

# Mapping discrete and dimensional emotions onto the brain: controversies and consensus

Stephan Hamann

Department of Psychology, Emory University, Atlanta, GA 30322, USA

**A longstanding controversy in the field of emotion research has concerned whether emotions are better conceptualized in terms of discrete categories, such as fear and anger, or underlying dimensions, such as arousal and valence. In the domain of neuroimaging studies of emotion, the debate has centered on whether neuroimaging findings support characteristic and discriminable neural signatures for basic emotions or whether they favor competing dimensional and psychological construction accounts. This review highlights recent neuroimaging findings in this controversy, assesses what they have contributed to this debate, and offers some preliminary conclusions. Namely, although neuroimaging studies have identified consistent neural correlates associated with basic emotions and other emotion models, they have ruled out simple one-to-one mappings between emotions and brain regions, pointing to the need for more complex, network-based representations of emotion.**

## Controversies over emotion theories

The nature of emotion has been debated for more than twenty-five centuries. A major debate has focused on the nature of the basic units of emotion and whether these units are essentially dimensional or discrete [1]. Discrete emotion views (see [Glossary](#)) propose that there exists a limited number of distinct emotion types, each with specific characteristic properties, as opposed to a continuum of emotional states [2]. One highly influential type of discrete emotion theory, basic emotion theory [3–8], proposes a limited set of basic emotions (for example, happiness, sadness, anger, fear, disgust, and surprise) that are universal, biologically inherited, and have unique physiological and neural profiles that distinguish them from one another. By contrast, the other major theoretical position, represented by dimensional theories of emotion, conceptualizes emotions as arising from combinations of more fundamental dimensions, such as emotional arousal (strength or intensity of emotion) and emotional valence (degree of pleasantness or unpleasantness), in combination with cognitive processes, such as appraisal and attribution of meaning [9–14]. For example, dimensional views propose that a basic emotion such as fear emerges from a combination of negative valence, high arousal, and other

attributes that are not specific to the category of fear *per se* [15–21].

Despite the large empirical literature that has attempted to determine which of these theories better

## Glossary

**Emotion:** although there is no consensus definition of the term ‘emotion’, it is frequently defined in terms of a temporary change in affect or feeling state, elicited by an affectively salient situation, that involves coordinated, multiple systems, including physiology, brain activity, behavior, and (in humans) conscious experience. These changes typically facilitate adaptive behavioral responses, such as approach or avoidance. A key related distinction is between emotion recognition (perceiving an emotion in another individual) and emotional experience (one’s own emotions).

**Valence:** refers to the degree of pleasantness/positivity vs unpleasantness/negativity associated with an emotion.

**Arousal:** in the context of emotional arousal, refers to the strength of experienced emotion, ranging from calm to excited.

**Psychological construction theories of emotion:** emphasize that types of emotion emerge from a construction process, in which basic psychological operations, such as perception, attention, and memory, combine to create emotional meaning, influenced by social and linguistic factors.

**Dimensional theories of emotion:** propose that emotional states can be accurately represented by a small number of underlying affective dimensions, most commonly two (arousal and valence).

**Discrete theories of emotion:** propose that there exists a small number of separate emotions, characterized by coordinated response patterns in physiology, brain, and facial expression. Basic emotions are a subset of discrete emotions proposed to be the most elemental and adaptive, culturally universal, and to have an inherited, biological basis in the brain. Six basic emotions have been most frequently suggested: happiness, sadness, anger, disgust, fear, and surprise.

**Activation likelihood estimation (ALE):** a type of coordinate-based meta-analysis used to summarize the location of consistent findings across multiple neuroimaging studies. Findings from individual studies, represented as sets of three-dimensional coordinates of maximal brain activation, are modeled as probability distributions. The degree of spatial overlap across multiple studies is then estimated, producing a map that indicates the degree of convergence of results across studies.

**Multivariate pattern analysis (MVPA):** a neuroimaging analysis method that uses powerful pattern-classification algorithms to decode information concerning cognitive and affective representations from patterns of activity distributed across multiple brain loci. This contrasts with standard analysis methods that focus on establishing relationships between individual brain loci and cognitive variables. MVPA can accurately decode the information currently represented in an individual’s brain, such as information about a viewed object or information being retrieved from long-term memory.

**Conceptual act model:** a particular psychological constructionist model of emotion that proposes that emotions arise from the combination of ‘core affect’ (mental representation of bodily changes, associated with arousal and valence) with a categorization process that determines the emotional meaning of core affect, incorporating past experience and the current situation. Categorization creates a meaningful interpretation of core affect, which can be experienced as a discrete emotion, such as fear, an affective feeling of high arousal, or even as a non-affective sensation, such as stomach upset. Although the conceptual act model has a dimensional basis, categorization can produce a wide variety of emotion states, including discrete emotions, albeit via a completely different process than in basic emotion models.

Corresponding author: Hamann, S. ([shamann@emory.edu](mailto:shamann@emory.edu)).

explains emotional phenomena, there remains a lack of consensus. Basic emotion theory proposes that basic emotions should be reflected in consistent and discriminable patterns of psychophysiological responses and brain activity [8,22–25]. Accordingly, several studies have sought consistent and specific psychophysiological correlates for basic emotions [23,26,27]. Intuitively, it might seem that differences between basic emotions, such as fear and sadness, should be relatively easy to detect, because of the apparent ease with which we discern these emotions in ourselves and others. Contrary to this intuition, reviews of early studies concluded that psychophysiological studies failed to find consistent and specific correlates for basic emotions [28,29], in contrast with the relatively well-established psychophysiological correlates of emotional arousal (e.g., skin conductance response) and valence (e.g., facial muscle activity) dimensions [14,30,31]. Although recent psychophysiological studies using multivariate approaches [23,26] have been much more successful in differentiating emotions, this lack of early success motivated a search for alternative approaches.

Most recently, this debate has moved to the domain of neuroimaging. Because emotions are ultimately mediated by brain mechanisms, it should be possible to map basic emotions onto their underlying brain mechanisms [32–34]. Neuroimaging can assess activity and functional connectivity across the entire brain as different emotional states unfold across time and is potentially better able to identify biological correlates for basic emotions. The mapping of emotions to their brain representations is part of a larger goal of mapping all mental processes onto their corresponding brain mechanisms [35–39].

This review highlights recent developments and trends in this continuing debate over emotion theories and how they can be mapped onto patterns of brain activity. The first part of this review discusses key relevant findings from neuroimaging studies and meta-analyses regarding the brain representation of emotions and evaluates the degree to which these findings support different emotion theories. In the second part, broader implications are discussed, including the important role of functional networks and multivariate analyses, the complementary role of other approaches, such as neuropsychological lesion and nonhuman animal studies, and finally, outstanding issues and directions for future research. As will be discussed, neuroimaging studies have identified consistent neural correlates associated with basic emotions and other emotion models. Rather than being specialized for one type of emotion, however, individual brain regions often contribute to multiple emotions, ruling out simple one-to-one mappings between emotions and brain regions and pointing to the need for more complex, network-based representations of emotion.

### Neuroimaging meta-analyses

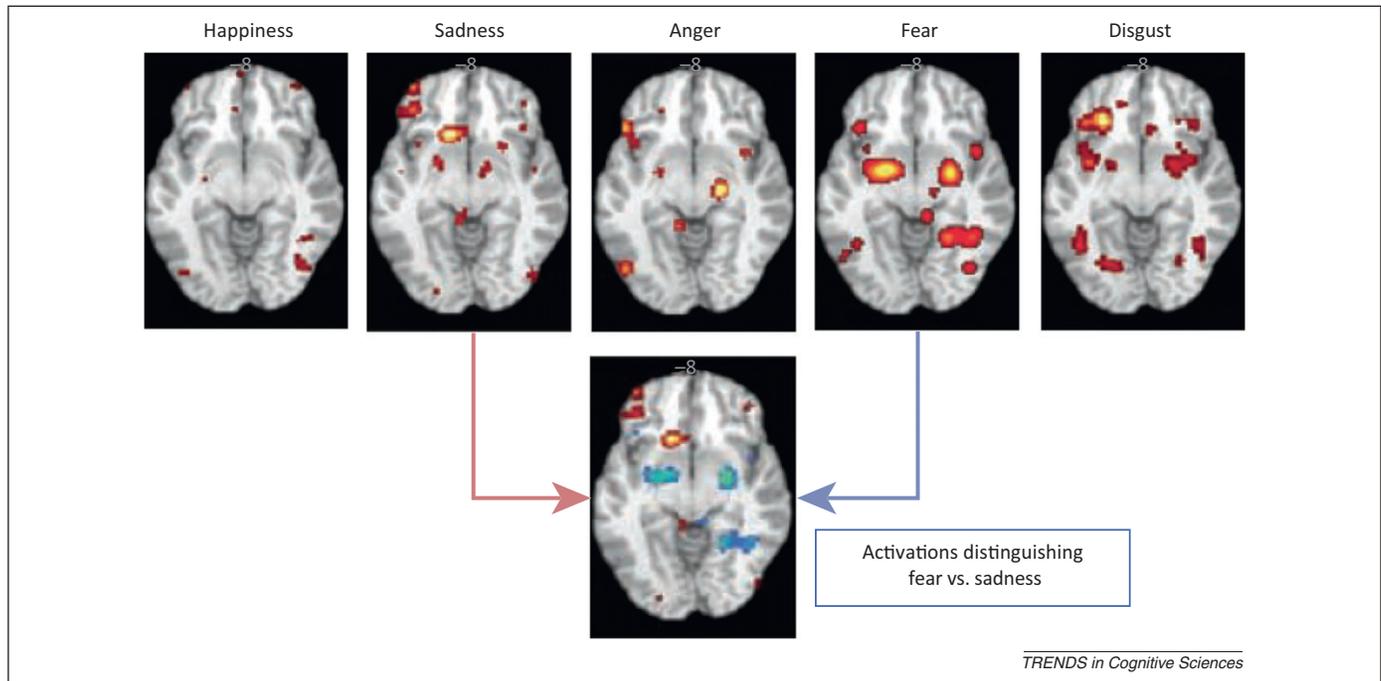
Ideally, in order to identify consistent and discriminable activations associated with each emotion category, multiple emotions should be elicited and contrasted within a single neuroimaging study [40,41]. However, very few neuroimaging studies have examined multiple emotions in the same experiment in a manner that allows testing

directly the predictions of basic emotion theories [42,43], leading to a search for other methods to test these predictions. Meta-analyses that assess function-location correspondences can help overcome this limitation in the neuroimaging literature, because they allow activation patterns associated with individual emotions to be compared across different studies [37,44]. Neuroimaging meta-analyses can be used to identify neural patterns associated with emotion states that are consistent and discriminable and, because they pool effects across multiple studies, meta-analyses can substantially reduce problems frequently associated with neuroimaging studies, including low experimental power and small sample sizes [37].

Current neuroimaging meta-analysis methods, such as the activation likelihood estimation (ALE) method, quantitatively assess relationships between functions such as emotion and regional brain activations [37,45–47]. Such meta-analyses are referred to as function-location meta-analyses, because the primary interest is in the spatial distribution and clustering of functional activation across brain loci. The ALE method collects neuroimaging studies that have targeted specific functions and extracts the three-dimensional brain activation coordinates from reported statistical contrasts (e.g., sadness vs neutral emotion), converting them into representative statistical brain maps [37]. Spatial overlap (consistency) of activations within emotion conditions and significant differences between emotion (discriminability) can then be assessed using permutation methods.

These meta-analysis methods were used to evaluate the predictions of basic emotion theories in a recent study by Vytal and Hamann [43]. This study combined a large number of neuroimaging studies of basic emotions to determine the extent to which basic emotion categories are associated with consistent and discriminable brain activation patterns, as predicted by basic emotion theories. To assess consistency, brain activation loci were identified, the activity of which was most consistently and strongly associated with the five basic emotions happiness, sadness, anger, fear, and disgust (few neuroimaging studies of surprise exist, precluding analysis of this basic emotion). Activation maps for each pair of emotions were contrasted statistically to test the prediction that each basic emotion category is discriminable from other emotions.

The results generally supported the predictions of basic emotion views. Each basic emotion was characterized by consistent neural correlates, and each emotion could be discriminated from each of the other emotions in pairwise contrasts. Figure 1 illustrates representative consistency maps and a map of regions differentiating two emotions (only one axial level is shown; for full results, see [43]). Importantly, the structure-function correspondences identified in the meta-analysis overlapped substantially with structure-function correspondences identified using other approaches, such as neuropsychological lesion studies and nonhuman animal models, providing converging evidence for basic emotion accounts. For example, among the most prominent findings were that fear was consistently associated with activation of the amygdala; disgust with activation of the insula, ventral prefrontal cortex, and amygdala; sadness with activation of medial prefrontal cortex; anger



**Figure 1.** Brain maps of consistent and discriminable activations for basic emotions. The upper row shows brain regions (regions in color) where activations across multiple neuroimaging studies spatially converged in an activation likelihood analysis (see text for description), indicating regions of consistent activation for each emotion relative to a non-emotional baseline condition (adapted from [43]). Activation clusters are shown superimposed on a single representative axial slice from a standard structural MRI image (at  $z = -8$  in a standard anatomical space [100]; see [43] for full results). The lower row illustrates a selected result from the discriminability analysis, which statistically contrasted the maps for each basic emotion in pairwise analyses. Blue-scale clusters indicate regions where activation was significantly more likely for fear than sadness (e.g., the cluster in the bilateral amygdala); red-scale clusters show the opposite contrast. All pairwise contrasts between basic emotions yielded regions of differential activation likelihood, indicating discriminable activations for each emotion.

with activation of orbitofrontal cortex; and happiness with activation of rostral anterior cingulate cortex.

These findings are generally consistent with those of two earlier, smaller meta-analyses that used different meta-analysis methods [48,49]. Both studies found limited evidence for consistent associations between brain regions and basic emotions, and one study was also able to discriminate between several basic emotions, although its analysis method was not able to localize these differences to specific brain regions [49].

#### Dimensional and psychological construction theories

A recent study by Lindquist and colleagues [18], the most extensive meta-analysis to date, reexamined the neuroimaging evidence for basic emotions and arrived at different conclusions. Their meta-analysis combined three different meta-analysis methods to assess consistency and discriminability: an analysis similar to ALE, a frequency analysis, and logistic regressions that determined which emotion categories predicted increased regional activity. In line with previous meta-analyses [43,48,49], evidence for consistent regional brain activations corresponding with each basic emotion category was found, even though the degree of consistency varied across basic emotions. However, no evidence was found to support the prediction that basic emotions have discriminable neural correlates. Instead, every region that was activated for a given basic emotion was also activated for at least one other basic emotion.

As an alternative to the basic emotions view, the authors detailed a conceptual act model [17,19,20], a particular psychological construction view, in which emotions emerge from combinations of more basic psychological and

neural components (Box 1). The meta-analytical findings were interpreted as being more consistent with the conceptual act model and with the interactions of six functional groups identified in a previous meta-analytic study [50], mediating processes such as core affect and conceptualization (Figure 2). Unlike other meta-analyses that focus on mappings between individual, predefined emotions to brain activation patterns, this earlier study used a novel, entirely data-driven approach that avoided the use of emotion labels, such as fear and sadness. By examining patterns of co-activation across a large number of neuroimaging studies investigating emotion, they identified six functionally distributed groups that they proposed to act as functional building blocks for the generation of emotional states.

Comparing the Lindquist *et al.* meta-analyses with the previous meta-analysis by Vytal and Hamann [43] reveals key similarities, but also differences that illustrate how the interpretation of evidence for basic emotions depends on the criteria used for evaluation. Both studies found evidence of consistent activations for basic emotions, using similar meta-analysis methods (closely related activation likelihood and activation density analyses). Each study tested emotion discriminability in markedly different ways, however, contrasting activation maps between all possible emotion pairs in one case [43], and contrasting each emotion with the average of all other emotions in the other [47].

Although methodological differences complicate direct comparisons between studies [51], a reasonable conclusion is that the meta-analytic evidence for discriminable neural correlates for basic emotions is currently inconclusive. In general, whereas meta-analytical findings have been

### Box 1. Dimensional and discrete models of emotion

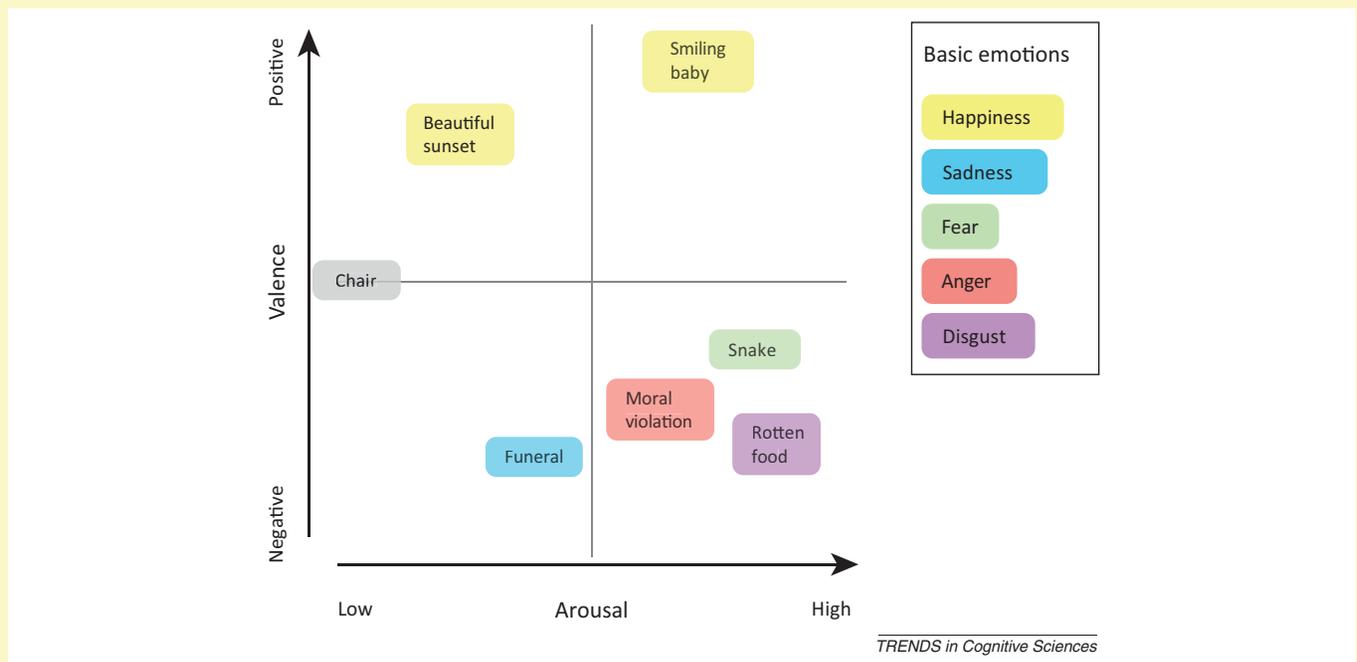
Dimensional and discrete theories of emotion are in fact families of related theories. For example, the most commonly proposed dimensional view posits the dimensions of valence (pleasantness) and arousal (emotion strength) [13,56], but alternate dimensions, such as positive and negative activation, have also been proposed. Dimensional theories differ in the minimum number of dimensions needed to represent emotion and the ways in which dimensions combine with other processes to create emotional experience [20,21]. An attractive feature of dimensional approaches is their parsimony and their applicability across multiple domains. Dimensional approaches have also proven to be empirically powerful, successfully accounting for a wide range of emotion effects [14,90,91].

Similarly, although the most commonly encountered discrete emotion theory, basic emotions, posits six basic emotions (happiness, sadness, fear, anger, disgust, and surprise), other discrete emotion theories have posited a different number of emotions [5,22,25]. Basic emotion theories propose that these emotion categories are biologically inherited and are basic in the sense that they cannot be broken down into constituent psychological elements.

To illustrate how basic emotions can be represented within a dimensional framework, Figure 1 shows how basic emotions can be represented as a combination of dimensions. For example, the figure illustrates different instances of basic emotions (e.g., seeing a snake

or baby), with each instance located in the two-dimensional space created by the arousal and valence dimensions. For reference, an emotionally neutral stimulus with minimal arousal and neutral valence is also shown (chair; grey rectangle). Different instances of a basic emotion can also vary in arousal and valence. For example, viewing a sunset may elicit happiness that is more moderate in arousal and valence than viewing a baby. The combination of arousal and valence information is not proposed to be the sole determinant of emotion states. Rather, these emotional building blocks are combined with other information and processes, such as appraisal and attribution, to create emotions [92].

A current debate in the neuroimaging literature is between discrete emotion views and psychological construction views (which originate from dimensional accounts). Psychological construction views, such as the conceptual act model, reject the notion that basic emotions are innate and biologically basic. Instead, they propose that instances of fear and other emotions result from the combination of 'core affect' (representations of somatic changes, experienced as valence and arousal) with other basic psychological processes that are not specific to emotion, such as attention and categorization [18]. This process of psychological construction results in specific instances (exemplars) of emotions that individuals experience as fear or other emotions, through a process markedly different from the triggering of biological affect programs posited by basic emotion theories [4].

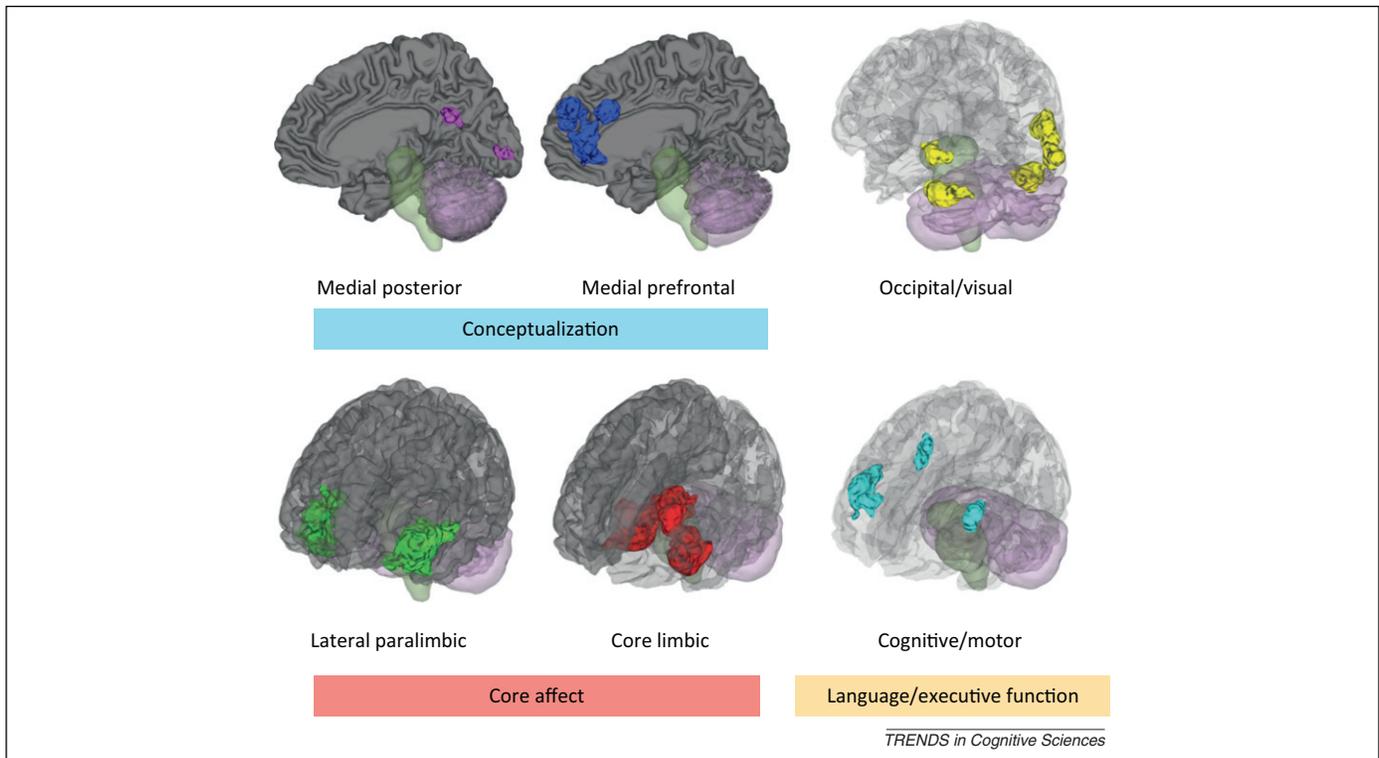


**Figure 1.** Representing basic emotions within a dimensional framework. Dimensional models can represent instances of basic emotions, for example, fear elicited by a snake (green rectangle), in terms of variation along affective dimensions (e.g., arousal and valence), in combination with additional evaluative processes, such as appraisal and attribution that specify affective significance [92].

interpreted as supporting aspects of basic emotion theories and the conceptual act model [18], neither view has so far received unequivocal support [52]. New neuroimaging studies directly contrasting the predictions of different emotion models using the same data will be key to evaluating the relative merits of emotion models.

Until recently, meta-analyses of the affective dimensions of arousal and valence dimensions were precluded because of the relative lack of neuroimaging studies manipulating both dimensions independently [53,54]. Several such studies have recently appeared [55–60], however, setting the stage for a meta-analysis and providing preliminary information regarding the neural correlates of

these dimensions. An initial view proposed that arousal and valence are represented separately, with amygdala activity tracking increasing arousal and different orbitofrontal cortex (OFC) regions representing positive (medial OFC) and negative (lateral OFC) valence [54,61]. Subsequent studies suggested a more complex representation that involves multiple regions in the representation of arousal, valence, as well as their interaction, implying that these dimensions are not represented independently [53,55–57]. For example, amygdala activation correlated with increasing arousal for words regardless of valence, but activity in other regions, including medial OFC and putamen, tracked the interaction of arousal and valence [53].



**Figure 2.** Six distributed functional groups consistently activated in a data-driven meta-analysis of emotion processing. In contrast to all of the other neuroimaging meta-analyses considered in this review, Kober *et al.* [50] adopted a unique data-driven approach to determine the functional organization of brain regions consistently activated in emotion tasks from 162 studies. A multi-level kernel-based approach was used to assess functional groupings across progressively broader spatial scales, from voxels to functional subregions, regions, and finally the functional group level across the brain. This procedure identified six functionally distributed groups that have been proposed by the conceptual act model to function as building blocks for the generation of emotional states [19,29]. The construction process is hypothesized to start with core affect (a lateral paralimbic group and a core limbic group), which is then processed further and given situation-specific conceptual meaning (medial posterior and medial prefrontal cortical groups) and participates in interactions with a functional group mediating attentional and linguistic function. A sixth group (occipital/visual group) is involved jointly in visual processing and attention to emotional stimuli. These functional groups derived from neuroimaging data independent of emotion categories represent an intriguing alternative approach to conceptualizing neural mechanisms mediating emotion.

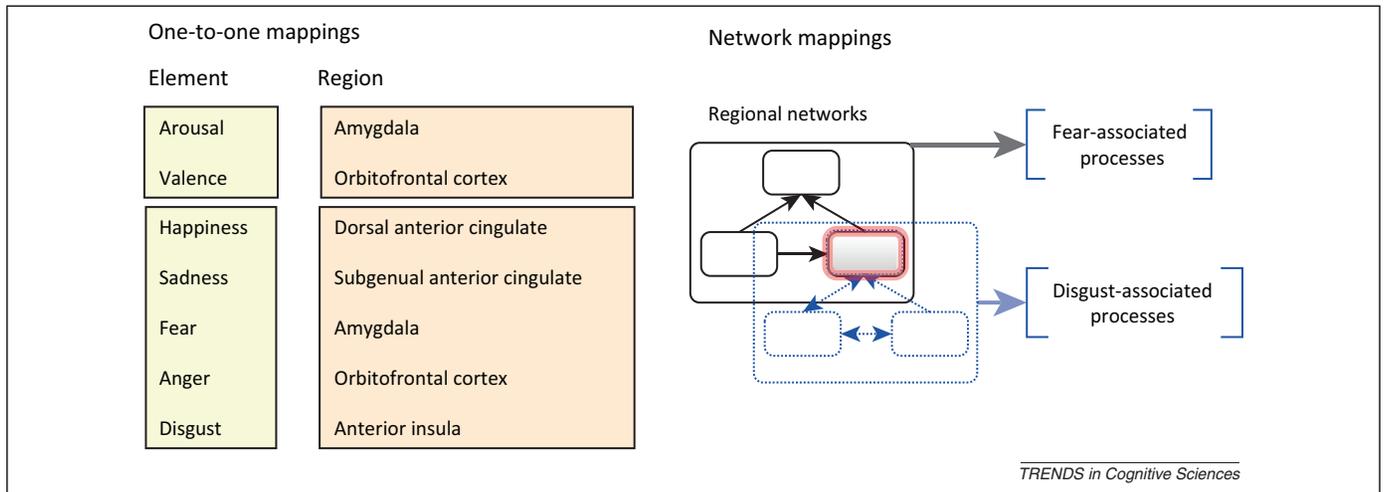
### Mapping emotions to networks

Two key issues in the debate over emotion models concern the type of mapping proposed and whether this mapping is at the level of individual brain regions or networks of regions. In the most basic type of mapping, each element of a given model is mapped onto a single brain region's activity [62]. Many neuroimaging studies investigating emotion have adopted this approach. The most commonly proposed mappings between the elements of dimensional and basic emotion theories and individual brain regions are illustrated in Figure 3 (left panel) [19,45,49]. However, it has become increasingly clear from meta-analyses and other evidence that approaches to mapping emotions that rely on simple one-to-one mappings between emotion categories and individual brain regions are ultimately insufficient and that more complex mappings are required to account for affective neuroimaging findings [52,62–64]. Individual regions can participate in multiple emotions and individual emotions can map onto activity in multiple regions [32]. For example, meta-analyses have demonstrated that disgust maps onto activations across multiple brain regions [18,43] and that the amygdala is associated with both positive and negative emotion [65–67], as well as with the basic emotions fear and disgust [68,69].

Meta-analytical findings suggest that a more appropriate level of mapping between emotion and brain is the level of functional networks [18,43]. This dovetails with other reviews that have suggested a need to shift to networks of

interacting regions as the basic unit of analysis and mapping for both emotion and cognition [62,63,70–72]. As illustrated in Figure 3 (right panel), such regional networks can share brain regions, and their components can carry out different aspects of emotion processing, with each component assuming different functional roles and computations depending on the particular configuration of the currently active network. For example, the amygdala (red box) is recruited during both fear and disgust [18,43]. This co-activation may reflect overlapping processes, or alternatively, the amygdala may take on different processing attributes in different network configurations [73]. A region's function can also be further influenced by individual differences, such as age and sex [74], as illustrated by a recent meta-analysis that found substantial sex differences in amygdala involvement (and other regions) for positive and negative emotions [75].

These considerations have prompted suggestions that the evaluation criteria and even the definitions of basic emotion categories may require modification. For example, the fact that brain regions can take on different functions at different times has led to the suggestion that regions can be functionally specialized for a given emotion, playing an emotion-specific role at a particular time, but playing other (possibly non-affective) roles at other times [52,64]. This contrasts markedly with the functional specificity criterion that has been used to evaluate basic emotion models [18,43], which requires a given region to have a specific



**Figure 3.** Levels of mapping between emotion models and the brain. The left panel illustrates the most commonly proposed one-to-one mappings between elements of emotion theories and individual brain regions [18,43]. For example, amygdala activation typically correlates with emotional arousal [101], whereas activation in the orbitofrontal cortex correlates with emotional valence [102]. As noted in the text, these one-to-one mappings run afoul of numerous experimental findings that show that, for example, fear consistently activates regions other than the amygdala, and the amygdala in turn is associated with several emotion processes. Such difficulties with one-to-one mappings have motivated a shift to more complex interrelationships, such as functional networks. For example, in the right panel, network mappings may involve individual brain regions (small rectangles) participating in networks that carry out the processing mediating different emotions. An individual region, such as the amygdala (red rectangle) can participate in multiple networks and that region's role can be modulated according to the currently active network configuration. These network dynamics have important implications for evaluating the neuroimaging evidence for different emotion theories.

emotional function to the exclusion of other emotional or cognitive functions, at all times. A further suggestion has been that basic emotion categories may turn out to map onto multiple subtypes of coherent emotion networks, and that splitting these categories to better reflect such networks may be needed [52,76] (but see also [19,29]). These possibilities suggest promising avenues for future development of basic emotion models.

### Multivariate approaches to emotion representations

Neuroimaging meta-analyses have not yet begun to examine correspondences between emotions and networks of brain regions, although multivariate meta-analysis methods have been applied in cognitive domains [77–79] and the data-driven meta-analysis discussed previously [50]. Therefore, whether basic emotion theories will gain greater support from assessment of network-level correspondences remains an open question. Multivariate pattern classifiers and machine learning algorithms represent a promising approach to detect highly distributed activation patterns corresponding to emotions [80–82]. An excellent illustration of this approach is a recent neuroimaging study that identified multivariate activation patterns in medial prefrontal cortex and left superior temporal sulcus corresponding to basic emotion categories (anger, disgust, fear, happiness, and sadness) [83]. Moreover, these patterns were invariant across stimulus modality (face, body, or voice) and the perceived intensity of the emotion [83], and were found in the absence of systematic emotion related activation differences in these regions, highlighting the ability of these methods to detect otherwise inaccessible representations [84].

This shift towards network-level analyses and multivariate approaches to analyzing brain activity parallels the same trend in studies of emotion that use psychophysiological approaches [23,85]. For example, a recent pattern classification study found substantial evidence for discriminable patterns corresponding to basic emotions, which

were also invariant across different emotion induction methods and individuals [23].

Neuroimaging meta-analyses have important limitations and cannot fully substitute for a lack of directly relevant neuroimaging studies [37,47]. New neuroimaging

### Box 2. Neuropsychological lesion studies and emotion models

Because neuroimaging studies cannot establish that brain regions are essential for a function such as emotion, the insights gained from neuroimaging studies require further corroboration from studies that can support causal claims, such as those that use neuropsychological approaches with patients with focal brain lesions [93,94]. Conversely, lesion studies have sometimes reported functional dissociations, where damage to a specific region is associated with impairment in one emotion but not others, which suggest one-to-one mappings between specific emotions and individual regions, contrary to the findings of neuroimaging studies.

An illustrative recent case is that of patient SM, a woman with focal bilateral amygdala lesions. SM shows a striking absence of overt fear behavior and a highly impaired experience of fear in daily life [95,96]. These impairments in fear experience were also specific to fear, as SM was normal when tested on other basic emotions. Impairments in experiencing fear were found across a wide range of potent fear-eliciting stimuli and situations, including exposure to live snakes and spiders, fear-inducing films and locations, and real-life experience sampling. Single case studies have well-known limitations, and these findings require replication before more definitive conclusions can be drawn regarding basic emotion views more generally. Also, other studies of patients with bilateral amygdala lesions have not described similar emotion deficits [89,97]. Differences in emotion assessment methods and the fact that patient SM sustained her lesions during development may explain these differing findings. Similarly emotion-specific impairments have also been reported for the perception and experience of the basic emotion disgust in patients with lesions to the basal ganglia and insula [98,99]. Such isolated impairments in individual emotions and other neuropsychological findings can serve an important role as constraints on the development of theories based on neuroimaging findings and illustrate the ongoing challenge of integrating emotion findings from different methods.

### Box 3. Emotion in nonhuman animals

Much of what is known regarding the neural systems mediating basic emotions such as fear comes from neuroscience studies of nonhuman animals [91]. A wide variety of techniques have addressed the neural mechanisms underlying fear behavior and fear learning, down to the level of molecular pathways. For example, in the well-studied case of fear conditioning, a neutral stimulus (tone) that predicts an aversive stimulus (shock) spurs associative learning in the lateral nucleus of the amygdala, which in turn projects to an array of structures (hypothalamus, periaqueductal grey, brainstem) that mediate signs of fear such as freezing. These basic neural fear circuits are similar across species from rodents to humans.

The fact that humans share these specialized, inherited neural mechanisms for fear and other emotions is often interpreted as key support for basic emotion theories [25]. However, an alternative view challenges this conclusion, pointing out that an emotion such as fear can elicit any of several adaptive responses, such as freezing, avoidance, or even attack [18]. Because these responses depend on different neural circuits, by this view, it cannot be concluded that the sole neural basis of fear has been identified, unless one posits multiple neural bases of fear. Regardless of the outcome of such debates, this example illustrates the complexities of making inferences from animal models of emotion to humans. The degree to which animal findings will inform ongoing debates regarding the nature of human emotion will depend on the strength and validity of such between-species inferences.

studies targeted at the issues discussed in this review will thus be essential. In addition, neuroimaging methods assess correlations between emotions and brain regions, not whether these regions are functionally essential [41,86,87]. Thus, converging evidence from studies using other methods, such as neuropsychological studies of patients with focal brain lesions (Box 2) and studies with nonhuman animals (Box 3), will ultimately need to be integrated with findings from neuroimaging to arrive at comprehensive emotion models [88,89].

### Concluding remarks

This review has highlighted the contributions of neuroimaging studies to recent debates over emotion theories and the representation of emotion in the human brain (see Box 4 for future research directions). Neuroimaging analyses have identified several brain regions consistently associated with different basic emotions, as well as neural

### Box 4. Questions for future research

- Network-based and multivariate pattern classification approaches are a promising next step in elucidating how elements of emotion models map onto brain processes. How can these approaches best be implemented at the level of individual neuroimaging studies and in meta-analyses?
- Neuropsychological lesion approaches can provide converging evidence addressing whether regions identified by neuroimaging are critical for specific emotion processes. However, the emotion processes associated with a given brain region can differ markedly depending on the method used, raising the challenge of integrating and reconciling findings across approaches.
- Direct comparison of the predictions of competing emotion models in neuroimaging studies are needed to complement the findings of meta-analyses.
- How can elements of basic emotion and psychological construction accounts be combined into hybrid accounts?

correlates for affective dimensions. Such information allows prediction of the brain regions likely to be activated during a particular emotion. However, predicting and distinguishing emotions from neuroimaging data has proven more challenging, in part because brain regions can participate in multiple emotions. This lack of one-to-one correspondences between emotions and brain regions has motivated a theoretical shift towards the investigation of network-level emotion representations, which can provide more appropriate and sensitive tests of the predictions of emotion models. Relatedly, multivariate methods have shown promise in decoding distributed networks representing basic emotions. Neuroimaging has played an important role in informing emotion theories, but the ultimate assessment of these theories requires a broader evaluation of evidence from multiple approaches, such as neuropsychological lesion and nonhuman animal studies. In conclusion, the contributions of neuroimaging to debates over the nature of emotions and their brain representation have already motivated new theoretical directions which in turn are likely to influence the evolution of new theories about how the human brain generates and represents emotions.

### Acknowledgments

I would like to thank Lisa Feldman Barrett, Luiz Pessoa, and Andrea Scarantino for helpful comments and suggestions during the preparation of this article. This work was supported by the Hope for Depression Research Foundation and the Atlanta Center for Behavioral Neuroscience.

### References

- 1 Russell, J.A. (2009) Emotion, core affect, and psychological construction. *Cogn. Emot.* 23, 1259–1283
- 2 Barrett, L.F. et al. (2007) The experience of emotion. *Annu. Rev. Psychol.* 58, 373
- 3 Ekman, P. (1992) An argument for basic emotions. *Cogn. Emot.* 6, 169–200
- 4 Ekman, P. and Cordaro, D. (2011) What is meant by calling emotions basic. *Emot. Rev.* 3, 364–370
- 5 Tomkins, S.S. (1962) *Affect, Imagery, Consciousness: Vol. 1. The Positive Affects*, Springer
- 6 Barrett, L.F. (2011) Was Darwin wrong about emotional expressions? *Curr. Dir. Psychol. Sci.* 20, 400–406
- 7 Panksepp, J. (2010) Affective consciousness in animals: perspectives on dimensional and primary process emotion approaches. *Proc. R. Soc. Lond. B: Biol. Sci.* 277, 2905–2907
- 8 Panksepp, J. (2007) Neurologizing the psychology of affects: how appraisal-based constructivism and basic emotion theory can coexist. *Perspect. Psychol. Sci.* 2, 281–296
- 9 Gable, P. and Harmon-Jones, E. (2010) The motivational dimensional model of affect: Implications for breadth of attention, memory, and cognitive categorisation. *Cogn. Emot.* 24, 322–337
- 10 Gerber, A.J. et al. (2008) An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia* 46, 2129–2139
- 11 Barrett, L.F. (1998) Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cogn. Emot.* 12, 579–599
- 12 Russell, J.A. and Barrett, L.F. (1999) Core affect, prototypical emotions, and other things called emotion: Dissecting the elephant. *J. Pers. Soc. Psychol.* 76, 805–819
- 13 Fosner, J. et al. (2009) The neurophysiological bases of emotion: an fMRI study of the affective circumplex using emotion-denoting words. *Hum. Brain Mapp.* 30, 883–895
- 14 Lang, P.J. and Bradley, M.M. (2010) Emotion and the motivational brain. *Biol. Psychol.* 84, 437–450
- 15 Feldman Barrett, L. and Russell, J.A. (1999) The structure of current affect: controversies and emerging consensus. *Curr. Dir. Psychol. Sci.* 8, 10–14

- 16 Russell, J.A. (2003) Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172
- 17 Wilson-Mendenhall, C.D. *et al.* (2011) Grounding emotion in situated conceptualization. *Neuropsychologia* 49, 1105–1127
- 18 Lindquist, K.A. *et al.* (2012) The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* 35, 121–143
- 19 Barrett, L.F. (2012) Emotions are real. *Emotion* 12, 413–429
- 20 Barrett, L.F. (2009) Variety is the spice of life: a psychological construction approach to understanding variability in emotion. *Cogn. Emot.* 23, 1284–1306
- 21 Fontaine, J.R.J. *et al.* (2007) The world of emotions is not two-dimensional. *Psychol. Sci.* 18, 1050–1057
- 22 Ekman, P. and Cordaro, D. (2011) What is meant by calling emotions basic. *Emot. Rev.* 3, 364–370
- 23 Stephens, C.L. *et al.* (2010) Autonomic specificity of basic emotions: evidence from pattern classification and cluster analysis. *Biol. Psychol.* 84, 463–473
- 24 Lench, H.C. *et al.* (2011) Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitation. *Psychol. Bull.* 137, 834–855
- 25 Izard, C.E. (2009) Emotion theory and research: highlights, unanswered questions, and emerging issues. *Annu. Rev. Psychol.* 60, 1–25
- 26 Rainville, P. *et al.* (2006) Basic emotions are associated with distinct patterns of cardiorespiratory activity. *Int. J. Psychophysiol.* 61, 5–18
- 27 Kreibig, S.D. *et al.* (2007) Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology* 44, 787–806
- 28 Cacioppo, J.T. *et al.* (2000) The psychophysiology of emotion. In *Handbook of Emotions* (2nd ed.) (Lewis, M. and Haviland-Jones, J.M., eds), pp. 173–191, Guilford Press
- 29 Barrett, L.F. (2006) Are emotions natural kinds? *Perspect. Psychol. Sci.* 1, 28–58
- 30 Bradley, M.M. *et al.* (2001) Emotion and motivation II: Sex differences in picture processing. *Emotion* 1, 300–319
- 31 Critchley, H.D. (2009) Psychophysiology of neural, cognitive and affective integration: fMRI and autonomic indicators. *Int. J. Psychophysiol.* 73, 88–94
- 32 Poldrack, R.A. (2010) Mapping mental function to brain structure: how can cognitive neuroimaging succeed? *Perspect. Psychol. Sci.* 5, 753–761
- 33 Bandettini, P.A. (2009) What's new in neuroimaging methods? *Ann. N. Y. Acad. Sci.* 1156, 260–293
- 34 de Gelder, B. (2010) The grand challenge for frontiers in emotion science. *Front. Psychol.* 1, 187 <http://dx.doi.org/10.3389/fpsyg.2010.00187>
- 35 Kriegeskorte, N. (2008) Representational similarity analysis – connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2, 4 <http://dx.doi.org/10.3389/neuro.06.004.2008>
- 36 Kriegeskorte, N. *et al.* (2006) Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3863–3868
- 37 Eickhoff, S.B. *et al.* (2009) Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum. Brain Mapp.* 30, 2907–2926
- 38 Barrett, L.F. and Wager, T.D. (2006) The structure of emotion: evidence from neuroimaging studies. *Curr. Dir. Psychol. Sci.* 15, 79–83
- 39 Mauss, I. and Robinson, M. (2009) Measures of emotion: a review. *Cogn. Emot.* 23, 209–237
- 40 Clithero, J.A. *et al.* (2011) NeuroImage within- and cross-participant classifiers reveal different neural coding of information. *NeuroImage* 56, 699–708
- 41 Huettel, S.A. (2011) Event-related fMRI in cognition. *NeuroImage* 62, 1152–1156
- 42 Damasio, A.R. *et al.* (2000) Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* 3, 1049–1056
- 43 Vytal, K. and Hamann, S. (2010) Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *J. Cogn. Neurosci.* 22, 2864–2885
- 44 Robinson, J.L. *et al.* (2009) Metaanalytic connectivity modeling: delineating the functional connectivity of the human amygdala. *Hum. Brain Mapp.* 31, 173–184
- 45 Laird, A.R. *et al.* (2005) ALE meta-analysis: controlling the false discovery rate and performing statistical contrasts. *Hum. Brain Mapp.* 25, 155–164
- 46 Salimi-Khorshidi, G. *et al.* (2009) Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. *NeuroImage* 45, 810–823
- 47 Wager, T.D. *et al.* (2009) Evaluating the consistency and specificity of neuroimaging data using meta-analysis. *NeuroImage* 45, S210–S221
- 48 Phan, K.L. *et al.* (2002) Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage* 16, 331–348
- 49 Murphy, F.C. *et al.* (2003) Functional neuroanatomy of emotions: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 207–233
- 50 Kober, H. *et al.* (2008) Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *NeuroImage* 42, 998–1031
- 51 Hamann, S. (2012) What can neuroimaging meta-analyses really tell us about the nature of emotion? *Behav. Brain Sci.* 35, 150–152
- 52 Scarantino, A. (2012) Functional specialization does not require a one-to-one mapping between brain regions and emotions. *Behav. Brain Sci.* 35, 161–162
- 53 Lewis, P.A. *et al.* (2007) Neural correlates of processing valence and arousal in affective words. *Cereb. Cortex* 17, 742–748
- 54 Anderson, A.K. *et al.* (2003) Dissociated neural representations of intensity and valence in human olfaction. *Nat. Neurosci.* 6, 196–202
- 55 Gerdes, A.B.M. *et al.* (2010) Brain activations to emotional pictures are differentially associated with valence and arousal ratings. *Front. Hum. Neurosci.* 4, 175 <http://dx.doi.org/10.3389/fnhum.2010.00175>
- 56 Nielen, M.M. *et al.* (2009) Distinct brain systems underlie the processing of valence and arousal of affective pictures. *Brain Cogn.* 71, 387–396
- 57 Colibazzi, T. *et al.* (2010) Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion* 10, 377–389
- 58 Sakaki, M. *et al.* (2012) Beyond arousal and valence: the importance of the biological versus social relevance of emotional stimuli. *Cogn. Affect. Behav. Neurosci.* 12, 115–139
- 59 Weierich, M.R. *et al.* (2010) Novelty as a dimension in the affective brain. *NeuroImage* 49, 2871–2878
- 60 Viinikainen, M. *et al.* (2010) Nonlinear relationship between emotional valence and brain activity: evidence of separate negative and positive valence dimensions. *Hum. Brain Mapp.* 31, 1030–1040
- 61 Hamann, S. (2003) Nosing in on the emotional brain. *Nat. Neurosci.* 6, 106–108
- 62 Pessoa, L. (2008) On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* 9, 148–158
- 63 Pessoa, L. (2012) Beyond brain regions: network perspective of cognition-emotion interactions. *Behav. Brain Sci.* 35, 158–159
- 64 Scarantino, A. and Griffiths, P. (2011) Don't give up on basic emotions. *Emot. Rev.* 3, 444–454
- 65 Kim, S.H. and Hamann, S. (2007) Neural correlates of positive and negative emotion regulation. *J. Cogn. Neurosci.* 19, 776–798
- 66 Hamann, S. and Mao, H. (2002) Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *Neuroreport* 13, 15–19
- 67 Ewbank, M.P. *et al.* (2009) The amygdala response to images with impact. *Soc. Cogn. Affect. Neurosci.* 4, 127–133
- 68 Sambataro, F. *et al.* (2006) Preferential responses in amygdala and insula during presentation of facial contempt and disgust. *Eur. J. Neurosci.* 24, 2355–2362
- 69 Mechias, M.-L. *et al.* (2010) A meta-analysis of instructed fear studies: implications for conscious appraisal of threat. *NeuroImage* 49, 1760–1768
- 70 Poldrack, R. (2006) Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63
- 71 Adolphs, R. (2010) Conceptual challenges and directions for social neuroscience. *Neuron* 65, 752–767
- 72 Adolphs, R. (2010) What does the amygdala contribute to social cognition? *Ann. N. Y. Acad. Sci.* 1191, 42–61

- 73 McIntosh, A.R. (2004) Contexts and catalysts: a resolution of the localization and integration of function in the brain. *Neuroinformatics* 2, 175–182
- 74 Kret, M.E. and De Gelder, B. (2012) Sex differences in processing emotional signals of others. *Neuropsychologia* 50, 1211–1221
- 75 Stevens, J.S. and Hamann, S. (2012) Sex differences in brain activation to emotional stimuli: a meta-analysis of neuroimaging studies. *Neuropsychologia* 50, 1578–1593
- 76 Scarantino, A. (2009) Core affect and natural affective kinds. *Philos. Sci.* 76, 940–957
- 77 Poldrack, R.A. (2011) The future of fMRI in cognitive neuroscience. *NeuroImage* 62, 1216–1220
- 78 Kaul, C. *et al.* (2011) The gender of face stimuli is represented in multiple regions in the human brain. *Front. Hum. Neurosci.* 4, 238 <http://dx.doi.org/10.3389/fnhum.2010.00238>
- 79 Diana, R.A. *et al.* (2008) High-resolution multi-voxel pattern analysis of category selectivity in the medial temporal lobes. *Hippocampus* 18, 536–541
- 80 Pantazatos, S.P. *et al.* (2012) Decoding unattended fearful faces with whole-brain correlations: an approach to identify condition-dependent large-scale functional connectivity. *PLoS Comput. Biol.* 8, e1002441
- 81 Anders, S. *et al.* (2011) Flow of affective information between communicating brains. *NeuroImage* 54, 439–446
- 82 Shackman, A.J. *et al.* (2011) The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* 12, 154–167
- 83 Peelen, M.V. *et al.* (2010) Supramodal representations of perceived emotions in the human brain. *J. Neurosci.* 30, 10127–10134
- 84 Bach, D.R. *et al.* (2011) A stable sparse fear memory trace in human amygdala. *J. Neurosci.* 31, 9383–9389
- 85 Kolodyazhniy, V. *et al.* (2011) An affective computing approach to physiological emotion specificity: toward subject-independent and stimulus-independent classification of film-induced emotions. *Psychophysiology* 48, 908–922
- 86 Kwok, S.C. (2011) Where neuroimaging and lesion studies meet. *J. Neuroimaging* <http://dx.doi.org/10.1111/j.1552-6569.2011.00600.x>
- 87 Friston, K.J. (2005) Models of brain function in neuroimaging. *Annu. Rev. Psychol.* 56, 57–87
- 88 Adolphs, R. (2008) Fear, faces, and the human amygdala. *Curr. Opin. Neurobiol.* 18, 166–172
- 89 Phelps, E.A. (2006) Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.* 57, 27–53
- 90 Kensinger, E. (2006) Processing emotional pictures and words: effects of valence and arousal. *Cogn. Affect. Behav. Neurosci.* 6, 110–126
- 91 Lang, P.J. and Davis, M. (2006) Emotion, motivation, and the brain: reflex foundations in animal and human research. *Prog. Brain Res.* 156, 3–29
- 92 Gendron, M. and Barrett, L.F. (2009) Reconstructing the past: a century of ideas about emotion in psychology. *Emot. Rev.* 1, 316–339
- 93 Berntson, G.G. *et al.* (2010) The insula and evaluative processes. *Psychol. Sci.* 22, 80–86
- 94 Tranel, D. *et al.* (2006) Altered experience of emotion following bilateral amygdala damage. *Cogn. Neuropsychiatry* 11, 219–232
- 95 Feinstein, J.S. *et al.* (2011) The human amygdala and the induction and experience of fear. *Curr. Biol.* 21, 34–38
- 96 Hamann, S. (2011) Affective neuroscience: amygdala's role in experiencing fear. *Curr. Biol.* 21, R75–R77
- 97 Anderson, A.K. and Phelps, E.A. (2002) Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. *J. Cogn. Neurosci.* 14, 709–720
- 98 Sprengelmeyer, R. (2007) The neurology of disgust. *Brain* 130, 1715–1717
- 99 Sprengelmeyer, R. *et al.* (1998) Neural structures associated with recognition of facial expressions of basic emotions. *Proc. R. Soc. Lond. B: Biol. Sci.* 265, 1927–1931
- 100 Tzourio-Mazoyer, N. *et al.* (2002) Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 289, 273–289
- 101 Hamann, S.B. *et al.* (2002) Ecstasy and agony: activation of the human amygdala in positive and negative emotion. *Psychol. Sci.* 13, 135–141
- 102 Goodkind, M.S. *et al.* (2011) Tracking emotional valence: the role of the orbitofrontal cortex. *Hum. Brain Mapp.* 33, 753–762