x = 49$, $y = 17$, and $z = 3$; green curve, $x = 53$, $y = 15$, and $z = 3$. (Right) Distribution of peak CP values for individual subjects for the (A) MPFC and (B) right IFG regions. Again, values were averaged from the analyses with neutral and happy trials.

tained at the group level were consistently observed across subjects.

If fMRI signals in a brain region encode behavioral choice, then responses evoked during trials in which subjects reported seeing fear should be equivalent to one another. Thus, similar responses should be observed during hits (in which a target fearful face is correctly reported as fearful) and during false alarms (in which a nontarget stimulus is reported as fearful). Likewise, responses evoked for trials in which subjects reported not seeing fear should be equivalent to each other. Thus, similar responses should be observed during correct rejects (in which a nontarget stimulus is correctly reported as nonfearful) and during misses (in which a target fearful face is incorrectly reported as nonfearful). Deviations from this pattern would indicate that the region might be more closely tied to physical stimulus properties (containing or not fear) or performance (correct vs. incorrect trials); see below. Fig. 4 displays the responses to hits, false alarms, and misses relative to the responses to correct rejects for the MPFC and right IFG regions. Overall, the pattern is consistent with the encoding of behavioral choice (see Fig. 5, which is published as supporting information on the PNAS web site, and *Supporting Text* for further analyses).

We also contrasted stimuli containing fearful targets (fear-neutral pairs) to those containing only neutral faces (neutral-neutral pairs). Several brain regions evoked greater responses during trials containing fearful targets than during trials containing only neutral faces (Table 2, which is published as supporting information on the PNAS web site). These regions included bilateral fusiform gyrus (at both more posterior and more anterior sites), bilateral superior temporal sulcus, and a site just dorsal to the right amygdala, which we interpret as ventral striatum. Such regions have been consistently reported during the perception of emotional faces in general, and fearful faces in particular (15). We also observed stronger activation during fear-containing trials in several other brain regions, including

bilateral anterior insula, the anterior cingulate cortex (ACC), and the right middle frontal gyrus. The overlap between stimulus-driven and choice-related activation was formally assessed by computing their voxel-wise intersection (by using masks). Although there was some overlap (regions marked with a “No” in Table 1), the PCC, MPFC, right IFG (at both sites), and left insula were not driven by stimulus differences.

The next question we addressed was whether the areas that reliably predicted behavioral choice also predicted performance accuracy. To determine the regions that predicted task

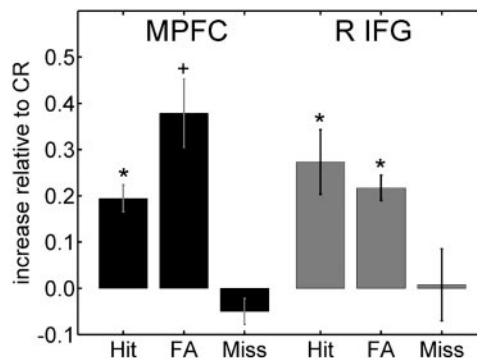


Fig. 4. Responses to hits, false alarms (FA), and misses, relative to correct reject (CR) responses. For both regions, hits and FA evoked similar responses that were stronger than responses evoked by misses and CR; the latter two trial types evoked similar responses. For both the MPFC and the right IFG, the pattern is thus consistent with the encoding of behavioral choice. An asterisk indicates $P < 0.05$ in a paired t test comparing that condition with miss trials; a plus sign indicates $P = 0.06$ in the paired t test comparing that condition with miss trials. Note that no other pairwise differences were observed (e.g., no significant differences were observed between hits and false alarms for the MPFC). Error bars are standard error of the mean pooled across subjects.

accuracy, we contrasted correct vs. incorrect trials. The resulting map revealed only two regions, bilateral anterior fusiform gyrus (at $y = -45$) and the ventral right amygdala/entorhinal cortex ($z = -29$; Fig. 6, which is published as supporting information on the PNAS web site). Thus, regions that predicted behavioral choice were distinct from those that predicted task accuracy.

Discussion

In the present paper, we computed whole-brain choice maps, which allowed us to characterize the distribution of CPs across regions and relate it to stimulus-evoked and performance-driven responses. Our results revealed a partial dissociation among regions that are sensitive to stimulus differences or performance and regions that are predictive of behavioral choice. We observed four regions that reliably predicted behavioral choice: PCC, MPFC, right IFG (at two sites), and left insula, which were neither sensitive to physical differences nor predictive of task accuracy. Thus, we suggest that these regions may have a role in perceptual decision making that is independent of visual stimulation and task accuracy (see *Supporting Text* and Figs. 7 and 8, which are published as supporting information on the PNAS web site, for further control analyses). The peak of the MPFC site was located in the most anterior aspect of the ACC and overlapped with a putative affective division of the ACC (16), which is connected to several emotional processing regions, including the amygdala, periaqueductal gray, nucleus accumbens, and hypothalamus, and has outflow to autonomic and visceromotor centers (17). Thus, the MPFC is well positioned to modulate brain responses according to the emotional significance of the incoming stimulus. The PCC is involved in emotion (18) and, in particular, the present finding that PCC signals are predictive of behavioral choice during the detection of fearful faces is consistent with a recent proposal that this region is involved in establishing an interface between attention and motivation (19). Interestingly, the PCC is also interconnected with the ACC, including its more anterior portion (20). The IFG has been proposed to be part of a network of regions that process emotionally arousing visual stimuli (21) and is connected to the amygdala and the more anterior portions of the ACC. Finally, the insula is a complex structure with important autonomic and other limbic functions, which is reciprocally connected with the amygdala (22). Overall, the PCC, MPFC, right IFG, and left insula may be part of a network of brain regions involved in emotional processing that participate in the decision that a fearful face has been seen.

Significant CP values were also observed in the left amygdala when these were computed based on fearful-neutral and neutral-neutral target-mask pairs. However, this site did not survive our stricter conjunction analysis. These findings suggest that amygdala responses may be predictive of behavioral choice when differences in arousal exist between the conditions but not when only valence differences exist (as in the context of happy faces). Moreover, a site in the inferior right amygdala predicted behavioral performance, suggesting that this region is important to the task, an interpretation consistent with a large body of data linking the amygdala to emotional perception (23).

CPs have been previously used in monkey neurophysiology to characterize the link between neuronal firing within one brain region and behavioral choice. In the original study by Britten *et al.* (8), a relatively modest but significant average CP of 0.55 was obtained for MT neurons tested in a direction discrimination task (8). Similar values have been obtained for MT neurons in depth perception (0.59) (3) and structure-from-motion tasks (0.57) (4). In one study (14), the mean CP value for MT neurons in a structure-from-motion task was

0.67. At the same time, single-cell unaveraged values in the 0.6–0.8 range have been reported (8). The present CP values were observed in a similar range (0.63–0.78). Our results thus demonstrate that fMRI signals can reliably predict behavioral choice in a rigorous and quantitative fashion at levels comparable with neuronal data.

In monkeys, activity that predicts behavioral choice or that appears to be linked to a perceptual decision has also been described in nonsensory regions, including the intraparietal area (24, 25), frontal eye fields (26), and medial (27) and dorsolateral (5, 28) prefrontal cortex. Although the involvement of some of these regions may have depended on the specific task used, it is noteworthy that we also observed choice-related activity in both medial and lateral prefrontal cortex (for some individuals, the latter site included responses in the inferior portions of the middle frontal gyrus).

Previous neuroimaging studies have also used signal detection theory to link behavior and fMRI signals. For example, Ress and Heeger (29) investigated a low-level visual detection task and showed that both hits and false alarms evoked greater responses than correct rejections and misses (see Fig. 3 as well as Fig. 6). Thus, in early visual areas V1, V2, and V3, activity corresponded to the subject's percept rather than the physically presented stimulus. A related approach was used by Grill-Spector *et al.* (30) to probe the neural correlates of detection and identification of faces in occipitotemporal cortex; their results revealed stronger responses in the fusiform gyrus when a face was detected vs. when it was missed.

In the present study, we showed that small fluctuations in fMRI responses were related to near-threshold fluctuations in the perceptual decision process. The present findings extend previous results from neurophysiology by showing that fMRI signals can quantitatively predict behavioral choice independent of stimulus properties and task accuracy. The present results show that certain perceptual decisions are not encoded by a single central brain region (see ref. 7). Instead, in the case of fear detection, decisions are represented over a network of nonsensory regions (see also ref. 5), which includes the PCC, MPFC, right IFG, and left insula. Our results do not imply, however, that the entire network is directly responsible for deciding that a fearful target is present. Choice-related signals in one region may reflect computations made at another site in the brain. Moreover, the complex decisions in our task are unlike simpler sensory decisions studied with monkeys. Thus, choice-related activations that predicted fear-present behavioral responses should be viewed to include a more general set of processes that are associated with fear detection. We suggest that the conscious decision that a fearful face is present is represented across a network of interconnected brain regions that prepare the organism to appropriately handle emotionally challenging stimuli and that regulate the associated emotional response. Our analysis demonstrates that we can quantify the link between the behavioral decision that fear is present and fMRI signals. Additional studies are needed to elucidate how and when these multiple regions are involved in perceptual decision making, and how their role might differ from other nonsensory regions (e.g., frontal eye fields, dorsolateral prefrontal cortex) shown to contribute to perceptual decisions (5, 7, 28).

We thank Mrim Boutla, Alexander Grunewald, and Shruti Japee, as well as the anonymous reviewers, for insightful discussions and feedback on the manuscript; Vishnu Murty for assistance in the analysis of the behavioral data; Alumi Ishai for making available a set of emotional faces; and Ziad Saad for the development of the Analyses of Functional Neuroimages (National Institute of Mental Health, National Institutes of Health)/MATLAB (Mathworks, Natick, MA) interface, which was instrumental in developing voxel-wise CP computations. This research was supported in part by Grant 1 R01 MH071589-01 from the National Institute of Mental Health, by the Brain Science Program (the Burroughs Wellcome Fund) of Brown University, and by the Ittleson Foundation.

1. Newsome, W. T., Britten, K. H. & Movshon, J. A. (1989) *Nature* **341**, 52–54.
2. Salzman, C. D. & Newsome, W. T. (1994) *Science* **264**, 231–237.
3. Uka, T. & DeAngelis, G. C. (2004) *Neuron* **42**, 297–310.
4. Grunewald, A., Bradley, D. C. & Andersen, R. A. (2002) *J. Neurosci.* **22**, 6195–6207.
5. Romo, R. & Salinas, E. (2003) *Nat. Rev. Neurosci.* **4**, 203–218.
6. Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A. & Ward, B. D. (2004) *Nat. Neurosci.* **7**, 295–301.
7. Heekeren, H. R., Marrett, S., Bandettini, P. A. & Ungerleider, L. G. (2004) *Nature* **431**, 859–862.
8. Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S. & Movshon, J. A. (1996) *Visual Neurosci.* **13**, 87–100.
9. Purushothaman, G. & Bradley, D. C. (2005) *Nat. Neurosci.* **8**, 99–106.
10. Adolphs, R. (2002) *Current Opin. Neurobiol.* **12**, 169–177.
11. Barlow, H. B., Levick, W. R. & Yoon, M. (1971) *Vision Res. Suppl.* **3**, 87–101.
12. Efron, B. & Tibshirani, R. J. (1993) *An Introduction to the Bootstrap* (Chapman & Hall, New York).
13. Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F. & Evans, A. C. (2002) *NeuroImage* **15**, 1–15.
14. Dodd, J. V., Krug, K., Cumming, B. G. & Parker, A. J. (2001) *J. Neurosci.* **21**, 4809–4821.
15. Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Shouten, J. L. & Pietrini, P. (2001) *Science* **293**, 2425–2430.
16. Bush, G., Luu, P. & Posner, M. I. (2000) *Trends Cognit. Sci.* **4**, 215–222.
17. Barbas, H. (2000) *Brain Res. Bull.* **52**, 319–330.
18. Maddock, R. J. (1999) *Trends Neurosci.* **22**, 310–316.
19. Small, D. M., Gitelman, D. R., Gregory, M. D., Nobre, A. C., Parrish, T. B. & Mesulam, M. M. (2003) *NeuroImage* **18**, 633–641.
20. Kobayashi, Y. & Amaral, D. G. (2003) *J. Comp. Neurol.* **466**, 48–79.
21. Yamasaki, H., LaBar, K. S. & McCarthy, G. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 215–222.
22. Mesulam, M.-M. (2000) in *Principles of Behavioral and Cognitive Neurology*, ed. Mesulam, M. (Oxford Univ. Press, New York), pp. 1–120.
23. Adolphs, R. (2002) *Curr. Opin. Neurobiol.* **12**, 169–177.
24. Shadlen, M. N. & Newsome, W. T. (2001) *J. Neurophysiol.* **86**, 1916–1936.
25. Platt, M. L. & Glimcher, P. W. (1999) *Nature* **400**, 233–238.
26. Gold, J. I. & Shadlen, M. N. (2000) *Nature* **404**, 390–394.
27. Hernandez, A., Zainos, A. & Romo, R. (2002) *Neuron* **33**, 959–972.
28. Kim, J. N. & Shadlen, M. N. (1999) *Nat. Neurosci.* **2**, 176–185.
29. Ress, D. & Heeger, D. J. (2003) *Nat. Neurosci.* **6**, 414–420.
30. Grill-Spector, K., Knouf, N. & Kanwisher, N. (2004) *Nat. Neurosci.* **7**, 555–562.

Supporting Information

Methods

Subjects

Nine volunteers (6 females) aged 23 (mean) \pm 8 (SD) years participated in the study, which was approved by the Institutional Review Board of both Brown University and Memorial Hospital of Rhode Island. All subjects were in good health with no past history of psychiatric or neurological disease and gave informed consent. Subjects had normal or corrected-to-normal vision.

Stimuli

Face stimuli were obtained from the Ekman set (1), a set recently developed by Ohman and colleagues (KDEF, Lundqvist, D., Flykt, A., and Ohman, A.; Karolinska Hospital, Stockholm, Sweden), as well as a set developed and validated by Alumit Ishai (2) at NIMH (Bethesda, USA). Forty instances of identity-matched fearful, happy, and neutral faces were employed.

fMRI data acquisition and analysis

fMRI data were collected using a Siemens 1.5 Tesla scanner. Each scanning session began with the acquisition of a high-resolution MPRAGE anatomical sequence (TR = 1900 ms, TE = 4.15 ms, TI = 1100 ms, 1-mm isotropic voxels, 256 mm field of view). Each subject performed 7-8 experimental runs, each lasting 4 min and 5 s. During each functional scan, 81 gradient echo echo-planar volumes were acquired with a TE of 38 ms and TR of 2980 ms. Each volume consisted of 37 axial slices with slice thickness of 3 mm and in-plane resolution of 3 mm x 3 mm.

fMRI data analysis: Standard maps

Two types of fMRI data analysis were performed: standard analyses using the general linear model (3) and choice probability maps. For the standard analyses, we employed AFNI tools (4) (<http://afni.nimh.nih.gov/afni>), unless indicated otherwise. Initially, both anatomical and functional data were normalized to the standard space defined by the Montreal Neurological Institute by using the BET and FLIRT tools from the FSL package (<http://www.fmrib.ox.ac.uk/fsl/>). For the functional data, the first 3 volumes of each run were discarded to allow for equilibration effects. The remaining volumes were then spatially registered to the volume acquired closest in time to the particular subject's high-resolution anatomy. Next, each volume was spatially smoothed with an 8-mm Gaussian filter (FWHM). Each subject's data were then analyzed with standard multiple regression methods (3). The linear models included constant and linear terms (for each run) that served as covariates of no interest (these terms controlled for drifts of MR signal). Regressors of interest were defined according to experimental conditions associated with the visual stimulus (e.g., fear containing trials and neutral trials) or behavioral responses (e.g., correct and incorrect trials). *F*-maps of the contrasts of interest were generated for each individual. Fixed-effects statistical group maps were then obtained by converting each individual's *F*-map into a *Z*-map and then combining these into a composite final group *Z*-map. We initially determined the voxels with a significant task-related response, namely, a supra-threshold response to any stimulus type. Because, at the group level, these responses were very robust, we employed a threshold of 10^{-6} for significance. These significant voxels were then used to create a mask that constrained the search space for specific contrasts. For these contrasts, we employed 0.001 as the threshold for significance.

Control Analyses

We now present several additional analyses that further investigate our fMRI data. In particular, control analyses are presented that address potential non-choice related origins of activations.

In neurophysiology studies, it is customary to determine choice probability values while holding the physical stimulus constant. For instance, choices involving depth judgments might be analyzed for zero-disparity stimuli only (5). In the present study, we adopted a different strategy. Because we were interested in investigating behavioral choice during a near-threshold fear-detection task, our analyses of choice probability involved both fear-containing and neutral stimuli. One potential concern with our results is that choice probability values thus confound behavioral choice with, for instance, stimulus-driven responses. For this reason, we included only voxels that exhibited significant choice probabilities in the *conjunction* of the choice maps obtained by separately considering neutral and happy face “catch” trials (see Methods). While this strategy is conservative, it largely eliminates the concern that choice probabilities reflected some low-level feature of the faces. In any case, here we provide additional analyses that further address the concern that other factors contributed to the response.

As described in the main text, we observed a network of regions that predicted behavioral choice. We quantified choice-related responses by computing the area under the fMRI ROC curves (choice probability values). Statistically significant choice probabilities demonstrate that it is possible for an ideal observer to determine trial type (reported target present vs. reported target absent) by inspecting a trial’s response amplitude. Thus, for voxels with significant choice probability values, the distributions of fMRI response magnitudes of reported target present and reported target absent exhibit a significant amount of *non* overlap. If fMRI signals in a brain region encode behavioral choice, then responses evoked during reported target present trials should be similar to one another (i.e., responses to hits and false alarms should be similar); likewise, responses evoked during reported target absent trials should be similar to one another (i.e., responses to correct rejects and misses should be similar). Deviations from this pattern of results indicate that the region more closely represents other properties. For example, large responses to hit and correct reject trials and weaker responses to misses and false alarms indicate that the region is sensitive to behavioral performance (correct vs. incorrect trials). At the same time, large responses to hits and misses and weak responses to correct rejects and false alarms show that the region is sensitive to the physical characteristics of the stimulus (fear-containing vs. neutral trials).

We performed several control analyses to rule out alternative explanations for significant choice probability values. We directly compared responses evoked by hits vs. false alarms as well as correct rejects vs. misses. To maximize statistical power, we included both high- and low-confidence trials, resulting in the following average number of trials: hits: 60; false alarms: 14; correct rejects: 61; misses: 14. fMRI responses to these trial types were obtained in the same manner as in the ROC maps (averages of time points 3, 6, and 9 s post stimulus). For each subject, we first computed the mean fMRI response on a voxel-wise manner. We then submitted the individual results to a second-level analysis that contrasted trial types via a paired *t* test. For the contrast of hits vs. false alarms, we only observed stronger activity for hits in the bilateral anterior fusiform gyrus (FG) and the inferior right amygdala. Similar differential activation was observed in the contrast of correct rejects vs. misses. These results indicate that fMRI signals in these regions were stronger during correct (hits and correct rejects) compared to incorrect (false alarms and misses) trials. Thus, responses in these regions were more predictive of performance

accuracy than of behavioral choice, a conclusion that is in direct agreement with the contrast of correct vs. incorrect trials (Fig. 6); note that the latter contrast was generated with standard multiple regression and convolved regressors. It should be pointed out that it is unlikely that significant differences were not detected in the above contrasts because of lack of statistical power as the choice probability maps demonstrate that significant differences between reported target present and reported target absent trials *can* be reliably detected.

Regions that evoked stronger responses to fear-containing stimuli compared to neutral stimuli were also *not* considered to represent behavioral choice. In such regions, responses to hits and misses (fear-containing trials) were larger than responses to correct rejects and false alarms (neutral trials). This was the case for all regions list in Table 2; see also regions marked with an “x” in Table 1.

Overall, four brain regions exhibited significant choice probability values that could not be explained by other responses properties: PCC, MPFC, right IFG (at both sites), and left insula. Thus, the responses observed in these regions most likely represent behavioral choice.

To reiterate, to rule out alternative interpretations of the choice-related signals, we determined voxels that reliably predicted behavioral choice but did *not* exhibit stimulus-driven or performance-related activation. Moreover, we also required that responses be similar during hits and false alarms (reported fear-present trials) and during correct rejects and misses (reported fear-absent trials), as would be expected if they reflected behavioral choice (Fig. 4 and 5).

Choice Responses for Constant Physical Stimuli

As stated above, in neurophysiology studies, it is customary to determine choice probability values while holding the physical stimulus constant. With our data, we followed a similar strategy by performing an additional control analysis in which we considered only fear-containing trials. For such trials, the data were sorted depending on behavioral choice (i.e., reported fear-present and reported fear-absent trials) and both individual and group choice probability maps were generated. Such maps (Fig. 7) exhibited a pattern of activation that overlapped with the maps shown in Fig. 1. Notably, we observed choice-related signals in the MPFC, PCC, anterior IFG, and left insula (at a slightly inferior slice than shown in Fig. 1 at $z = -3$; data not shown). These results strengthen the idea that the observed activations sites in Fig. 2 are due to behavioral choice and not to stimulus differences. However, in the present context, maps generated by holding the stimulus constant introduce a performance-related confound because this contrast compared correct (reported fear-present trials) vs. incorrect (reported fear-absent trials) trials.

Choice Responses Associated with Fear-Absent Trials

Behavioral choices in our task had two potential outcomes: reported fear present and reported fear absent. So far, we have concentrated solely on one of the outcomes (reported fear present) because the task was to indicate when fearful faces appeared, however briefly. However, we also investigated whether there were voxels in the brain for which the responses reliably predicted fear-absent responses. Such regions would exhibit choice probabilities significantly smaller than 0.5. One brain region that exhibited such behavior was the right posterior insula ($x = 40, y = -19, z = 18$). However, in this region, responses evoked during miss trials were greater than responses

evoked during correct reject trials (both trial types are linked to reported fear absent). Thus, these results show that the fMRI signals in this region are not specifically linked to behavioral choice. Two other brain regions are discussed below in the context of motor responses.

Choice and Motor Responses

In our task, subjects indicated whether a fearful face was present or not via a button press. Fear-present responses were linked to a right hand action and fear-absent responses were linked to a left hand action. Thus, behavioral choices were directly associated with motor responses. As expected, fear-present responses (associated with a right hand action) evoked strong signals in left (contralateral) central sulcus and right (ipsilateral) superior cerebellum. Moreover, significant choice probability values were observed in these regions. At the same time, fear-absent responses (associated with a left hand action) evoked strong responses in right central sulcus and left superior cerebellum; significant choice probabilities associated with fear-absent trials were also observed.

It is conceivable that some of the other significant choice probabilities that we observed could, in theory, reflect contributions of hand-specific motor responses or motor response preparation and not behavioral choice per se. We find this possibility unlikely for the following reasons. The key regions predictive of behavioral choice were the PCC, MPFC, right IFG, and left insula. The MPFC and the right IFG sites were quite anterior (Fig. 2) and located in regions not associated with motor response or preparation. The PCC and the insula are also not directly involved in these functions. A second key reason was that significant choice probabilities associated with deciding that fear was *absent* from the display were *not* observed aside from right motor cortex and left cerebellum (except for the right posterior insula; see above). Thus, we suggest that significant choice probability values that predicted reported target *present* trials were not due to hand-specific, motor-related processes.

To explicitly address the concern that choice-related activations were potentially due to motor-related processes, we studied an additional group of five participants. The task and all experimental details were exactly as before, with one exception. For this group, subjects were asked to indicate “fear” with a left-hand button press and to indicate “no fear” with a right-hand button press. The observed pattern of choice probabilities was very similar to the one obtained previously with the reverse hand-response mapping. Critically, the lateralized choice-related responses in the right IFG were lateralized in the same manner in the new group (Fig. 8).

Table 2. Contrast of fear-containing (fear-neutral pairs) and neutral (neutral-neutral pairs) stimuli

Fear > Neutral	MNI Coordinates			Z score
	X	Y	Z	
L Fusiform gyrus (pos)	-35	-64	-20	2.8
R Fusiform gyrus (pos)	18	-64	-15	3.3
L Fusiform gyrus (ant)	-48	-46	-22	3.6
R Fusiform gyrus (ant)	46	-49	-18	3.4
L Superior Temporal sulcus	-55	-55	9	3.4
R Superior Temporal sulcus	55	-52	24	3.2
Anterior Cingulate Cortex	4	30	42	4.1
R Middle Frontal gyrus	44	14	21	3.7
L Anterior Insula	-32	28	-3	4.0
R Anterior Insula	31	17	-3	3.3
R Ventral Striatum	12	-7	-11	3.4

L: left; R: right; pos: posterior; ant: anterior

Figure Captions

Figure 5. Percent signal change for three key brain regions exhibiting choice-related responses (see Table 1), as in Fig. 4. Graphs display the average signal change relative to correct reject trials (reported fear-absent during the presentation of neutral trials). For each subject and region, signals were obtained from the voxel with the peak choice probability value. Note that the responses evoked during reported fear-present trials (hits and false alarms) were similar and greater than responses evoked during reported fear-absent trials (misses and correct rejects); responses to misses and correct rejects were also similar to each other. Error bars represent the standard error of the mean. L: left; R: Right; CR: correct reject; FA: false alarm; PCC: posterior cingulate cortex.

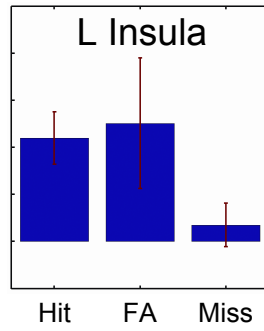
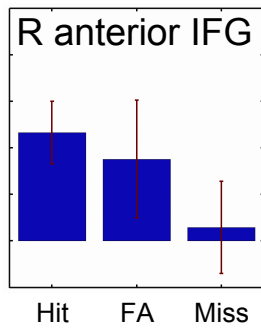
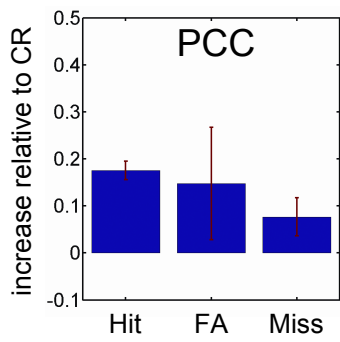
Figure 6. Performance-related activity as obtained by the contrast of correct and incorrect trials. *Top:* Bilateral anterior fusiform gyrus (FG). *Bottom:* Right amygdala (Amyg)/Entorhinal area (EA).

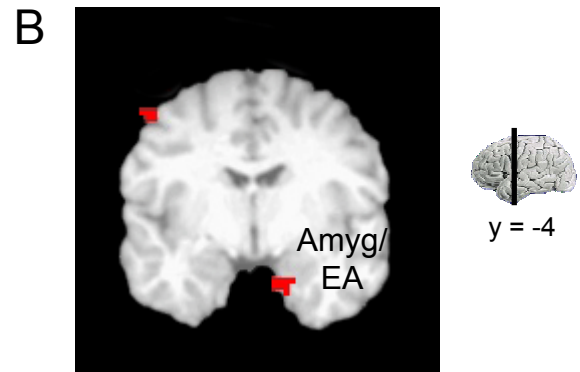
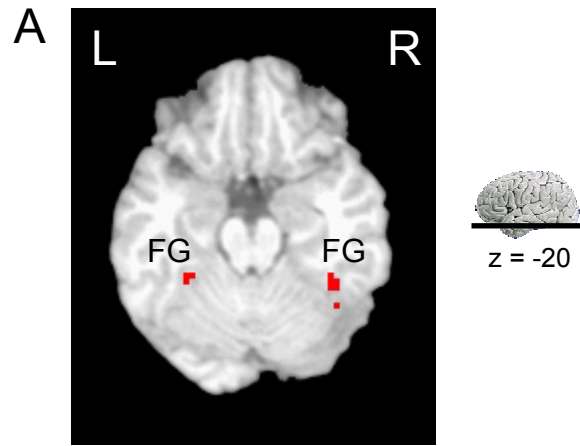
Figure 7. Choice responses for constant physical stimuli. Group choice probability maps generated for fear-containing stimuli only. The general pattern of activation is similar to the one shown in Fig. 2.

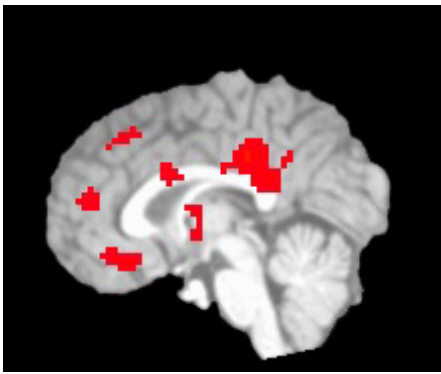
Figure 8. Group choice probability map with reversed hand-response mapping. An additional set of five participants was tested exactly as before aside from the reversed hand-response mapping. The pattern of choice-related activations was very similar to the one obtained before. In particular, right IFG voxels still reliably predicted behavioral choice when the response “fear” was indicated with the left hand. Note that choice-related signals were observed at sites slightly inferior to those observed before, although they were still lateralized to the right.

Supporting Material References

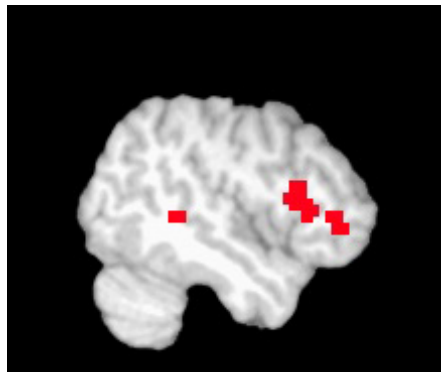
1. Ekman, P. & Friesen, W. V. (1976) *Pictures of facial affect* (Consulting Psychologists Press, Palo Alto, CA).
2. Ishai, A., Pessoa, L., Bickle, P. C. & Ungerleider, L. G. (2004) *Proc Natl Acad Sci U S A*.
3. Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Heather, J. D. & Frackowiak, R. S. (1995) *Human Brain Mapping* **3**, 165-189.
4. Cox, R. W. (1996) *Computers and Biomedical Research* **29**, 162-173.
5. Dodd, J. V., Krug, K., Cumming, B. G. & Parker, A. J. (2001) *Journal of Neuroscience* **21**, 4809-4821.







x = 0



x = 44

