# The Structure of Emotion

## Evidence From Neuroimaging Studies

**Lisa Feldman Barrett[1] and Tor D. Wager[2]**

*[1]Boston College and [2]Columbia University*

**ABSTRACT—*One common point of debate in the study of emotion is whether the basic, irreducible elements of emotional life are discrete emotion categories, such as* anger, fear, sadness, *and so on, or dimensions such as* approach *and* avoidance. *Resolving this debate will identify the basic building blocks of emotional life that are the most appropriate targets of scientific inquiry. In this paper, we briefly review meta-analytic work on the neuroimaging of emotion and examine its potential for identifying "natural kinds" of emotion in the brain. We outline criteria for identifying such natural kinds, summarize the evidence to date on category and dimensional approaches, and suggest ways in which neuroimaging studies could more directly address fundamental questions about the nature of emotion.***

**KEYWORDS—*emotion; affect; amygdala; neuroimaging***

"In nature's infinite book of secrecy / A little I can read."

—Shakespeare, *Antony & Cleopatra*

What are the basic building blocks of emotional life that a science of emotion should focus on? This question is almost as old as psychology itself, and it remains unanswered. The "basic emotion" approach argues that certain categories of emotion, described by such English words as *anger*, *sadness*, *fear*, *happiness*, and *disgust*, are biologically basic—inherited, reflex-like modules that cause a distinct and recognizable behavioral and physiological pattern (e.g., Ekman, 1972; Panksepp, 1998). The "dimensional" approach argues that anger, sadness, fear, and so on are categories that characterize more highly elaborated responses constructed from more fundamental, biological properties such as *valence* (pleasure/displeasure) and *arousal* (high activation/low activation; Russell & Barrett, 1999), *positive* and *negative activation* (e.g., Watson & Tellegen, 1985), or *approach* and *withdrawal* (e.g., Lang, Bradley, & Cuthbert, 1990). The pressing question is which typology is given by nature and consists of "natural kinds," such that it is possible to make inductive discoveries about them?

Natural kinds give psychobiological evidence of their existence. Neuroimaging techniques (functional magnetic resonance imaging, or fMRI, and positron emission tomography, or PET) have recently opened the door to searching directly for the circuitry that supports emotional processing in humans. To indicate a natural kind, patterns of neural activation must be consistent (i.e., show increased activation regardless of the induction method used) and specific (e.g., a fear circuit should be architecturally separable from an anger circuit even though the two may share some brain areas in common). To the degree that consistency and specificity criteria are satisfied, an emotion category or affect dimension can be said to have a "brain marker." In principle, it should be possible to map patterns of activity within a connected set of brain areas, but in practice, most of the imaging research to date has searched only for the most salient or distinctive feature (e.g., brain area) in the circuitry for a given emotion construct.

Meta-analytic summaries (statistical summaries of empirical findings across many studies) of the first 10 years or so of research are now available (Murphy, Nimmo-Smith, & Lawrence, 2003; Phan, Wager, Taylor, & Liberzon, 2002; Wager, Phan, Liberzon, & Taylor, 2003) and allow us to begin to search for evidence of natural kinds. The meta-analyses of Murphy et al. and Phan et al. focused on the neural activations for the five emotion categories—happiness, sadness, anger, fear, and disgust. Wager et al., along with Murphy et al., focused on the neural activations for two affective dimensional models (positive/negative affect and approach/withdrawal motivation). None of the meta-analyses assessed the valence/arousal affective model, in part because many neuroimaging studies fail to measure arousal separately. There were many potential methodological issues in conducting these meta-analyses (e.g., the sole reliance on reporting the peak activation within a broader area of activation, the inherent limitations in the signal-source resolution and spatiotemporal resolution of

**Address correspondence to Lisa Feldman Barrett, Department of Psychology, Boston College, Chestnut Hill, MA 02167; e-mail: barretli@bc.edu.**

**TABLE 1**

*Summary of Emotion Category–Brain-Area Activation Correspondences Found in Two Studies (Phan, Wager, Taylor, & Liberzon, 2002; Murphy, Nimmo-Smith, & Lawrence, 2003)*

| | Phan et al. (2002) | | Murphy et al. (2003) | |
|---|---|---|---|---|
| | *N* | Brain activations | *N* | Brain activations |
| Anger | 5 | None | 8 | Lateral orbitofrontal cortex |
| Sadness | 14 | Subcallosal anterior cingulate cortex (ACC) | 14 | Rostral supracallosal anterior cingulate and dorsomedial prefrontal cortex |
| Disgust | 5 | Basal ganglia | 7 | Insula/operculum and globus pallidus |
| Fear | 13 | Amygdala | 26 | Amygdala |
| Happiness | 11 | Basal ganglia | 11 | Rostral supracallosal anterior cingulate/ dorsomedial prefrontal cortex |

**Note.** Sample sizes from Murphy et al. (2003) were taken from Fig. 3, which reports the number of studies included in follow-up chi-square analyses (Murphy, personal communication, 2004). The subcallosal ACC is considered the ''visceral'' part of the ACC; it is connected to the medial orbital frontal cortex and is associated with autonomic control. The supracallosal ACC is considered to the ''cognitive'' aspect of the ACC; it is connected to the dorsomedial prefrontal cortex and dorsolateral prefrontal cortex and is associated with attention and working-memory functions. The globus pallidus is part of the basal ganglia (see Fig. 1).

current neuroimaging techniques, and the heterogeneity of the studies included). Nonetheless, these meta-analyses provide a starting point for evaluating whether categories or dimensions capture natural kinds of emotional phenomena in the brain. They also highlight issues that will help make the results of future studies more cohesive and interpretable.

## EVIDENCE FOR BASIC EMOTIONS IN THE BRAIN

The main findings of the emotion-category–brain-location analyses from Murphy et al. (2003) and Phan et al. (2002) are summarized in Table 1, and the locations of some of the key brain regions are shown in Figure 1. Both meta-analyses agreed that right and left amygdalae were preferentially activated with fear, and that rostral (or forward) portions of the anterior cingulate cortex (ACC) were preferentially activated by sadness. These findings are in agreement with lesion and animal studies that have linked fear and depression to the amygdala and to the subcallosal portion of the ACC (i.e., the rostral portion of the ACC that is below the corpus callosum), respectively (e.g., LeDoux, 2000). Both analyses suggested that disgust produced frequent activation in the basal ganglia (particularly the globus

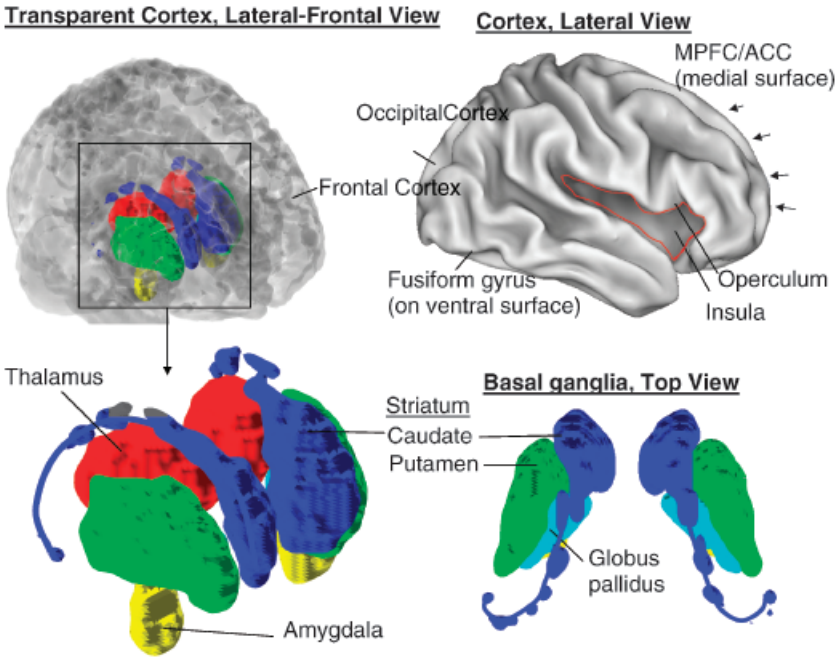


**Fig. 1.** Key brain regions implicated in emotion-related processing. Shown are the insular cortex (red outline); thalamus (red); amygdala (yellow); basal ganglia, composed of the caudate (blue), putamen (green), and globus pallidus (cyan); medial surfaces of the medial prefrontal cortex (MPFC) and anterior cingulate cortex (ACC) are not shown.

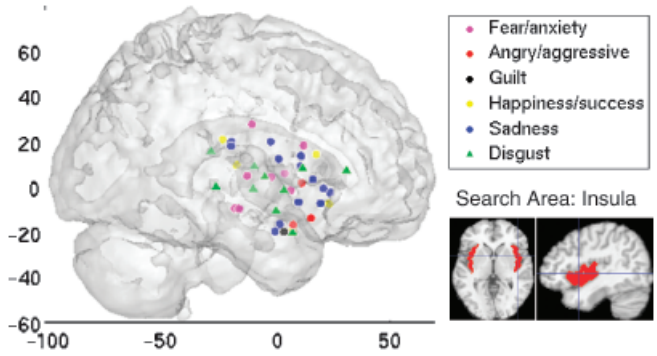


**Fig. 2.** Reported peaks in the insula (from Wager et al., 2003). The insular region of interest is shown in red in the inset panels. Peaks in the right insula are shown on the transparent brain (left panel), but the pattern of results is similar for the left insula. These findings suggest that disgust-related stimuli do not uniquely and consistently activate the insula. Rather, insular activations were significantly more frequent with negative or withdrawal-related emotions across specific emotion categories and when emotion tasks required concurrent cognitive judgments (Phan et al., 2002).

pallidus). Murphy et al. also reported disgust-specific activations in the insula, a large cortical region in the frontal lobes adjacent to orbitofrontal cortex, that contains processing regions for taste, smell, and somatic as well as visceral activity, whereas Phan et al. found that insula activity was associated with negative emotions generally (Fig. 2). At first glance, then, these findings seem to suggest that certain basic emotion categories may be natural kinds in the brain.

### Consistency

The fear–amygdala correspondence was the most consistent finding across both studies, yet Phan et al. (2002) reported that only 60% of studies involving fear showed increased activation in the amygdala, and Murphy et al. (2003) reported just under 40%. Small sample sizes and stringent statistical thresholds may account for these modest percentages, but the consistency of these emotion–brain correspondences is called into question by the additional observation that increases in amygdala activation were reliably related to the method by which emotions were induced. For example, in humans, the amygdala is particularly responsive to faces and other visual stimuli (Phan et al., 2002). In fact, if one considers only studies in which participants viewed facial caricatures of fear, the fear–amygdala correspondence increases by about 20% in each meta-analysis. These findings call into question the conclusion that activation of the amygdala in humans reflects engagement of a core "fear system" in the brain.

A similar case can be made for sadness. The percentage of studies that showed a sadness–anterior cingulate correspondence (60% in Phan et al., 2002, and 50% in Murphy et al., 2003) was modest, and the sadness–ACC correspondence may be accounted for by induction method. Many of the studies involving sadness stimuli (e.g., at least 10 of the 14 studies summarized in Murphy et al.) involved cognitive demand. Phan et al. reported that cognitively demanding emotional tasks (such as being asked to remember a sad event to induce a feeling of sadness, or rating emotional stimuli) specifically engaged rostral portions of the ACC, as compared to passive emotional tasks (where stimuli are merely viewed and experienced). Thus, these activations may not reflect sadness per se but something more complex about the cognitive and motivational processes involved in interacting with emotionally valenced stimuli.

### Specificity

The emotion-category–brain-localization correspondences were not only inconsistent; there is mounting evidence against their specificity. Given space constraints, we turn to the fear–amygdala correspondence for illustrative purposes. A number of studies are now available to show that the amygdala is engaged by positive objects and rewards or novelty (for a discussion, see Barrett, 2006a). Simple perceptual cues (e.g., eye gaze; Adams, Gordon, Baird, Ambady, & Kleck, 2003) modulate whether or not viewing facial caricatures of fear elicits amygdala activation, and even those with amygdala damage can correctly identify those caricatures when attention is directed toward the eyes of the stimulus caricature (Adolphs et al., 2005). Taken together, the findings are more consistent with the view that the amygdala is involved with computing the affective significance of a stimulus (i.e., the extent to which the stimulus predicts an impending threat or reward; for a discussion, see Barrett, 2006a).

## EVIDENCE FOR BROAD AFFECTIVE DIMENSIONS IN THE BRAIN

The main findings from affect-location analyses from both meta-analyses are presented in Table 2.

### Consistency

Differences in analytic strategy make a detailed comparison of the Wager et al. (2003) and Murphy et al. (2003) meta-analyses impossible, but several broader observations are possible. Both analyses observed greater left-sided activations (left-lateralization) for approach- (vs. withdrawal-) related affect (localized to the frontal cortex in Wager et al.). Wager et al. reported a number of activation foci for withdrawal-related affect that were not observed by Murphy et al., including regions in the amygdala, the left medial prefrontal cortex and anterior cingulate, the basal ganglia, and the insula. Both PET and fMRI imaging methods produced identical results. Moreover, there was a dissociation between approach-active and withdrawal-active regions in the medial prefrontal cortex.

The two meta-analyses showed less consistency in their assessment of the positive versus negative activation model. Murphy et al. (2003) reported null effects. Wager et al. (2003)

**TABLE 2**

*Summary of Affect Dimension–Location Correspondences in Two Studies (Murphy, Nimmo-Smith, & Lawrence, 2003; Wager, Phan, Liberzon, & Taylor, 2003)*

| | Murphy al. (2003) | Wager et al. (2003) |
|---|---|---|
| Positive affect | No difference in activation patterns | Left lateral frontal cortex, basal ganglia |
| Negative affect | No difference in activation patterns | Insula |
| Approach | Greater left- than right-sided activations across anterior and posterior regions | Left lateral frontal cortex, anterior medial prefrontal cortex |
| Withdrawal | No lateralization differences | Amygdala, left medial prefrontal cortex and rostral anterior cingulate, right striatum (basal ganglia), left insula, left fusiform and superior occipital cortices |

reported findings similar to (but considerably weaker than) those observed for the approach/avoidance model.

### Specificity

The most striking observation from Wager et al. (2003) is that many of the same regions showing emotion-category effects also showed specialization for the broader category of withdrawal-related affects. For example, fear-related stimuli may activate the dorsal (or top) portion of the amygdala because they are part of a broader class of aversive stimuli that engage this region. The obvious next questions are whether affect dimensions or emotion categories better predict the observed patterns of activation, and whether another potential function, perhaps based on stimulus salience, fits the pattern of data better. This would be an important direction for future research.

## TOWARD AN UNDERSTANDING OF EMOTION IN THE BRAIN

Every emotional event has neural correlates, and discovery of these correlates is a major task of science. An important step is to determine whether discrete emotion categories, such as fear, and/or broader affective categories, such as approach and avoidance, have consistent and specific neural correlates—brain markers—that justify thinking of them as natural kinds. A conservative approach to the existing evidence precludes drawing firm conclusions about natural emotion kinds in the human brain, but it is possible to offer several tentative observations.

Thus far, emotion category–location correspondences are neither consistent nor specific. This observation is largely consistent with evidence from other instrument-based measures of emotion in humans indicating that it is currently not possible to characterize each emotion category by a biological signature (Barrett, 2006a). Although there is good evidence from the animal literature that specific behaviors (e.g., freezing) may depend on specific nuclei (groups of neurons) in the amygdala and brainstem (e.g., LeDoux, 2000; Panksepp, 1998), there currently is insufficient evidence that these form the basis for basic emotion circuits (because each behavior is not necessarily associated with any single emotion category). Of course, it is possible to find caveats to explain why evidence for basic emotion kinds has not materialized (e.g., human research may fail to elicit strong and differentiated emotional responses in the lab), and distinct, natural kinds of emotions might reveal themselves if only scientists could find the right eliciting stimuli, employ better measurement tools, or use more sophisticated and precise research designs. Nonetheless, the available imaging research highlights an important observation: While one may not want reject the idea of emotions as natural kinds defined by neural circuits in the brain, it is not prudent to accept that idea too quickly either.

The limitations of emotion research notwithstanding, neuroimaging studies do show some consistent effects across studies. Of the three existing dimension models, there is emerging evidence from the Murphy et al. (2003) and Wager et al. (2003) analyses for the reliability of the approach/withdrawal model, but aspects of its consistency (e.g., variations across induction method) and specificity have not been fully tested. Furthermore, the valence/arousal model remains to be meta-analytically evaluated, and dimension-specific activation patterns need to be tested for specificity against nonemotional categories such as salience and self-relevance.

Future research will undoubtedly clarify, and perhaps even reshape, the empirical landscape that we have mapped. First, the human brain is being visualized with ever-increasing precision, allowing the study of functional regions within the brain that are much smaller than the very broad regions that were the focus of Phan et al.'s (2002) and Murphy et al.'s (2003) emotion-category analyses. For example, the higher-resolution analysis of Wager et al. (2003) provided the observation that separate, nearby areas in the rostral medial prefrontal cortex were frequently activated by approach- and withdrawal-related affect (a finding replicated in a meta-analysis of the orbitofrontal cortex by Kringelbach & Rolls, 2004), and may explain why Murphy et al. found medial prefrontal cortex correspondences for both happiness and sadness. Furthermore, finer-grained analysis of the cortex (e.g., single-neuron activations, neurochemical analyses, more fine-grained temporal analyses) as well as PET imaging of the brainstem and other subcortical areas may help to resolve the question of which emotion structure reflects what actually occurs in the brain.

Second, researchers need to move from studying singular brain areas to identifying circuits, because a given brain area may be involved in more than one functional circuit. Studies that examine the functional connectivity between brain areas, supplemented by neuroanatomical studies of neuronal circuitry in humans and primates, allow the possibility of investigating whether the pattern of information transfer, rather than architectural distinctiveness, characterizes distinct emotion circuits. In this respect, neuroimaging plays a unique and complementary role to lesion studies in animals, because neuroimaging alone allows the simultaneous measurement of the entire brain and reveals dynamic patterns of functional connectivity across diverse systems.

Third, researchers are beginning to take more care in measuring or controlling for variables (e.g., the autonomic arousal, motor-response tendencies, and cognitive demand associated with emotional stimuli in a given processing instant) that may confound, or obscure, the evidence on emotion structure. Without such controls, activations may result more from the idiosyncratic demands of the tasks than from the emotional or affective state being induced.

Fourth, researchers must work to formalize the inferential process before neuroimaging data can be used to clarify key debates in the psychological literature. The "brain mapping" approach is typically designed to test the probability of activation in a brain area given a certain emotion or affect (e.g., what parts of the brain activate with fear?), but it is also necessary to test the probability that an emotion (or affect) is present given activation in a given brain area (e.g., given amygdala activation, is there fear?). Claims about the psychological meaning of activation are made in virtually every neuroimaging study, and these claims are often made based on general ideas about what brain areas "do," leading to intuitive, rather than empirical, tests of specificity.

Finally, progress in understanding the structure of emotion not only requires conducting better studies with better research tools; it may also require learning to ask different sorts of questions about emotion. While it is important to ask questions like "how many emotions are there?" and "what is the brain marker for fear," it may also be important to consider the possibility that emotion words, such as *fear*, *anger*, *sadness*, *disgust*, and *happiness*, do not refer to specific mechanisms in the mind or the brain (Barrett, 2006b). If emotions are psychological events like memories, then they are best thought of as products of distinct but interacting psychological processes with accompanying neural systems, and scientists might begin to design experiments to systematically map how instances of emotion are synthesized from component psychological processes that we know to be implemented in the human brain. Doing so may provide us with a better translation of the pages of "nature's infinite book" that are the workings of the brain, and allow us to answer the age-old question of what emotions really are.

**Recommended Reading**

Murphy, F.C., Nimmo-Smith, I., & Lawrence, A.D. (2003). (See References)

Phan, K.L., Wager, T.D., Taylor, S.F., & Liberzon, I. (2002). (See References)

Wager, T.D., Phan, K.L., Liberzon, I., & Taylor, S.F. (2003). (See References)

***Acknowledgments***—We thank K. Luan Phan and Fionnuala Murphy for providing us with additional information about their meta-analysis. We also thank Michael Ross, Kevin Ochsner, James Russell, David DeSteno, and Christopher Wright for their helpful comments on earlier drafts of this article. This article was supported by NIMH Grant K02 MH001981 to Lisa Feldman Barrett.

## REFERENCES

Adams, R.B., Gordon, H.L., Baird, A.A., Ambady, N., & Kleck, R.E. (2003). Effects of gaze on Amygdala sensitivity to anger and fear faces. *Science*, *300*, 1536–1537.

Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P., & Damasio, A.R. (2005, January). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68–72.

Barrett, L.F. (2006a). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*, 28–58.

Barrett, L.F. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*, 20–46.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation* (pp. 207–283). Lincoln: University of Nebraska Press.

Kringelbach, M.L., & Rolls, E.T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, *72*, 341–372.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, *97*, 377–398.

LeDoux, J.E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.

Murphy, F.C., Nimmo-Smith, I., & Lawrence, A.D. (2003). Functional neuroanatomy of emotion: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 207–233.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.

Phan, K.L., Wager, T.D., Taylor, S.F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, *16*, 331–348.

Russell, J.A., & Barrett, L.F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*, 805–819.

Wager, T.D., Phan, K.L., Liberzon, I., & Taylor, S.F. (2003). Valence, gender, and lateralization of functional brain anatomy in emotion: A meta-analysis of findings from neuroimaging. *Neuroimage*, *19*, 513–531.

Watson, D., & Tellegen, A. (1985). Towards a consensual structure of mood. *Psychological Bulletin*, *98*, 219–235.