The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology

Jonathan Posner\textsuperscript{a,b}, James A. Russell\textsuperscript{c}, and Bradley S. Peterson\textsuperscript{a,b}

\textsuperscript{a}Columbia College of Physicians \& Surgeons
\textsuperscript{b}New York State Psychiatric Institute
\textsuperscript{c}Boston College

Abstract

The circumplex model of affect proposes that all affective states arise from cognitive interpretations of core neural sensations that are the product of two independent neurophysiological systems. This model stands in contrast to theories of basic emotions, which posit that a discrete and independent neural system subserves every emotion. We propose that basic emotion theories no longer explain adequately the vast number of empirical observations from studies in affective neuroscience, and we suggest that a conceptual shift is needed in the empirical approaches taken to the study of emotion and affective psychopathologies. The circumplex model of affect is more consistent with many recent findings from behavioral, cognitive neuroscience, neuroimaging, and developmental studies of affect. Moreover, the model offers new theoretical and empirical approaches to studying the development of affective disorders as well as the genetic and cognitive underpinnings of affective processing within the central nervous system.

The reigning experimental paradigm in affective neuroscience research posits that emotions can be divided into discrete and independent categories and that specific neural structures and pathways subserve each of these emotional categories. This theory of basic emotions has yielded significant advances in the understanding of affect and yet, in the fields of clinical psychology and psychiatry, it has left unsettled many important questions. The theory of basic emotions, for example, has not explained the near ubiquitous comorbid illnesses among mood disorders, nor has it resolved confusion over the neurophysiological underpinnings of affective disorders. Moreover, basic emotion theory is largely incompatible with recent findings in behavioral genetics and temperament research. Given these empirical and heuristic limitations of the theory of basic emotions, we propose that a shift is needed in the conceptual approaches taken to the study of emotion. We propose that clinicians and researchers move away from a strictly basic emotion model of affective states, where each emotion is thought to emerge from independent neural systems, to more dimensional models of emotions, in which all affective states are understood to arise from common, overlapping neurophysiological systems.

Although poorly represented in psychiatry, dimensional models have a long history in psychology (Larsen \& Diener, 1992; Russell, 2003; Schlosberg, 1952; Watson, Wiese, Vaidya, \& Tellegen, 1999). One particular dimensional approach, termed the circumplex model of affect, proposes that all affective states arise from two fundamental neurophysiological systems, one related to valence (a pleasure–displeasure continuum) and the other to arousal.
or alertness (Russell, 1980). Each emotion can be understood as a linear combination of these two dimensions, or as varying degrees of both valence and arousal (see Figure 1). Joy, for example, is conceptualized as an emotional state that is the product of strong activation in the neural systems associated with positive valence or pleasure together with moderate activation in the neural systems associated with arousal. Affective states other than joy likewise arise from the same two neurophysiological systems but differ in the degree or extent of activation. Specific emotions therefore arise out of patterns of activation within these two neurophysiological systems, together with cognitive interpretations and labeling of these core physiological experiences.

In addition to reviewing the literature relevant to dimensional theories of emotion, we will discuss the implications that dimensional approaches hold for psychiatric research and clinical practice. The circumplex model of affect in particular offers a conceptual and experimental framework for exploring the neural basis of affect that is also likely to provide insight into the neurophysiology of affective disorders. It also provides a theoretical basis for understanding the widespread comorbidities among mood and anxiety disorders, and its dimensional approach makes available powerful statistical and methodological tools for use in genetic, neuroimaging, and neurobiological studies of affective disorders.

Theories of Basic Emotions

The dominant theory of emotion in psychiatric and neuroscience research posits that humans are evolutionarily endowed with a discrete and limited set of basic emotions (Ekman, 1992; Panksepp, 1998; Tomkins, 1962, 1963). Each emotion is independent of the others in its behavioral, psychological, and physiological manifestations, and each arises from activation within unique neural pathways of the central nervous system (CNS). Fear, by this account, produces aversive feelings and behaviors that are related to activation within specific neural pathways. Other emotions are produced through activation in separate neural pathways distinct from the fear system, and likewise they produce subjective feelings, peripheral nervous systems patterns, and behaviors associated with that specific emotional state. This is a theory in which each specific emotion maps to one neural system.

Animal studies

The conceptualization of emotions as discrete and independent has arisen largely from affective research with animals. By selectively stimulating neural pathways and observing subsequent behaviors, or conversely by eliciting behaviors in highly constrained experimental circumstances and measuring neural activity, animal researchers have constructed taxonomies of the basic emotions and proposed specific neural pathways associated with each putative basic emotion (Panksepp, 1998). Although this experimental approach has helped researchers begin to explore the neural basis of emotion, it is inherently limited in the information that it provides about affective experiences and the neural systems that support them. Researchers are forced to attribute affective states to animals based on the behaviors that the animals exhibit. Affective behaviors, however, are neither sufficient nor necessary to characterize emotional states (Kagan, 2003; Panksepp, 1998). Anxiety, for example, can be felt without any overt changes in behavior, just as affective behaviors, such as frowning and smiling, can be elicited without any obvious changes in subjective feeling. An animal, therefore, could experience feelings without demonstrating any overt changes in behavior, and conversely, through experimental manipulation, could display affective behaviors without any associated feeling. As Damasio (2003) notes, paramecium will flee from predators, but to argue that single-celled organisms experience fear makes little sense.

Aware of these limitations in their experimental paradigms and the possible difficulties with interpreting their results, emotion researchers assert that the validity of many of the findings...
from animal research must ultimately be confirmed through human studies of emotion (Panksepp, 1998). This confirmation, however, has proved elusive, and inconsistencies between the findings from human studies of emotion and those predicated from animal research abound (Berridge, 2003; Davidson, 2003). In short, and at the risk of oversimplification, these inconsistencies might be summarized as follows: animal research emphasizes the role of subcortical and other evolutionarily primitive structures in the processing of emotions, whereas human research demonstrates the importance of neocortical structures in emotional experience (Berridge, 2003). Given animal researchers' necessary reliance on overt behavior, their results could reasonably be interpreted as delineating neural systems related primarily to affective behaviors rather than subjective feelings as described in the human literature. Indeed, abundant data from lesion studies and neuroimaging investigations demonstrate that activation of the prefrontal cortex participates centrally in the experiencing of positive and negative emotions (Davidson, Ekman, Saron, Senulis, & Friesen, 1990).

Facial expressions and the human physiological correlates of basic emotions

In addition to conducting animal studies of putative affective processes, basic emotion theorists have investigated emotional processes in humans by exploring facial expressions and peripheral physiological responses to affective stimuli. These investigators have assumed that patterns of autonomic activation and facial innervation are specific to each basic emotion (Ekman, 1992; Ekman, Levenson, & Friesen, 1983). Evidence to support this hypothesis, however, is limited. In a meta-analysis of studies correlating autonomic activity with affective states, for example, Cacioppo concludes that the basic emotions have not been found to be associated with specific patterns of autonomic activation (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000). The physiological measures associated with a single basic emotion often differ significantly depending on the nature of the eliciting stimuli (Hamm, Gerlach, Globisch, & Vaitl, 1992). Similar problems plague animal models of emotion, in that dissimilar physiological responses are observed in association with a single basic emotion (Iwata & LeDoux, 1988). Conversely, disparate basic emotions in both animal and human studies are often associated with similar physiological responses (Cacioppo et al., 2000).

Affective researchers have also proposed that each basic emotion is associated with a characteristic facial expression. If this were true, facial expressions would then provide overt criteria for classifying the basic emotions because the basic emotions could be defined simply by the presence of a characteristic facial expression. This proposal, however, has been largely discredited. As Ekman notes, not all emotions are accompanied by a characteristic facial expression (Ekman, 1993). Moreover, certain facial expressions are associated with more than one emotion (e.g., a smile, is associated variously with happiness, pride, and condescending sarcasm). This poor specificity in the emotional correlates of facial expressions suggests that the taxonomy of facial expressions, as described by Ekman and Izard, does not describe adequately the taxonomy of emotions. Facial expressions may sometimes communicate information about, among other things, an individual's affective state, but they do not delineate it (Camras, 1992; Fernandez–Dols & Ruiz–Belda, 1997).

Developmental studies

A final approach to studying the basic emotions has been the detailed analysis and exploration of infants' affective responses, both in controlled experimental settings and in their naturalistic environments. Investigators have argued that the rudiments of the basic emotions are present in infants, are observable soon after birth, and surface prior to the emergence of language or other necessary cognitive mechanisms described in dimensional models of emotion (Kopp & Neufeld, 2003). These investigators argue that the cognitive interactions necessary for specific emotional responses, as described by dimensional models, cannot possibly be present in preverbal infants, who lack sophisticated cognitive capacities. We suspect, however, that this
developmental approach assumes a similar empirical perspective to the study of emotion as that taken by animal models of affect: the behaviors of nonverbal subjects are observed and then interpreted by researchers accustomed to classifying affective behaviors into specific emotional categories. In other words, although the cognitive processes necessary for subjectively experiencing and labeling specific emotions are not present in the infants, such hardware is present in the researchers who are labeling the behavior.

We also suspect that although developmental studies have not yet excluded the necessity of cognition in affective responses, when taken together with animal research in which specific affective responses are elicited through electrical stimulation of neural pathways, this body of literature may suggest that discrete neural pathways underlie specific affective behaviors. These affective behaviors, however, are not identical to subjective feelings. Eliciting a startle response in an infant or animal, for example, is not equivalent to eliciting fear. Similarly, identifying the neural pathways associated with a startle response should not de facto be regarded as the same as identifying the neural circuitry that subserves fear.

**Summary of basic emotion theory**

Recent findings from studies of affective neuroscience have posed significant challenges to the theory of basic emotions. The neural foundations of basic emotions have not yet been validated, peripheral physiological correlates for the basic emotions have not been established, and specific facial expressions associated with each basic emotion have not been identified. Taken together, these shortcomings leave the theory of basic emotions without a sufficient empirical foundation thus far for defining which emotions are indeed basic (Ortony & Turner, 1990).

Basic emotion theorists have primarily explored the behavioral and expressive manifestations of emotion. Investigations of the subjective, or experiential, components of emotion, rather than supporting a one to one correspondence between a discrete emotion and an underlying neural system, have instead suggested that emotions arise from cognitive interpretations of core physiological experiences (Cacioppo et al., 2000; Russell, 2003). These studies of the subjective components of emotion have motivated the development of dimensional models of emotion to understand and explore the core physiological bases of affective experiences. Such models are of particular interest to studies of affective and clinical neurosciences because they suggest new interpretations and novel experimental approaches to neuroimaging, genetic, and developmental studies of emotion and affective illnesses. We will review these approaches and show how they pertain to one particular dimensional model of emotion, the circumplex model of affect, which we believe will provide a useful conceptual framework for further exploring and understanding affect and affective disorders.

**The Circumplex Model of Affect**

Clinicians and researchers have long noted the difficulty that people have in assessing, discerning, and describing their own emotions (Saarni, 1999). This difficulty suggests that individuals do not experience, or recognize, emotions as isolated, discrete entities, but that they rather recognize emotions as ambiguous and overlapping experiences. Similar to the spectrum of color, emotions seem to lack the discrete borders that would clearly differentiate one emotion from another (Russell & Fehr, 1994). Indeed, researchers exploring the subjective experience of emotion have noted that emotions are highly intercorrelated both within and between the subjects reporting them (Russell & Carroll, 1999; Watson et al., 1999). Subjects rarely describe feeling a specific positive emotion without also claiming to feel other positive emotions (Watson & Clark, 1992). These intercorrelations among emotions, often obscured in experimental paradigms of basic emotions, are addressed head-on by dimensional models of...
affect. Dimensional models regard affective experiences as a continuum of highly interrelated and often ambiguous states.

Extensive and detailed study of the intercorrelations among emotional experiences, using statistical techniques such as multidimensional scaling and factor analysis of subjective reports of emotional words, faces, and experiences, has repeatedly yielded two-dimensional (2-D) models of affective experience (Larsen & Diener, 1992). These dimensions have been conceptualized in different ways: as the dimensions of positive and negative affect (Watson et al., 1999), tension and energy (Thayer, 1989), approach and withdrawal (Lang, Bradley, & Cuthbert, 1998), or valence and arousal (Russell, 1980). Despite the differing descriptive labels applied to these dimensions, the 2-D structure is found consistently across a large number of studies. In interpreting this 2-D structure, proponents of the circumplex model of affect suggest that all affective states arise from two independent neurophysiological systems, which, for the purposes of discussion here, we term the valence and arousal systems. Each and every affective experience is the consequence of a linear combination of these two independent systems, which is then interpreted as representing a particular emotion (see Figure 1). Fear, for example, is conceptualized by circumplex theorists as a neurophysiological state typically involving the combination of negative valence and heightened arousal in the CNS. The subjective experience of fear arises out of cognitive interpretations of these patterns of physiological activity that occur in the context of eliciting stimuli. As emotions are experienced and communicated, cognitive interpretations are employed to identify the neurophysiological changes in the valence and arousal systems and conceptually organize these physiological changes in relation to the eliciting stimuli, memories of prior experiences, behavioral responses, and semantic knowledge (Russell, 2003).

Emotions can therefore be seen as the end product of a complex interaction between cognitions, likely occurring primarily in neocortical structures, and neurophysiological changes related to the valence and arousal systems, which presumably are subserved largely by subcortical structures. We now briefly review the experimental evidence that supports this model, paying particular attention to the implications that this theoretical and experimental paradigm holds for basic and clinical research in the affective neurosciences.

**Empirical support from psychometric studies**

Using the statistical tools of factor analysis and multidimensional scaling, a wide variety of psychological assessments have demonstrated the emergence of two underlying, or latent, dimensions of emotion when individuals label and communicate either their own affective states or the affective states of others (Feldman Barrett & Fossum, 2001; Larsen & Diener, 1992; Russell, 1980; Watson, Clark, & Tellegen, 1988). Researchers have consistently reproduced the 2-D structure of the circumplex model using similarity ratings of facial expressions (Abelson & Sermat, 1962; Cliff & Young, 1968; Russell & Bullock, 1985; Schlosberg, 1952) and emotion-denoting words (Bush, 1973; Kring, Barrett, & Gard, 2003; Russell, 1980). Many of these findings have been replicated in cross-cultural samples (Russell, 1983). Moreover, self-reports of affective states studied over various time frames, languages, and response formats have repeatedly yielded 2-D models of emotion (Feldman Barrett & Russell, 1998; Russell, 1980; Watson et al., 1988). Simply stated, these robust and often replicated findings indicate that subjects who report feeling sad are also likely to report feeling angry, down, guilty, and so forth, whereas subjects who report feeling good are also likely to report experiencing other positively valenced emotions (Watson & Clark, 1992).

Biological correlates have been found repeatedly for the two dimensions of emotional experience predicted by psychometric studies of emotions. Using standardized emotional probes, researchers have observed that peripheral physiological responses to affective stimuli vary incrementally with subjective ratings of valence and arousal. Indeed, investigators have
correlated increases in skin conductance and heart rate accelerations with subjective ratings of arousal (Lang, Greenwald, Bradley, & Hamm, 1993). A similar correlation has also been reported for functional magnetic resonance imaging signal intensity in the visual cortex and ratings of arousal when viewing emotionally evocative pictures (Bradley, Sabatinelli, Lang, Fitzsimmons, King, & Desai, 2003). EEG studies have demonstrated similar increases in cerebral activation in relation to subjective ratings of arousal (Keil, Muller, Gruber, Wienbruch, & Stolarova, 2001), and animal studies have established that EEG measures of arousal increase with increasing electrical stimulation of the reticular formation (RF; Heilman, Watson, & Valenstein, 1993; Moruzzi & Magoun, 1949).

Similarly, valence ratings have been correlated with facial electromyographic measurements of the corrugator and zygomatic musculature (Cacioppo, Petty, Losch, & Kim, 1986; Lang et al., 1993). Corrugator activity increases incrementally with negative valence ratings regardless of the specific affective state described by the subject. Conversely, zygomatic muscle activity increases incrementally with positive valence ratings.

Investigators using the statistical tool of factor analysis to explore further the physiological correlates of affective experience generally report two-factor solutions to account for variance in the labeling of affective experience. The first factor, associated strongly with hedonic judgments, correlates highly with measures of activity in the zygomatic and corrugator muscles of the face that participate in emotional expression. The second factor, associated strongly with ratings of arousal or emotional excitement, correlates highly with measures of skin conductance and thus with activity in the sympathetic nervous system (Bradley, 2000; Lang et al., 1993). These correlations are largely inconsistent with the position of basic emotion theorists that specific physiological responses to affective stimuli are associated with discrete affective states, and instead complement the 2-D approach of the circumplex model of affect.

Valence–neural circuitry

A large body of evidence from studies in affective neuroscience has established the role of the mesolimbic dopamine system in processing reward and pleasure. The mesolimbic system projects from the ventral tegmental area to the nucleus accumbens and has abundant connections to the prefrontal cortex, amygdala, and hippocampus (see Figure 2). The role of the mesolimbic system in reward has perhaps been most convincingly demonstrated in studies of both animals and humans in which repetitive electrical self-stimulation of the orbitofrontal cortex and ventral striatum (which includes the nucleus accumbens and caudate nucleus) is easily sustained, even at the risk of starvation and dehydration (Mora, Avrith, & Rolls, 1980; Rolls, Burton, & Mora, 1980). Moreover, studies of drug addiction demonstrate increased cerebral blood flow to these same regions following intoxication with a variety of drugs, and these increases in cerebral blood flow correlate strongly with subjective ratings of euphoria (Drevets, Gautier, Price, Kupfer, Kinahan, Grace, Price, & Mathis, 2001; Ingvar, Ghatan, Wirsen–Meurling, Risberg, Von Heijne, Stone–Elander, & Ingvar, 1998; Mathew, Wilson, Humphreys, Lowe, & Wiethe, 1992; Nakamura, Tanaka, Nomoto, Ueno, & Nakayama, 2000; Volkow, Mullani, Gould, Adler, Guynn, Overall, & Dewey, 1988).

Although the mesolimbic dopamine system is associated primarily with pleasure and reward, it also is thought to subserve dysphoric states in several ways. Studies of drug addiction, for example, suggest that hypoactivation or understimulation of the mesolimbic system is associated with a range of negative emotions (Diana, Pistis, Muntoni, & Gessa, 1996) that drive drug-seeking behavior (Diana, Melis, Muntoni, & Gessa, 1998; Goldstein & Volkow, 2002). As addictive drugs are consumed, increasing neural activity in the mesolimbic system stimulates reward pathways, thereby attenuating the dysphoria associated with drug abstinence and withdrawal. In addition, activity in the ventral striatum, a central component of the mesolimbic system, is known to increase from baseline as subjects anticipate and respond to
aversive stimuli (Jensen, McIntosh, Crawley, Mikulis, Remington, & Kapur, 2003; Seymour, O’Doherty, Dayan, Koltzenburg, Jones, Dolan, Friston, & Frackowiak, 2004). These studies together suggest that differential activity within the ventral striatum may signal differing responses within the mesolimbic system to emotional valences, both positive and negative. In addition to the contribution of dopaminergic systems to activity within valence systems, serotonergic projections from the dorsal raphe nucleus to the ventral striatum may contribute to the modulation of dysphoric feelings (Daw, Kakade, & Dayan, 2002). Indeed, interactions between these dopamine and serotonin pathways are thought to underlie the therapeutic effects of certain antidepressants (Dremencov, Gispan–Herman, Rosenstein, Mendelman, Overstreet, Zohar, & Yadid, 2004).

Finally, numerous studies have reported associations of positive and negative emotions with asymmetry of activity in the frontal lobe, particularly in the prefrontal cortex (Davidson, 1984; Heilman, 2000; Sackeim, Greenberg, Weiman, Gur, Hungerbuhler, & Geschwind, 1982). Lateralized activity of the prefrontal cortex in patients with frontal strokes (Babinski, 1914; Gainotti, 1972; Goldstein, 1948; Heilman, 2000; Robinson, Starr, & Price, 1984) and in numerous EEG studies generally suggest that greater activation of the right frontal lobe accompanies the experience of more negatively valenced emotions, whereas greater left frontal activation accompanies the experience of more positively valenced experiences (Ahern & Schwartz, 1985; Davidson, 1984; Jones & Fox, 1992; Tucker, Stenslie, Roth, & Shearer, 1981). Whether and how this asymmetry of activity in the prefrontal cortex affects the mesolimbic system is unclear, but the abundant interconnections of the mesolimbic system and prefrontal cortex suggest that lateralized prefrontal activity likely has consequences within the mesolimbic reward system.

In sum, the mesolimbic system has long been associated with pleasure and reward, but this system also likely plays a significant role in the experience of negative emotions. It may therefore represent a neural substrate for the valence dimension proposed by the circumplex model of affect.

**Arousal–neural circuitry**

The RF is thought to regulate arousal levels of the CNS through its connections with the limbic system and thalamus (Heilman, 2000; Jones, 2003; Figure 3). Sensory stimuli are relayed from the thalamus to the amygdala, where investigators have suggested that the neural representations of the emotional significance reside (LaBar & LeDoux, 2003; LeDoux, 1996; Sander, Grafman, & Zalla, 2003). Indeed, the amygdala seems to respond both to appetitive and aversive stimuli, with greater activations accompanying the presentation of more arousing stimuli (Anderson, Christoff, Stappen, Panitz, Ghahremani, Glover, Gabrieli, & Sobel, 2003; Davis & Whalen, 2001; Small, Gregory, Mak, Gitelman, Mesulam, & Parrish, 2003). These assessments of emotional arousal are then relayed to the RF through the amygdaloreticular pathways (Koch & Ebert, 1993; Rosen, Hitchcock, Sananes, Miserendino, & Davis, 1991). Some investigators have also suggested that these determinations of emotional arousal are relayed from the amygdala to the RF through the association cortices of the parietal lobe (Figure 3; Heilman, 2000).

When encountering emotionally arousing stimuli, increasing activity in the RF increases activity in the intralaminar nuclei of the thalamus via excitatory projections (Jones, 2003; Steriade, 1996). Immunohistochemical studies suggest that ascending neural projections from the RF are composed largely of neurons producing the excitatory neurotransmitter, glutamate (Jones, 1995; Kaneko, Itoh, Shigemoto, & Mizuno, 1989). An additional smaller body of inhibitory interneurons within RF produces γ-aminobutyric acid and activation of these neurons cause decreased activity within the RF, thalamus, and cerebral cortex (Jones, 1995). The glutamatergic projections from the ascending RF have an excitatory effect on the intralaminar...
nuclei of the thalamus, which in turn increase activity broadly throughout the cerebral cortex, but particularly in the primary and association cortices (Heilman et al., 2003; Jones, 2003). Descending tracts from the RF form the spinoreticular pathways (Jones, 2003) that modulate muscle tone and activity in the sweat glands, both of which are highly correlated with subjective ratings of emotional arousal (Lang et al., 1993). Conversely, lesions to the reticular network produce emotional hypoarousal, to the extent that bilateral lesions can produce coma (Heilman et al., 2003). Likewise, amygdala lesions in humans impair the ability to recognize affective stimuli; in nonhuman primates, these lesions produce abnormal behavioral responses in which both aversive and appetitive stimuli are treated similarly to nonarousing and nonemotional stimuli (Aggleton & Passingham, 1981; Horel, Keating, & Misantone, 1975; Myers & Swett, 1970; Weiskrantz, 1956). Conversely, states of emotional hyperarousal, such as panic and mania, are associated with increased activity in both the amygdala and RF networks (Rauch, Savage, Alpert, Fischman, & Jenike, 1997; Rauch, Shin, & Wright, 2003; Skinner, Homma, & Garcia–Rill, 2004).

Cognitive interactions with the valence and arousal systems

We have suggested that the mesolimbic and reticular networks support the phylogenetically primitive sensations of pleasure and arousal. According to the circumplex model of affect, cognitive schemas identify and instantiate these core neurophysiological sensations. We suspect that the prefrontal cortex primarily subserves these cognitive functions, interpreting and acting on signals from the mesolimbic and reticular networks to facilitate the conscious recognition and experience of changes in these core neurophysiological systems. Damasio has suggested that the prefrontal cortex is essential for weighing the future consequences of affectively salient decisions, particularly as they relate to reward and punishment (Damasio, 1994). We suggest a similar role for the prefrontal cortex, within the context of the circumplex model of affect, whereby the prefrontal cortex integrates, organizes, and structures the primitive sensations of pleasure and arousal with knowledge of the temporal contingencies that link prior experiences of stimuli within varying life contexts with expectations for the future (Fuster, 1997). The prefrontal cortex, for example, likely integrates the sensations of pleasure and elevated arousal when winning a lottery ticket with knowledge of the present context, recollection of previous financial hardship, and expectations of future advantage that the windfall will bring, to create the conscious experience of joy. The cognitive functions of the prefrontal cortex support the creation and conscious recognition of specific emotions by associating and integrating core neurophysiological sensations with specific internal and external cues.

Social psychologists have long noted the importance of cognitive appraisals in affective experience, with contextual and idiosyncratic factors significantly influencing self-reports of affective experience (Fridja, 1986; Lazarus, 1991). Self-reports of affective states can vary dramatically when contextual cues are altered, even when the core physiological responses to those cues are identical (Cacioppo et al., 2000). The experience of the threat of physical harm, for example, can variously produce a pleasurable excitement, as when on a roller coaster or sky diving, or alternatively intense fear, as when riding in a car that is out of control or when falling from a precipice. The core physiological sensation in response to the physical threat is likely similar across these experiences. The differing emotional response comes from the differing integrations of memories of past consequences in similar contexts with assessments of varying contingencies in the present context that are likely to produce either similar or different consequences in the immediate future.

Preliminary human imaging studies support the likely importance of the prefrontal cortex in integrating and modifying information gleaned from emotional systems. In a functional imaging study exploring the effects of cognitive appraisals on emotional experiences, activity...
in the medial and lateral prefrontal cortex increased as subjects consciously altered their appraisals of aversive stimuli (Ochsner, Bunge, Gross, & Gabrieli, 2002). Moreover, activity in brain regions associated with the experience of negative emotional valence correlated inversely with activity in the prefrontal cortex as subjects tried consciously to minimize their negative affective responses to the aversive stimuli. These findings suggest that cognitive reappraisals alter activity in emotional circuits within the subcortex through changes in activity of the prefrontal cortex. Likewise, lesion studies have noted the role of the prefrontal cortex in organizing and communicating affective states, highlighting the importance of the prefrontal cortex in making decisions that have strong emotional consequences (Damasio, 1994). Thus, consistent with the circumplex model of affect, preliminary empirical findings suggest that the prefrontal cortex may serve the function of interpreting the sensations of pleasure and arousal within varying situational contexts, culminating in the subjective experience of a specific emotion within a particular situational and historical context.

Implications for Affective Neuroscience

Comorbidity

The circumplex model of affect offers a conceptual framework for addressing important empirical questions in clinical psychiatry and psychology. Mood and anxiety disorders, for example, co-occur at remarkably high rates, both within individuals and within families, and in both clinical and epidemiological samples (Gorman, 1996; Levine, Cole, Chengappa, & Gershon, 2001). Similar genetic vulnerabilities have been implicated in both classes of disorders, and their pharmacological treatments overlap substantially (Axelson & Birmaher, 2001; Weizman & Weizman, 2000). Together, these shared clinical features suggest the presence of a common underlying pathophysiology for mood and anxiety disorders. Nevertheless, a useful theoretical or empirical account of what this underlying pathophysiology may be has yet to be established. In part, failure to account for comorbidities among affective illnesses may stem from adherence to a theory of basic emotions in which a specific affective state, such as “anxiety,” is thought to be a discrete entity that is easily distinguished from sorrow, sadness, shame, or other dysphoric emotions. The circumplex model, in contrast, suggests that all of these dysphoric emotions emerge from highly intercorrelated activity within the valence system of the CNS. Graphically, these dysphoric emotions are all positioned within the left half of the circumplex representing negatively valenced experiences (Figure 1). Varying arousal levels and differing cognitive appraisals, given past experiences and the current context, further differentiate these sensations of negative valence.

We suspect that the cognitive schemas that are used to interpret the underlying neurophysiological activations in valence and arousal systems of the CNS further shape and distort distinctions between certain emotions. For some individuals, overlapping cognitive schemas may diminish their capacity to distinguish similarly valenced emotions, such as sadness, anxiety, shame, and fear (Feldman, 1995). Given such cognitive distortions, we suspect that these individuals experience anxiety and depression as similar or highly related affective states. Within the graphical representation of their affective circumplexes, we expect that their ratings for anxiety and depression will overlap, instead of being more clearly distinguished by differing arousal levels (Figure 4). This relative collapsing or compression of the circumplex along the arousal dimension has been termed “valence focus” (Feldman, 1995). The structure of the circumplex for these individuals is determined almost exclusively by changes in emotional valence, with little differentiation along the arousal dimension.

A common underlying pathophysiology for depression and anxiety disorders is supported by neuroimaging studies of these conditions. Both illnesses are characterized by an increased prevalence of negative affect, which we believe is caused in part by abnormal activity in the mesolimbic system and prefrontal cortices, the circuits that we postulate to subserve,
respectively, emotional valence and its cognitive overlay (Davidson, 2002; Henriques & Davison, 1990, 1991; Rauch et al., 1997; Schaffer, Davidson, & Saron, 1983; Wu, Buchsbaum, Hershey, Hazlett, Sicotte, & Johnson, 1991). Anxiety disorders have been associated repeatedly with increased activity of the right prefrontal cortex and amygdala, similar to the abnormalities reported in depressed subjects (Drevets, 2003; Rauch et al., 1997; Shin, Kosslyn, McNally, Alpert, Thompson, Rauch, Macklin, & Pitman, 1997). Complementing these findings from functional imaging studies, both depressed and anxious subjects demonstrate morphological abnormalities in mesolimbic and prefrontal brain regions (Bremner, 2002; Drevets et al., 1998; Robinson, Wu, Munne, Ashtari, Alvir, Lerner, Koreen, Cole, & Bogerts, 1995). Considerable evidence from preclinical studies in animals and neuroimaging studies in humans suggests that antidepressant medications alter activity in the orbitofrontal cortex and mesolimbic system (Diler, Kibar, & Avci, 2004; Drevets, Bogers, & Raichle, 2002; Kennedy, Evans, Kruger, Meyer, McCann, Arifuzzman, Houle, & Vaccarino, 2001). We suspect that antidepressant medications normalize activity within this valence system of the CNS, thereby accounting for the efficacy of these agents in treating both anxiety and depressive disorders. Antianxiety medications and other CNS depressants, on the other hand, likely return heightened arousal levels to the normal range, thereby providing symptomatic relief to patients with either anxiety disorders or anxious depression. These agents do not help, and they can even exacerbate melancholic depression, as they have little effect of the valence system and may add to the hypoarousal associated with melancholia (Kravitz, Fawcett, & Newman, 1993; Rickels, Schweizer, Case, DeMartinis, Greenblatt, Mardos, & Garcia Espana, 1998).

The arousal dimension of the circumplex model is also useful in understanding the patterns of other common psychiatric comorbidities. Attention-deficit/hyperactivity disorder (ADHD), bipolar disorder, and anxiety disorders all commonly co-occur within individuals and within families (Pliszka, 1998; Sasson, Chopra, Harrari, Amitai, & Zohar, 2003). Although ADHD is not classified as an affective disorder, affective symptoms are frequently noted in these patients. Indeed, the DSM-IV includes affective symptoms as an associated feature of ADHD, and familial predispositions for ADHD have been associated repeatedly with predispositions for affective disorders (Doyle & Faraone, 2002; Faraone, Glatt, & Tsuang, 2003). Consistent with the circumplex model of affect, several investigators have noted that ADHD, bipolar disorder, and anxiety disorder are all characterized by hyperarousal (Biederman & Spencer, 1999; Swann, Katz, Bowden, Berman, & Stokes, 1999). Indeed, aberrant patterns of activation within arousal systems of the CNS have been noted in physiological studies of persons with these disorders (Mefford & Potter, 1989; Skinner et al., 2004; Ucles, Lorente, & Rosa, 1996). Delayed myelination of the RF, for example, has been suggested in children with ADHD in findings from a transcranial magnetic stimulation study (Ucles et al., 1996). Individuals with anxiety disorders demonstrate excessive activations of their amygdala when viewing fearful faces and phobic objects (Rauch et al., 2003), and morphological abnormalities of the amygdala have been reported in adolescents and adults with bipolar disorder (Blumberg, Kaufman, Martin, Whiteman, Gore, Charney, Krystal, & Peterson, 2003; DelBello, Zimmerman, Mills, Getz, & Strakowski, 2004). These findings imply that overlapping neural circuits in the limbic and reticular networks may be dysfunctional in each of these seemingly disparate Axis I disorders.

**Neuroimaging studies and emotion: Methodological concerns**

Neuroimaging studies, while contributing greatly to our capacity to probe subjective affective states, have produced widely inconsistent findings regarding the neural substrates of emotion. The role of the amygdala in the processing of emotional stimuli, for example, has proved elusive, with early neuroimaging studies correlating amygdala activation with fear responses and other negative emotions, but underestimating the role of the amygdala in subserving positive emotions. Continuing study of the role of the amygdala in emotion paradigms has
increasingly made clear that the amygdala also activates in response to appetitive stimuli (Baxter & Murray, 2002). Some investigators speculate that early studies mistakenly associated amygdala activity exclusively with negative emotions because negative emotions are relatively easier to elicit in experimental settings than are positive emotions (Davis & Whalen, 2001). Aversive stimuli likely have a greater biological relevance than appetitive stimuli, and they are thus more easily stimulated in a laboratory setting using standardized affective probes.

The circumplex model suggests that evidence of amygdala activity during both aversive and appetitive stimuli can be understood readily if the amygdala is a part of the arousal system of the CNS. Neuroimaging paradigms exploring olfactory and gustatory stimuli support this view by demonstrating that amygdala activity increases with arousal, regardless of the particular positive or negative valence associated with experience of the stimuli (Anderson et al., 2003; Small et al., 2003). This putative role of the amygdala in emotional arousal can be confounded, however, if the aversive stimuli used in affective paradigms are more arousing than the appetitive stimuli. Indeed, more robust activity in the amygdala often accompanies the presentation of aversive stimuli that are more arousing than the corresponding appetitive stimuli. To assess adequately the role of activity in the amygdala in emotional processing, affective probes must be controlled accurately along the arousal and valence dimensions, allowing for a proper comparison of aroused positive and aroused negative affective states.

Additional difficulties in interpreting the results from neuroimaging studies of affect and emotion may stem from the subtraction method employed in most functional imaging studies, in which the neural activity associated with a cognitive process is explored by employing two tasks that differ only in one or a small number of cognitive processes of interest. This approach is problematic for affective studies because of the difficulty in generating adequate control stimuli. In a traditional subtraction study, for example, fear-inducing stimuli, such as the viewing of snakes or a loaded gun, are often compared with “neutral” or nonemotional stimuli, such as household objects. When comparing brain activity induced by these two conditions, basic emotion theory suggests that the resulting activations will relate primarily to the neural pathways that subserve fear. The circumplex model, in contrast, suggests that both stimuli will activate valence and arousal systems. Viewing a snake might induce fear, which is a negatively valenced, highly aroused state; viewing a household object, on the other hand, might induce (among other possibilities) boredom, which is a negatively valenced, low arousal state.

Following subtraction of the gray scale intensities of the images acquired during each condition, activation of the negative valence systems in both conditions will theoretically cancel one another either partially or completely. The resulting activation that is measured following the subtraction procedure will then index only differential activity across the two conditions in the arousal system, or some unpredictable combination of activity in the arousal and valence systems, but not specifically activity associated only with “fear” as a reified emotional entity.

Of particular note is that neither the interpretation of the findings based on the theory of basic emotions, nor the interpretation based on the affective circumplex, can be falsified within this type of experimental paradigm. A parametric design for these functional imaging studies, on the other hand, would allow investigators to compare affective stimuli over multiple levels or through incremental changes along each of the dimensions of the affective circumplex. For example, parametric manipulation of affective probes that vary incrementally in the degree of arousal that they induce can be mapped against accompanying changes in neural activity, thereby allowing an assessment of the activity in neural structures or pathways that correlate with the degree of emotional arousal that the stimuli produce in individual subjects.

**Emotional temperament**

Temperament research posits the existence of stable psychological profiles that begin in infancy and that persist throughout adulthood. Affective reactivity is the psychological
characteristic that has been most intensively studied in relation to temperament (Kagan, 2003), and this body of research has pointed to negative and positive affectivity as stable and heritable temperamental traits (Clark, Watson, & Mineka, 1994; Watson & Clark, 1992). Across a broad range of settings, children with proclivities to experiencing negative affective states exhibit more frequent negative emotions, such as fear, anger, and sadness, and consequently they are at increased risk for developing both depression and anxiety disorders (Clark et al., 1994). Conversely, children with positive affective tendencies are known to exhibit more frequent positive emotions, such as happiness, pride, and excitement, and are at lower risk for developing most anxiety disorders (Clark et al., 1994). Temperamental predispositions similar to positive and negative affectivity have been described in behaviorally inhibited children (Kagan, 2003) and in nonhuman primates (Kalin, Larson, Shelton, & Davidson, 1998; Kalin & Shelton, 2003). Likewise, longitudinal studies suggest that behaviorally inhibited children are more prone to developing mood disorders than are noninhibited children (Mick & Telch, 1998; Rosenbaum, Biederman, Gersten, Hirshfeld, Meminger, Herman, Kagan, Reznick, & Snidman, 1988). These findings imply that temperamental predispositions correlate not with specific, discrete emotions, but rather with broad domains of emotionality; in particular, they correlate with individual predispositions to coloring of their experiences with a particular emotional valence. Children with temperamental predispositions toward negative affectivity are not at increased risk for developing a single, specific affective illness, as would be predicted by basic emotion theory, but rather they are at increased risk for developing a range of affective disorders that have in common the proclivity to be troubled by negatively valenced emotions and experiences.

The circumplex model proposes that temperamental differences in positive and negative affectivity likely arise from variations in the bias and reactivity of the valence system within the CNS. Negative affectivity likely reflects a predisposition to imbuing experiences with negatively valenced emotions, suggesting that the set point of the valence system may be displaced from the origin of the circumplex, or that this putative neurophysiological system may be overly reactive to aversive stimuli. Positive affectivity, on the other hand, likely reflects a predisposition to imbuing experiences with positively valenced emotions, possibly because the origin of the circumplex is positively displaced along the valence axis. Indeed, temperamental predispositions toward positive and negative affectivity are associated with differences in prefrontal asymmetry, with behaviorally inhibited children demonstrating exaggerated activity in the right prefrontal cortex relative to uninhibited children (Davidson & Fox, 1989). These electrophysiologic signatures are stable over time, suggesting the existence of a neural basis for positive or negative affective temperaments.

Genetics and the affective circumplex

To date, efforts to identify genes that predispose to affective disorders have been unsuccessful, possibly because the genetic approaches and the assumptions underlying them thus far have been invalid. The vast majority of genetic studies of these conditions, whether they are family linkage, sib-pair, or genetic association studies, have assumed that affective and anxiety illnesses are discrete, readily classified, diagnostic entities. If the associations among these illnesses that the circumplex model postulates are correct, then the illnesses, in fact, are not at all discrete; rather, they lie on a neurophysiological continuum and are a consequence of disturbances in the underlying dimensions of valence, arousal, or both. In that case, statistical tests of linkage or association are inherently flawed and cannot possibly detect valid genetic determinants with any degree of reliability.

Rather than continuing with genetic studies using the traditional, categorical diagnoses of affective disorders, the genetic basis for these affective illnesses should be explored using less traditional dimensional measures of affective experience, such as those that constitute the
dimensions of the affective circumplex. The associations of genetic markers with measures of reactivity of the valence or arousal systems, using either behavioral or neurophysiological measures of these systems in the CNS, may prove a more fruitful line of inquiry into the study of the molecular genetic basis of emotional illnesses. The negative affectivity associated with depressive and anxiety disorders, for example, can be assessed using existing physiological measures, and these individual differences in negative affectivity can be correlated with variability in quantitative trait loci within the genome. Use of dimensional measures has been shown to be more powerful statistically than the use of categorical variables in identifying of loci of genetic susceptibility for complex disorders (Almasy & Blangero, 1998; Amos, 1994). This dimensional approach should therefore shed light on the genetic basis for individual differences in vulnerabilities to affective illness.

Developmental correlates of the circumplex

Behavioral studies of the affective circumplex have demonstrated that concepts and experiences of emotion differ across ages and developmental stages (Bullock & Russell, 1984). Whereas adults are able to distinguish clearly across a wide range of emotions, children seem to group emotions under more general evaluative labels (Widen & Russell, 2003). Children are capable of discerning positive and negative emotions, but within these two broad categories, they seem limited in their capacity to discern subtle distinctions among classifications of their affective experiences (Harris, 2000; Wellman, Harris, Banerjee, & Sinclair, 1995). Children tend to equate the facial expressions of negative emotions such as anger, fear, and disgust, for example, whereas adults generally recognize differences between these affective states (Bullock & Russell, 1984). The affective circumplex that reflects the conceptual structure of emotions in children lacks the differentiation evident in the circumplex from adults. This is graphically represented through multidimensional scaling procedures of similarity ratings from children of common affective facial expressions (Figure 5; Russell & Bullock, 1985). These findings suggest that although adults are fully able to differentiate affective states in both the valence and arousal dimensions, the circumplex in children is poorly differentiated, with affective states grouped primarily according to extremes of valence and with little discrimination on the arousal dimension.

Clinically, this paucity of differentiation in the affective circumplex is evidenced in the limited capacity of children to express their own affective states. Children tend to describe their emotions solely in terms of valence ("I feel bad" or "I feel good"), lacking the subtleties evident in adult descriptions ("I feel excited" or "I feel content"; Wellman et al., 1995). As children mature, their capacity to conceptualize and express their own affective states becomes increasingly sophisticated (Saarni, 1999; Sroufe, 1979) and they are able to provide increasing levels of nuance and shading to descriptions of their emotional experience. They thereby become progressively better at differentiating the affective circumplex in each of its four quadrants. What is originally described as feeling "bad," "good," "excited," or "bored" may now be expressed, for example, as "dejected," "serene," "ecstatic," or "ennui," suggesting not that the underlying neurophysiology has necessarily changed, but rather that their conceptualization and interpretation of the core physiological experiences that underlie the circumplex is more sophisticated, resulting in more varied reports of emotional experience.

This developmental trajectory of emotional experience may be useful in understanding a common finding in epidemiological studies of mood disorders in children. Often the first sign of psychological distress in children at risk for mood disorders is the development of anxiety (Warner, Weissman, Mufson, & Wickramaratne, 1999). These symptoms may be reported either as feeling bad or worried, or they may be reported as fearfulness in the form of simple phobias. The presence of such anxiety symptoms predicts the subsequent development of other psychiatric illnesses, particularly depression, in later life (Breslau, Schultz, & Peterson,
Developmental studies of the affective circumplex suggest that these findings may in part pertain to differentiation and elaboration of emotional concepts. A child's report of feeling worried, for example, may be indexing a general negative valencing of emotional life, with or without adequately reflecting fluctuations in emotional arousal. The child or the child's parents may interpret this preponderance of negative valence as worry or anxiety. With cognitive maturation, however, this enduring negative valencing of affective experiences is interpreted with increasing nuance, allowing for more varied and subtle descriptions such as feelings of “despair,” “dysphoria,” or “hopelessness.” In effect, the core neurophysiological disturbances have not significantly changed, but the child's capacity to conceptualize and express these affective processes has matured, leading to more varied psychiatric symptoms and a divergence into one or more diagnoses.

Conclusions

We have argued for a conceptual shift in the theory and experimental approaches to the study of emotion and affective illness. We have suggested that the dimensional model of the affective circumplex helps explain current research and clinical findings that are at odds with models of basic emotions. The circumplex model of affect posits that the two underlying neurophysiological systems of valence and arousal subserve all affective states, and upon this substrate are layered various cognitive processes that interpret and refine emotional experience according to salient situational and historical contexts. Maturation of this cognitive overlay likely accounts for the progressive differentiation of the circumplex and the increasingly sophisticated nuanced of emotional experience across development. The circumplex model complements data from developmental, neuroimaging, and behavioral genetics studies of affective disorders. Moreover, the model addresses important lingering questions within clinical psychology, such as the high rates of comorbidity seen in various psychiatric disorders and the continued difficulty in uncovering genetic predispositions to Axis I disorders. Finally, the circumplex model offers novel experimental paradigms for pursuing future genetic, neuroimaging, and clinical research in the affective neurosciences.

Acknowledgements

This work was supported in part by NIMH Grants MH01232, MH59139, MH36197, MHK02-74677, and MH068318; a grant from the National Alliance for Research in Schizophrenia and Affective Disorders (NARSAD); NSF Grant BSC-0421702; and funding from the Thomas D. Klingenstein and Nancy D. Perlman Family Fund and the Suzanne Crosby Murphy Endowment at Columbia University.

References


Figure 1.
A graphical representation of the circumplex model of affect with the horizontal axis representing the valence dimension and the vertical axis representing the arousal or activation dimension.
Figure 2. The mesolimbic system. The mesolimbic system commences in the ventral tegmental area (VTA) with dopaminergic projections to the nucleus accumbens (NA). The NA has reciprocal connections to the amygdala (A), hippocampus (HC), caudate nucleus (C), and prefrontal cortex (PFC).
Figure 3.
The pathways of the arousal network.
Figure 4.
(a) The circumplex model with valence focus. This affective circumplex with valence focus demonstrates poor differentiation along the arousal axis. This creates overlap in similarly valenced emotions such as anxiety and sadness or happiness and contentedness. (b) The normal circumplex model without valence focus. A typical affective circumplex with similarly valenced emotions is differentiated along the arousal axis.
Figure 5.
The affective circumplex of children. Shown here is a representative circumplex that was generated from the ratings by children of the degree of similarity in facial expressions. Unlike an affective circumplex from adult ratings, various positive emotional expressions are grouped broadly into a “good emotion” category and negative expressions are grouped in a “bad emotion” category.