Distribution Agreement

Signature:

____________________________________ ______________

Christine D. Wilson-Mendenhall Date

Assessing Predictions from Psychological Construction Approaches to Emotion

By

Christine D. Wilson-Mendenhall
Doctor of Philosophy

Psychology

Lawrence W. Barsalou, Ph.D.
Advisor

Stephan Hamann, Ph.D.
Committee Member

Lynne Nygaard, Ph.D.
Committee Member

Kim Wallen, Ph.D.
Committee Member

Lisa Feldman Barrett, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Assessing Predictions from Psychological Construction Approaches to Emotion

By

Christine D. Wilson-Mendenhall
B.A., Emory University, 2003
M.A., Emory University, 2007

Advisor: Lawrence W. Barsalou, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2010

Abstract

Assessing Predictions from Psychological Construction Approaches to Emotion

By Christine D. Wilson-Mendenhall

To understand the complex mental phenomena underlying affect and emotion, researchers have begun investigating the neural systems that produce these states. In an initial manuscript, we examined the theoretical view that all emotions include two core properties (among other things): valence (state of pleasure or displeasure) and arousal (state of energy and mobilization). During a functional magnetic resonance imaging (fMRI) experiment, participants immersed themselves in imagined fear, happiness, and sadness scenarios. After the initial scenario immersion period, participants focused on and rated the valence or arousal of the felt emotion. This paradigm was designed so that neural activity during the scenario immersion event could be separated from the subsequent focus-rate event. Furthermore, the scenarios developed for each emotion category varied in valence and arousal, and also included familiar scenarios in which the valence was atypical for the category (e.g., pleasant fear of thrill-seeking) to maximize variance in valence. According to core affect views, brain regions correlating with valence or arousal ratings during the focus-rate events in our paradigm should be the same across experiences of all three emotions (fear, happiness, and sadness), and across typical vs. atypical valence. Based on previous literature, we further predicted that valence ratings would correlate with neural activity in orbitofrontal cortex and that arousal ratings would correlate with neural activity in the amygdala. The results supported our predictions, suggesting that valence and arousal are core properties of emotion experience. In a second manuscript, developed from the same experiment, we examined the emotion scenarios that had atypical valence (e.g., pleasant fear). Of primary interest was the role of large-scale brain networks that support 'default' internal socio-emotional simulation, salience detection, and task-oriented attention in processing these emotions. During scenario immersion and subsequent valence focus, more activation occurred in the default network for atypical than for typical emotion scenarios. During the subsequent valence focus period only, attention and salience networks became relatively more active for atypical emotion scenarios. Future research is necessary to understand the properties of atypical emotions (e.g., complexity, ambiguity, etc.) that underlie heightened activity in these networks when emotions are atypical.

Assessing Predictions from Psychological Construction Approaches to Emotion

By

Christine D. Wilson-Mendenhall
B.A., Emory University, 2003
M.A., Emory University, 2007

Advisor: Lawrence W. Barsalou, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2010

Table of Contents

Page

List of Tables and Figures

Page

General Introduction

The term 'emotion' has defined a research area that might broadly be described by its Latin roots as investigating what *moves* the mind and body. Speaking of emotion covers a wide range of phenomena varying in duration, intensity, and complexity (Russell, 2003). In this study, the focus will be on affective episodes. Affect as it is used here means to produce a change in the body state (Barrett & Bar, 2009; Duncan & Barrett, 2007). Awareness may not always accompany the bodily changes that characterize an affective episode. One can experience affect in many subtle forms that remain in the background of experience (Barrett & Bar, 2009; Damasio, 1994). Often, though, affect can be foregrounded when attention is focused on it, much like other aspects of experience.

A great deal of research has focused on examining the content, experience, and underlying mechanisms of typical episodes that people might categorize as happiness, sadness, fear, or anger. For example, participants will recall memories they remember as being intensely sad, such as the loss of a loved one, that may be accompanied by specific reactions like crying (Damasio et al., 2000). Many experiences, however, can be categorized as sadness. Consider the warm sadness one feels when thinking about pleasant memories. This experience is interpreted as sadness because it is conceptualized as loss, but the feeling tone of the memories is often pleasant. These types of episodes are usually not the first to come to mind when asked to identify instances of sadness. Nevertheless, most people can identify with this feeling, and would not find it strange to call it sadness. It is these types of experiences that exemplify the affective experience that accompanies everyday moments. While it is important to understand major life events that are accompanied by extreme emotion, it is also important to examine the day-to-day experience that underscores general emotional well-being.

Given the range of affective episodes or situations experienced, a central question is what genetic predispositions or knowledge is fundamental for their expression. Traditionally, the idea of natural kinds has been considered a useful construct for understanding what is primitive and given in the form of genetic predispositions. A natural kind is a grouping of instances that occur in the world and is not arbitrary (Barrett, 2006a). Natural kind categories are not constructed by the human mind; they are given by nature. In this framework, specific emotion categories are considered primitive. The terminology most frequently used to indicate category primacy is 'basic' or 'discrete.' Most theories include fear, anger, and joy as natural kinds, but the number and identity of categories differ among prominent theories (Ortony & Turner, 1990; Plutchik, 2002).

Another way to think about what is given innately is the construct of core knowledge. Core knowledge is a specialized system of knowledge that is hardwired in the brain (Spelke & Kinzler, 2007). The developmental psychology tradition of core knowledge recognizes four core systems that involve reasoning about objects, and a possible fifth system for reasoning about social partners and groups. Complex skills and abilities are thought to depend on these core foundations. Emotion theorists also use the construct of core knowledge to refer to a specialized, genetically predisposed, system that is the foundation for more complex behavior (Barrett, 2006a). However, core *affect* does not map onto proposed reasoning systems in development psychology. Instead, it posits the basic computations underlying valence (state of pleasure or displeasure) and arousal (state of energy and mobilization) as primitive abilities.

Both natural kinds and core knowledge are constructs that can be used to think about what predispositions have evolved in the brain. Importantly, these two frameworks shape research agendas in very different ways. Divergences have become even more apparent in the blooming field of social affective neuroscience. According to a natural kinds view of emotions, empirical work should focus on identifying diagnostic patterns of brain activity corresponding to each *discrete* emotion category. Alternatively, approaches in which core affect is primary suggest that an empirical focus should be on grounding mechanisms underlying valence and arousal in the brain. Because these approaches differ dramatically, they influence how neuroimaging studies are designed, analyzed, and interpreted. They also shape how brain function is viewed more generally, for example, by identifying brain regions as emotional or cognitive (cf. Duncan & Barrett, 2007; Pessoa, 2008).

The purpose of this study is to examine the hypothesis that valence and arousal are core knowledge, which is grounded in the brain. The valence and arousal dimensions of affective episodes will be manipulated such that they systematically vary within several discrete emotion categories. If core affect is a primitive process, then the same neural networks should support processing of valence and arousal regardless of emotion category. To our knowledge, this is the first study that examines these dimensions as *core* neural processes in emotion categorization.

Traditional Approaches: Discrete Emotions as Natural Kinds

The foundation of most discrete emotion views is evolutionary function, and many theorists credit Darwin as their inspiration (e.g., Ekman, 2003). Built-in evolutionary response systems provide an explanation for the classic problem of how a stimulus produces the physiological and behavioral responses that characterize emotional experience (LeDoux, 1996). For example, a threatening bear would trigger the evolved fear system. The fear system interprets the bear as 'dangerous' and activates the autonomic nervous system to facilitate escape. In this view, specific emotions systems evolved to deal with important problems related to survival, such as escaping predators. Behavior is considered adaptive if it promotes individual fitness or species fitness (e.g., prososcial activities; Buck, 1999).

Evolved discrete emotion systems are thought to interpret an ecologically valid stimulus quickly in an instinct-like process. The interpretation of 'danger' in the fear system, for example, is considered a process that occurs unconsciously. The term autoappraisal is often used to convey the automaticity of the process (Ekman, 2003). This autoappraisal generates a distinct set of neural/physiological, expressive, and feeling/motivational components (Izard, 2007). The end result is that the agent is prepared to act quickly in the situation and convey communicative signals. The term 'affect program' is sometimes used to describe this bottom-up process, which is thought to operate like a Fodorian encapsulated module at times (Cosmides & Tooby, 2000; Ekman, 2003).

Discrete emotion systems are each thought to serve different adaptive functions. Thus, they are triggered by different environmental stimuli, have different vocal and facial signals, and exhibit different neural and physiological profiles (Ekman, 1992). What follows is that dissociable patterns should exist in facial musculature, physiological activity, and brain states that correspond to these natural kinds. This question has been pursued in a number of ways (for review see Barrett, 2006a). One way has been to examine whether similar patterns for discrete emotions exist in multiple cultures. The assumption is that patterns that exist across cultures suggest an innate capacity for the emotions. In this vein, much early work focused on whether people could categorize posed facial expressions typical of emotion categories (for review see Ekman, 1994). As technology has become more sophisticated, physiological and neural activity has been measured to assess biological patterns within and between cultures.

Not all approaches to discrete emotions have a biological focus. A distinction has been made between the biological approach, and what has been called the psychological approach (cf. Ortony & Turner, 1990). The psychological approach also suggests that a basic set of discrete emotions exist and serve important functions. Instead of focusing on biological patterns, however, these theorists have focused on identifying situational patterns of experience in verbal reports (e.g., antecedent events, action tendencies, behavioral consequences) that correspond to discrete emotions (e.g., Mesquita & Frijda, 1992). Another divergence is that the psychological camp is more likely to support the idea that elements of affect programs are specified innately as opposed to the entire system (e.g., goal-directed action; cf. Ekman, 2003). The subjective report method assumes that insights into the way discrete emotion systems work can be made through conscious assessment of emotion states. Those in the biological tradition often disagree with this statement, suggesting that affect programs operate too rapidly to be available for conscious processing. Despite these criticisms, the psychological approach has provided descriptive information that is critical to any discrete emotion view, such as commonly experienced antecedent events thought to elicit discrete emotion states. Interestingly, antecedent events have been considered a distinguishing characteristic of discrete emotion states, but are not well defined in biological approaches (Ekman, 2003).

With the development of neuroimaging techniques, it is now possible to investigate the neural bases of emotional states. These methodologies have provided another way to test biological patterns. Most discrete emotion views predict that specific brain networks should be involved in discrete emotion states (Ekman, 2003; Izard, 2007; Panksepp, 2000). Very few theorists, however, have working theories that lay out the neurological profiles of different discrete emotions. The most developed theoretical perspective has emerged from animal work (Panksepp, 1998; 2000; 2005).

Panksepp (2000) specified a neural taxonomy that has three levels. The first level contains the 'reflexive affects,' including the startle reflex, gustatory disgust, pain and homeostatic distresses like hunger, and sensory pleasures. The term affect has been defined in this theory as a subjective experiential-feeling component that is difficult to express verbally (Panksepp, 2005). Reflexive affects are primarily represented in the brainstem. The second level contains the emotional affects, which are true 'blue-ribbon, grade-A' emotions. These affects map onto some of the discrete emotions proposed by other theorists. Emotion affects are considered sensory-motor emotional command circuits in the brain that coordinate behavioral, physiological, cognitive, and affective consequences. In this view, it is important to distinguish the emotional affects from the sensory pleasures and displeasures in the first level. Circuits for the emotional affects were originally characterized as involving 'intermediate' brain regions, including both higher limbic[1] regions in frontal, cingulate, and temporal cortices, and midbrain regions like the periaqueductal gray (PAG). In recent discussions, the subcortical circuits have been emphasized as the core of these systems (Panksepp, 2007). The final level in the taxonomy is called the 'higher sentiments' and includes social emotions such as shame, guilt, contempt, envy, humor, empathy, sympathy, and some forms of jealousy. Higher sentiments involve the forebrain only or mixtures with lower affects represented in subcortical regions.

Research on this theory has focused primarily on the emotional affects, which are considered natural kinds that have specifiable neural substrates. Seven emotional affect systems have been proposed to exist in mammalian brains, which correspond to the following experiential-feelings states: SEEKING, RAGE, FEAR, LUST, CARE, PANIC, and PLAY. These systems are based on specific behaviors that can be observed in animals. RAGE and FEAR have been linked to the discrete emotions of anger and fear that have been proposed by other basic emotion researchers. Correspondences with other discrete emotions discussed in the human

---

[1] Some researchers have argued that the term limbic should no longer be used because the criteria that defined the system are no longer considered valid (LeDoux, 1996). Others suggest that the term is still useful as a way to discuss families of regions that appear to be involved in emotion (Panksepp, 2000).

literature are less clear. PANIC has been proposed to map onto the separation distress that characterizes sadness, and rat 'laughter' vocalizations associated with PLAY have been compared to human social joy (Panksepp, 2005). In general, though, discrepancies with other discrete emotion theories are not considered problematic. These seven systems are considered to be the true natural kinds in animals and humans because they have identifiable neural correlates.

The emotional affect systems have been primarily been studied in rats. The main technique used is localized electrical and chemical stimulation to a brain site (Panksepp, 1998, 2007). Conclusions have been drawn about animals' emotional states on the basis of resulting behavior. For example, the core components of the RAGE system are thought to be hierarchically organized, starting with medial amygdala that projects to the medial hypothalamus, which then projects to midbrain PAG sites (Panksepp, 1998). Higher regions are dependent on lower regions for *expression* of the RAGE (i.e., lesions of the PAG diminish aggressive behavior elicited from the amygdala or hypothalamus but not vice versa). Aggression is the expressive behavior studied for RAGE, often in cats. Rage-like aggression is characterized by piloerection, autonomic arousal, hissing, and growling during attack, and is distinguished from predatory aggression, which is characterized by methodical stalking and well-directed pouncing. Similar distinctions have been made in rats. Whereas predatory aggression is evoked from stimulation of the dorsolateral hypothalamus, affective aggression is evoked from stimulation of the ventrolateral and medial hypothalamus. This neural distinction has been interpreted as support for a separate neural circuit dedicated to the emotion RAGE.

Importantly, the subcortical structures supporting RAGE aggression in cats and rats are presumed to be homologous in function to those in humans. It is strongly advocated that these animal models guide the study of human emotion.

Examining the Neural Evidence for Discrete Emotions

Neural evidence from human neuroimaging studies has recently taken a central role in evaluating the discrete emotion view. It should not be forgotten, however, that neuroimaging was preceded by many years of research on facial expression and physiology. The available evidence for discrete emotions has been critiqued multiple times over the past few decades (for reviews see Ortony & Turner, 1990; Russell, 1994; Barrett, 2006a). These reviews developed conceptual points that apply to all forms of evidence. Thus, it is important to preface discussion of current thinking in neuroscience with these conceptual foundations.

First, the emotions proposed as natural kinds are not well agreed upon among different theorists (Ortony & Turner, 1990). Most theories include fear, anger, and joy as natural kinds (Plutchik, 2002). Examples of less consistent categories include interest, shame, contempt, and wonder. In some cases, the categories proposed as natural kinds appear to map onto the methodologies being used. For example, in many early theories, an emotion was only basic if there was a corresponding facial expression that could be assessed across cultures. More recently, it has been suggested that specific forms of neural evidence offer the appropriate criteria for defining emotions as natural kinds (Panksepp, 2000).

Another issue that has been raised repeatedly involves interpreting findings from animal research. Research on emotional experience in human participants is bound more strictly by ethical standards than research conducted with animals. Invasive neural methodologies cannot be used with human participants unless part of a necessary medical procedure. Therefore, there is existing data from animal studies without comparable empirical results in human populations. Researchers who study emotions in animals are thus faced with an important problem when attempting to generalize to humans—their subjects cannot subjectively report their affective states. Thus, affective states are inferred and generalized on the basis of behavior, as described above in the characterization of Panksepp's RAGE system (Panksepp, 1998).

Grounding discrete emotion theory in animal models presents a related problem (cf. Ortony & Turner, 1990). Some proposed discrete emotions are presumed easier to study in animals because of stereotyped behavior patterns (e.g., freezing in fear; aggression in anger). However, other proposed basic emotions like shame or contempt are difficult to assess in animals. If the presence of affect programs in other animals is fundamental to theories of discrete emotions, then evidence from animal models suggest that some emotions proposed by classic theories (e.g., Ekman, 1992) are more basic than others. The idea that a continuum of 'basic-ness' exists calls into question the construct of a natural kind. However, this issue is not a problem for researchers who consider specific response patterns that can be observed in animals to be the criterion for discrete emotions states (e.g., Panksepp, 2000). Interestingly, these are the only theories that currently specify the neural bases of discrete emotion states in any detail.

Fueling this theoretical discussion is accumulating evidence that many studies do not find distinct patterns in the neural, muscular, or autonomic activity for discrete emotion states. Barrett (2006a) reviewed evidence that did not support the discrete emotion view across a number of domains, including subjective experience, facial and vocal signals, peripheral nervous system responses, voluntary behaviors, and neural circuitry. In this review, the two available meta-analyses of neuroimaging studies were examined. A third meta-analysis focusing on discrete emotions has since been published, which is included in the discussion here. All three reviews have used different meta-analytic techniques.

In the first neuroimaging meta-analysis, Phan, Wager, Taylor, and Liberzon (2002) examined 55 publications that investigated the mental processes in emotion, excluding low order sensory and motor tasks framed as emotional processing. They tabulated the frequency of studies that reported activation during tasks involving the emotion categories fear, sadness, disgust, anger, and happiness for a brain regions falling under the same label (e.g., amygdala). This technique is often referred to as the label-based method. They then computed chi-square tests to determine the relative contribution of a brain area to one emotion category above all others. Chi-square tests were significant for two emotion categories. A significant number of studies activated the amgydala during fear processing (60% of studies). The other significant results indicated that the subcallosal cingulate cortex was implicated in sadness (40% of studies). Seventy percent of studies reported activation in the basal ganglia for happiness, but the frequency was not distinguished from disgust (60% of studies). No significant patterns were observed for anger.

Murphy, Nimmo-Smith, and Lawrence (2003) examined 106 neuroimaging studies of emotion, which were defined as paradigms that use valenced stimuli or that assessed approach-avoidance tendencies. Studies were only included if they contained a high-level neutral control condition. The 3-D Kolmogorov-Smirnov statistic was used to compare the rough spatial distributions of activation patterns associated with the same emotion categories included by Phan et al. They found that the distribution for disgust differed significantly from all other emotion categories. The distribution for fear differed significantly from disgust, anger, and sadness, but not from happiness. The same profile was shown for sadness and anger, which differed from all other categories except happiness. To follow up this analysis, the most consistently activated region for each emotion category was reported. These regions were not necessarily the regions driving differences in spatial distributions. The amygdala was consistently activated in fear states, the insula/operculum and globus pallidus in disgust, the lateral orbitofrontal cortex in anger, the rostral supracallosal anterior cingulate/dorsomedial prefrontal cortex in happiness, and rostral supracallosal anterior cingulate/dorsal medial prefrontal cortex in sadness. These results are consistent with the results of Phan et al. for the categories of fear and sadness. Both reviews did not find a discriminable pattern for happiness.

Vytal and Hamann (2010) examined 85 neuroimaging studies using the same inclusion criteria as Murphy et al. They improved the spatial resolution by using a recently developed meta-analytic technique and increased power by including more (recently published) studies. The Activation Likelihood Estimation (ALE) technique allows one to compute a statistical map by modeling activation coordinates from different studies as the centers of Gaussian probability distributions and then computing voxel-wise probabilities of activation across studies (Laird et al., 2005). ALE maps were computed for each of the five discrete emotion categories and then pair-wise contrasts were examined by comparing the difference with a null distribution generated by random permutations. Discriminable patterns of activation were found for each emotion. Similar to the other meta-analyses, fear consistently activated the amygdala, which discriminated it from all other emotions except anger. In addition, the insula discriminated fear from all other emotions[2]. Also consistent with the other reviews, sadness consistently activated the subgenual anterior cingulate as well as the head of the caudate and middle frontal gyrus, both new findings. Murphy et al.'s finding that anger activated lateral orbitofrontal cortex[3] was also replicated, and in this review, anger was also discriminable in parahippocampal cortex. For the first time, happiness was discriminated from other emotions by rostral anterior cingulate and right superior temporal gyrus.

As meta-analytic methods have become more sophisticated and the number of neuroimaging studies has increased, stable patterns have been detected in the brain that discriminate between a subset of emotion categories. Naturally, these findings have been interpreted as evidence for discrete emotion theories. When

---

[2] Discrimination from disgust was limited to the posterior insula.

[3] This region was labeled as inferior frontal gyrus in the article. The coordinates and brodmann area given (47) suggest it is in the oribital sector of this gyrus, which some researchers call caudolateral or lateral orbitofrontal cortex.

considering what these patterns mean, however, many questions arise. Classic discrete emotions theories suggest that the auto-appraisal process involves interpretation of an antecedent event to produce a coordinated set of physiological and behavioral responses. A central question, then, is what functional roles the brain regions identified in these meta-analyses are playing in the appraisal process for each emotion? This question is further complicated by the fact that many of the regions reported (e.g., parahippocampal cortex) have been implicated in other cognitive neuroscience literatures including memory, learning, and pain. It seems possible, then, that the differences captured could reflect situational differences that are not core aspects of discrete emotion states. While this knowledge is still informative of the neural systems involved in conceptualizing emotional experience, it may not be best characterized as the functional correlates of the affect programs proposed in many discrete emotion theories.

A related issue is whether the emotional states studied in neuroimaging paradigms are truly the discrete emotion states that Ekman and others have proposed. A large proportion of the studies examined in the two most recent meta-analyses have used facial expressions as stimuli (Vytal & Hamann, 2010). Physiology researchers have emphasized the difficulty of eliciting discrete emotion states in the lab (Levenson, 2003). In the trade-off between control and ecological validity, facial expressions have been suggested to be high in control, but low in validity (i.e., tasks most similar to situations in which emotions are experienced in naturalistic environments). Other paradigms are considered more balanced in this regard, such reliving emotional states through autobiographical memories (e.g., Damasio et al., 2000).

Finally, many of the regions detected in the meta-analytic reviews are not consistent with the subcortical systems emphasized in animal models. The main exception to this statement is the consistent finding of amygdala activation for the category of fear, which is consistent with much animal work on fear conditioning (LeDoux, 1996). Some animal researchers dismiss fMRI findings due to the correlational nature of the BOLD signal (Panksepp, 2007). The human evidence carrying the most weight in these views are case studies reporting human subjective experiences and behavior during brain stimulation that was necessary for a medical procedure. Generalization of case studies is of course limited. Furthermore, the behavior and subjective experience following brain stimulation is known to vary greatly between different individuals (cf. Barrett et al., 2007).

Should the neural circuits identified as natural kinds in animals constitute the only ‘true’ neural evidence for mammalian discrete emotions? As mentioned previously, animal models are dependent on inferences from behavior. To investigate natural kind emotion categories, a behavior must be indentified that maps onto the proposed affective state. The periaqueductal gray (PAG) has been heavily emphasized in this theoretical view because evidence suggests that it supports coordinated skeletal, autonomic, and antinociceptive response profiles that have been described as fight or flight (Bandler & Shipley, 1994). A problem with this logic, however, is that animal behavior has been shown to be contextually sensitive to the situation (cf. Barrett, 2006a). For example, a predator situation, which would presumably activate the 'fear' system, could result in vigilance, freezing, attack, or flight depending on the situational context. An arguably bigger problem is the assumption that the proposed subcortical structures function identically in rats and humans. For example, the orbitofrontal cortex, which has direct, reciprocal connections with the PAG in primates, is hard to compare to the rodent orbitofrontal cortex (An, Bandler, Ongur, & Price, 1998; Kringelbach & Rolls, 2004). Even the taste system is connected differently in rodents and primates.

In summary, uncertainty still exists concerning which discrete emotions are primitive or basic and how corresponding affect programs are represented in the human brain. Given the lack of conclusive evidence for discrete affect programs, examining alternative proposals for the foundation of emotional experience seems warranted.

An Alternative Approach: Core Affect

An alternative hypothesis is that emotion categories can be deconstructed into more primitive constructs. In this view, core affect is proposed to be the primitive system or genetic predisposition upon which complex behavior is built. Core affect is a state of pleasure or displeasure with some level of arousal (Barrett & Bliss-Moreau, 2009; Russell, 2003; Russell & Barrett, 1999). It is a continuous and fluctuating state, which is not dependent on an object. In this way it is analogous, for example, to felt temperature (Russell, 2003). Its descriptors, valence and arousal, are orthogonal dimensions that create a space often described as a circumplex when prototypical emotion categories are mapped within it (Russell, 1980). More specifically, valence is the aspect of a core affect state that is experienced as ranging from very pleasant to very unpleasant, whereas arousal is the aspect of core affect that reflects a sense of mobilization, experienced as ranging from activated to still. Importantly, the experience of arousal as a degree of activation or deactivation in mind and body is not simply reducible to measures of bodily physiological activity (Barrett, Mesquita, Ochsner, & Gross, 2007).

Although the discrete emotion view is currently the dominant approach to studying emotion, the idea that core affect is fundamental to emotion experience is not new. There is a rich history of thinking about emotional experience in terms of dimensions such as valence and arousal, which includes such historical figures as Wilhelm Wundt (Russell, 2003; Barrett & Bliss-Moreau, 2009). Furthermore, theories of core affect can explain many affective phenomena (Russell, 2003). For example, mood can easily be understood as a prolonged core affect state with no immediate object.

Evidence for the core affect view has primarily been established using subjective reports of emotional experience. When data reduction techniques (e.g., factor analysis) are applied to ratings of subjective experience, resulting components reflect valence and arousal dimensions, not discrete emotion categories (for reviews see Barrett and Russell, 1999; Russell, 2003; Barrett, 2006b; Barrett & Bliss-Moreau, 2009). Such evidence suggests that these dimensions are *experientially primitive*. As mentioned previously, the use of subjective report has been criticized as a method for studying the mechanisms underlying emotional experience (e.g., LeDoux, 1996). Barrett (2006a) lays out several reasons why subject reports should not be dismissed. First, self-reports do not simply reflect the vocabulary available to express emotions, as shown by weak correlations between an individual's vocabulary structure of emotion words and verbal reporting of emotional experience. Perhaps more importantly, experiential reports have recently been shown to correlate with elements of sensory-motor behavior, such as perceptual sensitivity.

One criticism of core affect is that its two dimensions, valence and arousal, cannot capture the richness of emotional experiences (e.g., Fontaine, Scherer, Roesch, & Ellsworth, 2004; Panksepp, 2007). A common example of this argument is that prominent discrete emotions cannot be discriminated within the basic core affect framework. Both anger and fear, for example, are typically highly arousing, highly unpleasant states, which a simple dimensional model would not clearly discriminate. However, constructivist views, such as the conceptual-act model, suggest that core affect interacts with other neural systems to construct emotional experience (Barrett, 2006a; Lindquist & Barrett, 2008). Core affect is primitive and necessary for emotion, but it alone does not constitute emotion. The other major contributor to emotion is the human conceptual system, which forms concepts of external objects and settings as well as interoceptive concepts of mind and body to achieve goals (Barsalou, 1999; Barsalou, 2003; Barsalou, 2008).

Conceptual knowledge of agents, events, actions, relations, settings, and so on is used to constantly interpret present experience. Core affect, then, can be considered an element of a mental state, which would also contain many other forms of knowledge. Thus, additional knowledge of objects, relations, agents, and events would easily distinguish the categories fear and anger. Furthermore, it would not be surprising if other dimensions suggested to play a role in emotion experience (e.g., situational components like control, power, norms and value, goals) could be grounded in fundamental domains of conceptual knowledge.

Neurobiological Grounding of Core Affect

It has been proposed that core affect is *experientially* primitive. Is core affect, however, *biologically* primitive? If genetic predispositions exist for core affect, then a dedicated neural system should underlie its dimensional computations. For this reason, researchers are beginning to investigate whether neural systems map onto the subjective experience of valence and arousal.

One proposal is that the neural system supporting core affect binds external information interpreted by the sensory modalities with information arising from somatovisceral and homeostatic states of the body to represent the value of current experience (Barrett & Bliss-Moreau, 2009; Barrett & Bar, 2009). This system has recently been conceptualized as two interacting networks, largely based on patterns of connections that distinguish lateral and medial orbitofrontal cortex (OFC; Ongur & Price, 2000). The lateral network is considered a sensory integration circuit that produces and stores the value representation of the experience. It includes the central and lateral OFC, adjacent agranular insula, and the basolateral complex of the amygdala. The medial network is thought to guide autonomic, endocrine, and behavioral responses to an object or stimulus. This network includes medial OFC (extending into what is considered ventromedial prefrontal cortex), subgenual and pregenual anterior cingulate, amygdala, ventral striatum, hypothalamus, and regions in the midbrain and brainstem. Dynamic activation patterns across these networks are thought to produce the emergent phenomena of core affect.

These regions were all identified in a recent meta-analysis that examined patterns of co-occuring activations in 162 neuroimaging studies of emotion (Kober et al., 2008). Unlike previous reviews, this work did not attempt to differentiate activation patterns corresponding to proposed discrete emotion categories. Instead, it focused on identifying functional groups of areas that presumably interact to produce the experience of emotion. The brain regions in the proposed core affect system largely correspond to three interconnected functional groups identified in the meta-analysis.[4] Two of these functional groups, the 'medial PFC' and 'core limbic' groups, appear to map onto the medial network described above whereas as the 'lateral paralimbic' group corresponds more to the lateral network. Thus, the proposed core affect system clearly seems to be involved in emotional experience. The next step involves determining if these regions play a role in processing the arousal and valence dimensions central to core affect states.

Initial investigations of core affect dimensions have largely focused on the affective experience of basic sensations, particularly in the chemosensory domains of taste and smell. Because food is a natural elicitor of hedonic experience, it has been suggested that the pleasure felt during social interactions evolved from basic sensory pleasures (Berridge & Kringelback, 2008; Kringelback & Berridge, 2010). Furthermore, from an empirical standpoint, chemosensory stimuli are ideal because their chemical make-up can be systematically altered to manipulate valence and

---

[4] Six total functional groups were identified in the analysis.

intensity in a rigorous manner. Patterns emerging in this literature suggest a dissociation between coding of intensity (arousal) in the amygdala and valence in the orbitofrontal cortex (OFC).

In one of the first studies on this topic, Anderson et al. (2003) examined the affective properties of olfaction. They altered the intensity of a pleasant (citral) and unpleasant (valeric) acid to produce odors in the four quadrants of valence-arousal affective space (i.e., pleasant-high, pleasant-low, unpleasant-high, unpleasant-low). Participants were scanned while detecting whether an odor was present or not. The analysis strategy was to draw regions of interest on each participant's neuroanatomy in the amygdala and OFC. The results showed that high intensity odors activated the amygdala more than low intensity odors irrespective of valence category. Subjective ratings of intensity collected after scanning also correlated with activity in the right amygdala, extending into piriform cortex (primary olfactory cortex). A region in right medial OFC was more active for pleasant odors than unpleasant odors regardless of intensity, and a nearby anterior medial OFC region was positively correlated with valence ratings (i.e., more activity associated with greater pleasantness). In left lateral OFC, an interaction was found in which greater activation was observed for unpleasant relative to pleasant odors, but only at high intensities. A more anterior region in left lateral OFC was negatively correlated with valence (i.e., more activity associated with greater unpleasantness). Taken together, the findings were interpreted as evidence that intensity is encoded in lower-order primary olfactory cortex, and higher-order properties such as valence are represented in secondary olfactory cortex within OFC.

The valence profile observed above was replicated in a study that investigated multiple pleasant and unpleasant odors matched in intensity, including the medial-lateral distinctions in OFC (Rolls, Kringelbach, & de Araujo, 2003). Activity in these regions was also correlated with subjective valence judgments in the predicted directions. Interestingly, a distinct anterior region in left OFC also emerged in the negative valence correlations assessing unpleasantness in this study. Medial and lateral OFC activation is often interpreted in the context of reward theories. Lateral OFC has been proposed to map onto punishment and subsequent behavior change, whereas medial OFC has been proposed to map onto decoding and monitoring of reward (for reviews see Kringelback & Rolls, 2004; Kringelbach, 2005).

A number of chemosensory studies, however, have not observed a medial-lateral OFC valence profile. Small et al. (2003) manipulated arousal and intensity using different concentrations of sweet and bitter solutions. Based on each participant's individual ratings, solutions were formed that fell into four quadrants of gustatory affective space. Unlike previous studies, Small et al. found increased activation in right caudolateral OFC for pleasant tastes irrespective of intensity. A more anterior left OFC region was active for both high and low intensity unpleasant tastes relative to a tasteless solution. This region may correspond to the area that emerged in the correlational analyses in prior studies. The main effect of intensity in the amygdala was replicated, with whole-brain analyses further revealing the same effect in the mid-insula, cerebellum, and pons. These results suggest that the intensity effect may not be a function of early sensory processing, given that primary taste cortex is located in the insula/operculum. Nevertheless, taste and smell appear tightly coupled.

A clever study by Winston et al. (2005) further defined the role of the amydala in coding intensity. These researchers created high and low intensity concentrations of pleasant, neutral, and unpleasant odors. Similar to previous studies, the amgydala was more active for intense variations of both the pleasant and unpleasant smells. The amygdala, however, did not distinguish between the high and low intensities of neutral odors. Furthermore, the pleasant and unpleasant high intensity odors showed greater activity than the high intensity neutral odor. The term 'emotional salience' has been used to describe this pattern of results, which involves an interaction between valence and arousal.

In all the studies discussed thus far, the properties of the critical stimuli have been manipulated to modulate valence and intensity. Some studies, however, have used context to produce changes in subjective pleasure. For example, hungry participants have been scanned before and after being fed a liquid food stimulus until satiated (Kringelbach, O'Doherty, Rolls, & Andrews, 2003). After eating until satiated, participants did not find the food as pleasant as before eating it. An advantage of this design is that the stimulus properties do not change, including perceived intensity, which was verified empirically. Under these conditions, subjective valence ratings across the two time periods correlated positively with an area of left lateral OFC (and right lateral OFC trended towards significance). Regions of the left OFC were also significantly more active when participants were hungry than when satiated on a food. This effect was found regardless of which food was satiated (chocolate milk or tomato juice). Similar regions of OFC have been found in studies investigating the psychophysical synergy of appetizing unmami tastes (de Araujo, Krigelbach, Rolls, & Hobden, 2003), and other non-linear sensory effects such as combining taste and olfactory stimuli like a sucrose taste and strawberry smell (de Araujo et al., 2003). Furthermore, the OFC has also been implicated in subjective pleasantness ratings of other sensory pleasures such as sexual orgasms, drugs, chocolate, and music (for review see Kringelbach & Berridge, 2009).

Hedonic experience is also now discussed as a distinguishable component of reward processing (for reviews see Berridge & Kringelbach, 2008; Kringelback & Berridge, in press). More specifically, 'liking' or the pleasure component of reward has been distinguished from 'wanting,' the motivational component of reward. For example, one can want something without feeling pleasure, an ugly feature of some drug addictions. Interestingly, hedonic liking has been the least studied component in the neuroscience of reward, perhaps because of its subjective quality. Recently, however, researchers have capitalized on the observation that both newborn human infants and rats exhibit tongue protrusions for sweet tastes vs. gapes for bitter tastes. These behaviors have allowed researchers to begin studying 'liking' or sensory pleasure in animals. Neural methods linked with causation (e.g., stimulation and lesioning) have further suggested that 'hedonic hotspots' may exist in the nucleus accumbens, ventral pallidum, and brainstem of animals. As discussed above, OFC appears to be involved in coding pleasure (and possibly other cortical regions such as the insula, medial prefrontal, and cingulated cortices), but the causal properties of these cortical areas are still unclear. In general, much remains to be understood about how so called subcortical hedonic hotspots interact with cortical regions to represent pleasure.

To summarize, studies that focus on the affective properties of basic sensory experiences have provided initial insight into how core dimensions of emotion might be grounded in the brain. These studies point to central roles for the OFC in valence and the amygdala in arousal (or alternatively in the slightly different construct of emotional salience). It is important to recognize, however, that evaluating basic sensory stimuli differs in important ways from the complex situations that people usually describe as emotional. Affective dimensions of chemosensory experience are tightly tied to properties of the stimulus, which is why they can so easily be modified to fit in quadrants of affective space. Conversely, consider manipulating the hedonic experience and intensity of complex visual scenes. Simple changes in luminance, for example, will not suffice. It is possible that the intensity of a taste or smell differs considerably from a state of energy or activation that characterizes the concept of arousal as defined in the emotion literature.

Do studies investigating affective dimensions of complex stimuli find similar patterns? Perhaps surprising, amygdala involvement in arousal is the most consistent finding among stimulus domains. Early studies showed that the amygdala was active for both positive and negative visual scenes and words relative to a neutral control condition, suggesting a role in arousal (Hamann, Ely, Grafton, & Kilts, 1999; Hamann, Ely, Hoffman, & Kilts, 2002; Hamann & Mao, 2001). Consistent with these findings, amygdala activation was observed for high arousal stimuli irrespective of valence (positive, negative) or stimulus type (pictures or words; Kensinger & Schacter, 2006). Arousal ratings have also been shown to correlate with amygdala activity for both positive and negative words (Lewis, Critchley, Rotchtein, & Dolan, 2007). In the only study to manipulate arousal across emotion categories (disgust, fear, happiness, sadness), bilateral amygdala activation was observed for more intense facial expressions, defined as morphs closer to a typical category expression than a neutral expression, in all categories (Winston, O'Doherty, & Dolan, 2003). It is important to note that these findings are also consistent with the idea that the amygdala may not be coding arousal per se, but some aspect associated with emotionally salient stimuli (e.g., importance, uncertainty) relevant to the sense of mobilization that characterizes arousal.

Some findings, however, are inconsistent with this pattern. When Anders et al. (2008) categorized visual scenes and sounds into six categories by valence (positive, negative, and neutral) and arousal (high and low) on the basis of post scan ratings, they found complex valence-arousal interactions in the amygdala. These findings should be interpreted with caution, however, because several of the rating-defined categories were based on very few trials. Other studies have found that amygdala activity is negatively correlated with arousal ratings of words, but positively correlated with ratings of words or faces when coded for valence extremes (i.e., when highly pleasant and unpleasant stimuli receive the same code; Gerber et al., 2008; Posner et al., 2008). These findings were interpreted as support for an emotional saliency hypothesis, as opposed to an arousal hypothesis. The negative correlations with arousal are difficult to interpret, however, given that the duration of stimulus presentation was unusually long (18 seconds).

Only a few studies have examined valence as a continuous dimension that is dissociated from arousal. Lewis et al. (2007) modeled valence ratings extracted from word norms in three different ways. Participants viewed these words while being scanned and performing a self-reference task. Regressors were coded as (1) bipolor from pleasant to unpleasant, (2) as U-shaped deviations from neutral with highly pleasant and unpleasant rating receiving the same code, and (3) independently within positive and negative word classes. Regressors for arousal norms and a valence x arousal interaction term were also included. Activity in regions of lateral OFC was correlated with U-shaped coding of valence, as well as increasing pleasantness in the positive category and increasing unpleasantness in the negative category. The anterior and subgenual cingulate showed a slightly different pattern, correlating with the U-shaped coding and increasing unpleasantness. Finally, activity in the anterior insula was only correlated with increasing pleasantness. No regions correlated with the bipolor valence regressor.

OFC activation has not been observed in other studies using faces, visual scenes, or words. Two studies examined faces and words, respectively, using online valence ratings as bipolor regressors (Posner et al., 2008; Gerber et al., 2008). The main findings in these studies were that valence ratings correlated negatively with dorsal and lateral frontal OFC regions. The only positive correlation with valence was observed in the anterior insula for words, a somewhat similar finding to Lewis et al. Another study found similar prefrontal regions to be correlated with subjects' mean valence ratings (i.e., ratings were not analyzed on a trial-by-trial basis; Grimm et al., 2006).

It should be noted that many studies have examined the distinction between positive and negative stimuli, which are usually both high on the arousal dimension. In a meta-analysis of these studies, positive stimuli were found to activate the basal ganglia more than negative stimuli; negative stimuli activated the insula and cerebellum more than positive stimuli (Wager, Phan, Liberzon, & Taylor, 2003). Although positive and negative are often used synonymously with the term valence, the momentary experience of pleasure may actually be quite different the conceptualizing a stimulus as positive or negative.

In summary, studies of complex stimuli suggest a similar role for the amygdala in arousal. Less support, however, has been garnered for OFC involvement in valence coding. Thus, it remains unclear whether OFC function is restricted to sensory pleasures associated with the secondary gustatory and olfactory cortices. Studies of the affective properties of complex stimuli also highlight methodological challenges. First, it can be difficult to dissociate valence and arousal (e.g., many of the standard materials used to study emotion categories are highly arousing). Second, valence and arousal have been measured using norms, offline post-scanning ratings, and online ratings. Using norms may not adequately address individual differences, and offline ratings assume that participants perceive the stimulus the same way on different occasions. Third, it is often unclear if participants are evaluating whether a visual scene, for example, is a known to elicit pleasure, or whether they are evaluating the actual hedonic experience that the scene has induced in them (Levenson, 2003).

Overview of the Current Study

The current study builds on previous research in a number of ways. Our primary contribution was to examine the core affect properties of valence and arousal during experiences of several emotion categories. In an fMRI paradigm, participants immersed themselves an emotion, and then focused on and rated valence or arousal. Importantly, ratings were made across numerous scenarios categorized as one of three proposed discrete emotions: fear, sadness, or happiness. According to the core affect hypothesis, the regions correlating with valence and arousal ratings should be the same across the three discrete emotions.

To vary valence and arousal within each category, emotion scenarios were written and recorded that were typical (e.g., being disappointed by a significant other for sadness) and atypical (e.g., nostalgic remembering of a pleasant time for sadness) of a category. Although atypical scenarios are not often discussed in discrete emotion approaches, they do involve what is considered the antecedent event that triggers a specific discrete emotion program (e.g., loss for sadness). Atypical scenarios may simply be situations in which individual differences in previous experience (or what is sometimes referred to as 'learning history') may influence the appraisal process to a greater degree. To ensure that participants appraised the atypical scenarios similarly to the typical scenarios, all scenarios were explicitly categorized as fear, sadness, or happiness.

Using this procedure, several methodological concerns were addressed. Participants vividly imagined experiencing emotional scenarios to elicit a first person affective experience. Using this imagery technique addressed concerns that participant judgments reflect the known properties of a stimulus instead of the participant's feeling state. Participants also made online ratings of valence and arousal in the imagined scenarios. These ratings were used in correlation analyses instead of norms, which are less sensitive to individual differences, or offline ratings, which may not reflect the initial experience. Finally, the valence and arousal characteristics of the scenarios were constructed so the valence and arousal dimensions varied independently. This design allowed us to distinguish the contributions of arousal and valence.

These new methods were used to test the hypothesis that the same brain regions would correlate with valence and arousal during experiences categorized as fear, sadness, or happiness. Based on patterns in previous literature, a natural prediction was that regions of OFC would correlate with valence across categories whereas the amygdala would correlate with arousal across categories.

References

An, X., Bandler, R., Ongur, D., & Price, J.L. (1998). Prefrontal cortical projections to longitudinal columns in the midbrain periaqueductal gray in macaque monkeys. *The Journal of Comparative Neurology, 401*, 455-479.

Anders, S., Eeippert, F., Weiskopf, N., & Veit, R. (2008). The human amygdala is sensitive to the valence of pictures and sounds irrespective of arousal: an fMRI study. *Social Cognitive and Affective Neuroscience, 3*, 233-243.

Anderson, A.K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D.G., Glover, G., Gabrieli, J.D.E., & Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience, 6*(2), 196-202.

Bandler, R., & Shipley, M.T. (1994). Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? *Trends in Neurosciences, 17*(9), 379-389.

Barrett, L.F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society B, 364*, 1325-1334.

Barrett, L. F. (2006a). Emotions as natural kinds? *Perspectives on Psychological Science*, 1, 28-58.

Barrett, L. F. (2006b). Valence as a basic building block of emotional life. *Journal of Research in Personality*, 40, 35-55.

Barrett, L.F. & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology, 41*, 167-218.

Barrett, L. F., Lindquist, K., Bliss-Moreau, E., Duncan, S., Gendron, M., Mize, J., &

Brennan, L. (2007). Of mice and men: Natural kinds of emotion in the mammalian brain? *Perspectives on Psychological Science, 2*, 297-312.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology, 58*, 373-403.

Barsalou, L.W. (1999). Perceptual symbols systems. *Behavioral and Brain Sciences, 22*, 577-609.

Barsalou, L.W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes, 18*, 513-562.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.

Berridge, K.C., & Kringelbach, M.L. (2008). Affective neuroscience of pleasure: reward in humans and animals*. Psychopharmacology, 199*, 457-480.

Buck, R. (1999). The biological affects: A Typology. *Psychological Review, 106*(2), 301-336.

Cosmides, L. & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of Emotions* (pp.91-115). New York: Guilford Press.

Damasio, A.R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Penguin Books Ltd.

Damasio, A.R., Grabowski, T.J., Bechara, A., Damasio, H., Ponto, L.L.B., Parvizi, J., & Hichwa, R.D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience, 3*(10), 1049-1056.

De Araujo, I.E.T., Kringelbach, M.L., Rolls, E.T., & Hobden, P. (2003). Representation of umami taste in the human brain. *Journal of Neurophysiology, 90*, 313-319.

De Araujo, I.E.T., Rolls, E.T., Kringelbach, M.L., McGlone, F., & Phillips, N. (2003). Taste olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *European Journal of Neuroscience, 18*, 2059-2068.

Duncan, S., & Barrett, L. F. (2007). Affect as a form of cognition: A neurobiological analysis. *Cognition and Emotion* 21, 1184-1211.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*(3/4), 169-200.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115*(2), 268-287.

Ekman, P. (2003). *Emotions Revealed*. New York, NY: Henry Holt and Company.

Fontaine, J.R.J., Scherer, K.R., Roesch, E.B., & Ellsworth, P.C. (2004). The world of emotions is not two-dimensional. *Psychological Science, 18*(12), 1050-1057.

Gerber, A.J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J., & Peterson, B.S. (2008). An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotion faces. *Neuropsychologia, 46*, 2129-2139.

Grimm, S., Schidt, C.F., Bermpohl, F., Heinzel, A., Dahlem, Y., Wyss, M., Hell, D., Boesiger, P., Boeker, H., & Northoff, G. (2006). Segregated neural representation of distinct emotion dimensions in the prefrontal cortex—an fMRI study. *NeuroImage, 30*, 325-340.

Hamann, S. B., Ely, T.D., Hoffman, J.M., & Kilts, C.D. (2002). Activation of the human amygdala in positive and negative emotion. *Psychological Science, 13*(2), 135-141.

Hamann, S. B., Ely, T.D., Grafton, S.T., & Kilts, C.D. (2002). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience, 2*(3), 289-293.

Hamann, S. & Mao, H. (2002). Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *NeuroReport, 13*(2), 15-19.

Izard, C.E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science, 2*(3), 260-280.

Kensinger, E.A., & Schacter, D.L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience, 6*(2), 110-126.

Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T.D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *NeuroImage, 42*, 998-1031.

Kringelbach, M.L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. *Nature Reviews Neuroscience, 6*, 691-702.

Kringelback, M.L., & Berridge, K.C. (2009). Towards a functional neuroanatomy of pleasure and happiness. *Trends in Cognitive Sciences, 13*(11), 479-487.

Kringelbach, M.L., O'Doherty, J., Rolls, E.T., & Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral Cortex, 13*, 1064-1071.

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

Laird, A. M., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., Turkeltaub, P. E., Kochunov, P., & Fox, P.T. (2005). ALE meta-analysis: controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping, 25*(1), 155-164.

LeDoux, J. (1996). *The emotional brain: The mysterious underpinning of emotional life*. New York: Simon & Schuster.

Levenson, R.W. (2003). Autonomic specificity and emotion. In R.J. Davidson, K.R. Scherer, & H.H. Goldsmith (Eds.), *Handbook of Affective Sciences* (pp. 212-224). New York: Oxford University Press.

Lewis, P.A., Critchley, H.D., Rotshtein, P., & Dolan, R.J. (2007). Neural correlates of processing valence and arousal in affective words. *Cerebral cortex, 17*, 742-748.

Lindquist, K., & Barrett, L.F. (2008). Constructing emotion: The experience of fear as a conceptual act. *Psychological Science, 19*, 898-903.

Mesquita, B. & Frijda, N.H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin, 112*(2), 179-204.

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

Ongur, D., & Price, J.L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys, and humans. *Cerebral Cortex, 10*, 206-219.

Ortony, A. & Turner, T.J. (1990). What's basic about basic emotions. *Psychological Review, 97*(3), 315-331.

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

Panksepp, J. (2000). Emotions as natural kinds within the mammalian brain. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 137–156). New York: Guilford.

Panksepp, J. (2005). Affective consciouness: Core emotional feelings in animals and humans. *Consciousness and Cognition, 14*, 30-80.

Panksepp, J. (2007). Neurologizing the psychology of affects: how appraisal-based constructivism and basic emotion theory can coexist. *Perspectives on Psychological Science, 2*(3), 281-296.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neursocience, 9*, 148-158.

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

Plutchik, R. (2003). *Emotions and life: Perspectives from psychology, biology, and evolution*. Washington, DC: American Psychological Association.

Posner, J., Russell, J.A., Gerber, A., Gorman, D., Colibazzi, Yu, S., Wang, Z., Kangarlu, A., Zhu, H., & Peterson, B.S. (2009). The neurophysiological bases of emotion: an fMRI study of the affective circumplex using emotion-denoting words. *Human Brain Mapping, 30*, 883-895.

Rolls, E.T., Kringelbach, M.L., & de Araujo, I.E.T. (2003). Different represenations of pleasant and unpleasant odours in the human brain. *European Journal of Neuroscience, 18*, 695-703.

Ruge, H., Goschke, T., Braver, T.S. (2009). Separating event-related BOLD components within trials: The partial-trial design revisited. *Neuroimage, 47*, 501-513.

Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161-1178.

Russell, J.A. (1994). Is there a universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*(1), 102-141.

Russell, J.A. (2003). Core affect of the psychological construction of emotion. *Psychological Review, 110*(1), 145-172.

Small, D.M., Gregory, M.D., Mak, Y.E., Gitelman, D., Mesulam, M.M., & Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron, 39*, 701-711.

Spelke, E.S., & Kinzler, K.D. (2007). Core knowledge. *Developmental Science, 10*(1), 89-96.

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of Cognitive Neuroscience, 22*(12), 2864-2885.

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

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence*. Journal of Personality and Social Psychology, 76*(5), 820-838.

Weiskopf, N., Hutton, C., Josephs, O., & Deichmann (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-rain analysis at 3T and 1.5T. *NeuroImage, 33*, 493-504.

Wintson, J.S., Gottfried, J.A., Kilner, J.M., & Dolan, R.J. (2005). Integrated neural representations of odor intensity and affective valence in human amygdala. *The Journal of Neuroscience, 25*(39), 8903-8907.

Winston, J.S., O'Doherty, J., & Dolan, R.J. (2003). Common and distinct neural responses during direct and incidental processing multiple facial emotions. *NeuroImage, 20*, 84-97.

General Introduction Addendum

Data analysis revealed two distinct sets of findings in this experiment. Thus, it seemed most appropriate to write separate manuscripts to frame and discuss these findings. Both manuscripts have been written in formats appropriate for submission to journals, which will take place in the near future.

The first manuscript follows from the General Introduction and presents the results of the valence and arousal correlation analyses across emotion categories. This manuscript is written in short format appropriate for relevant journals and discusses correlations with valence and arousal that we observed in orbitofrontal cortex and amygdala, respectively.

The second manuscirpt was a product of the methods used to test the core affect hypothesis, in which we constructed atypical fear, happiness, and sadness scenarios to vary the valence dimension as much as possible (e.g., pleasant fear, unpleasant happiness, pleasant sadness). A clear pattern in the data was that these atypical instances of the emotions showed robust differences from the typical instances. We thus crafted a second manuscript to discuss greater activations for atypical relative to typical emotion states in large-scale brain networks, including default mode and attention networks.

Finally, these two manuscripts are followed by a general discussion, which address a few general themes of both. Also addressed are important themes for future research.

Deconstructing Emotion Experience: Neural Evidence for Core Affect

Christine D. Wilson-Mendenhall
Emory University

Lisa Feldman Barrett
Northeastern University

Lawrence W. Barsalou
Emory University

Abstract

Basic emotion theories propose that non-reducible affect programs in the brain produce a small number of discrete emotions such as fear or happiness. An alternative view is that core affect is a basic element of emotion experience, and that neural systems underlie core affect properties of valence and arousal. To test the core affect hypothesis, we developed fear, happiness, and sadness scenarios that varied in valence and arousal. During an fMRI session, participants imagined these scenarios from a first-person perspective, and then focused on and rated the valence or arousal of the induced emotion. The core affect view predicts that the brain regions correlating with valence or arousal ratings should be the same across experiences of all three emotions. Based on previous literature, we predicted that valence ratings would correlate with activity in orbitofrontal cortex (OFC) and that arousal ratings would correlate with activity in the amygdala. As predicted, across all three emotion categories, activity in medial OFC increased as rated unpleasantness decreased and rated pleasantness increased, whereas activity in left amygdala increased as rated arousal increased. Interestingly, a gradient in medial OFC was also observed in which voxels most sensitive to unpleasantness were more inferior to those most sensitive to pleasantness. The findings support the view that valence and arousal are basic properties of core affect that contribute to emotional experience.

Deconstructing Emotion Experience: Neural Evidence for Core Affect

One step towards understanding the complex phenomena we call 'emotion' is identifying the neural systems that underlie these mental states (Barrett, 2009; Barrett, Mesquita, Ochsner, & Gross, 2007). A key question is whether fundamental neural systems supporting emotion take the form of modules that produce some number of basic emotions states (e.g., fear, happiness) or take the form of circuits performing primitive valuation and salience operations that support core affect properties of valence and arousal.

A common view is that emotion programs evolved in the brain to detect and respond to situations critical for survival (Ekman, 1992; Izard, 2007; Panksepp, 2000). For example, a *fear* program evolved to detect dangerous situations and quickly trigger flight or fight responses. Although the number of these basic or discrete emotions is not agreed upon (cf. Ortony & Turner, 1990), the emotion categories fear, happy, sad, anger, and disgust are often among those proposed and have been the focus of three neuroimaging meta-analyses (Murphy, Nimmo-Smith, & Lawrence, 2003; Phan, Wager, Taylor, & Liberzon, 2002; Vytal & Hamann, 2010). The goal of these meta-analyses was to identify brain regions that discriminate a given basic emotion category from all others. The most recent meta-analytic review identified one or more brain regions showing this pattern for each of the five emotions, interpreting these findings as support for basic emotion theories (Vytal & Hamann, 2010).

Other methodologically rigorous meta-analyses, however, have shown that emotions falling in these categories activate many of the same neural systems

(Kober et al., 2009; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, submitted). Furthermore, reviews of evidence from other domains, including facial and vocal signals, peripheral nervous systems responses, and voluntary behaviors, offer little support for basic emotion theories (Barrett, 2006a; Ortony & Turner, 1990). An alternative view, Conceptual Act Theory (CAT), suggests that emotion states emerge from distributed neural circuitry serving an array of functions, with no brain system dedicated to a specific discrete emotion (Barrett, 2006a; Lindquist & Barrett, 2008; Wilson-Mendenhall, Barrett, & Barsalou, submitted). In this view and in other dimensional approaches, an essential element of emotion experience (but not the only one) is the neural systems that evolved to support core affect. Core affect is a continuous and fluctuating state of pleasure or displeasure with some level of arousal (Russell, 2003; Russell & Barrett, 1999). Initial evidence for core affect accrued from studies measuring subjective experience, which identified statistical components for valence and arousal using data reduction techniques (for reviews see Barrett, 2006b; Barrett & Bliss-Moreau, 2009; Russell, 2003; Russell & Barrett, 1999).

The neurobiological grounding of core affect properties, however, has been most widely studied in sensory domains, investigating the affective properties of taste, smell, and touch (Anderson et al., 2003; Kringelbach, O'Doherty, Rolls, & Andrews, 2003; Rolls, Grabenhorst, & Parris, 2008; Rolls, Kringelbach, & de Araujo, 2003; Rolls et al., 2003; Small et al., 2003; Winston et al., 2005). Consistently across studies, orbitofrontal cortex (OFC) has shown sensitivity to the valence of stimuli whereas amygdala has shown sensitivity to intensity. The neural systems that evolved to support the valuation and salience processing that guides basic sensory experiences (e.g., consume good-tasting foods; avoid painful touch) may also be critical for the affect experienced during social interactions (Berridge & Kringelback, 2008; Eisenberger, Lieberman, & Williams, 2003; Eisenberger & Lieberman, 2004). The neural correlates of subjective experiences for valence and arousal during complex, 'higher' emotional experiences like fear or sadness, however, remain unclear. Studies that have investigated the valence and arousal properties of more complex emotional word, scene, and face stimuli, often using norms collected from a separate set of participants, have produced mixed results. The most consistent finding is that the amygdala appears to be associated with high arousal states (Hamann, Ely, Hoffman, & Kilts, 1999; Kensinger & Schacter, 2006; Lewis, Critchely, Rotchetin, & Dolan, 2007; Winston, O'Doherty, & Dolan, 2003), though not all studies find this pattern (Anders et al., 2008; Gerber et al., 2008; Posner et al., 2008). Only a few studies have examined valence as a continuous dimension dissociated from arousal (Colibazzi et al., 2010; Lewis et al., 2007; Posner et al., 2008; Gerber et al., 2008). One of these studies found activity in OFC associated with valence (Lewis et al., 2007) whereas the others observed activation in more dorsal frontal regions.

The goal of the current study was to examine the core affect dimensions of valence and arousal experienced as qualities of complex social emotions categorized as fear, happiness, or sadness. To test the core affect hypothesis, we varied valence and arousal within each category, achieving a range that allowed us to measure the relation between subjective ratings and brain activity for each emotion. Participants imagined various fear, happiness, or sadness scenarios from a first-person perspective, and then focused on and rated the valence or arousal quality of their feeling during fMRI scans. If core affect is a fundamental element of emotion experience, valence and arousal ratings within each emotion category should independently correlate with activity in the same brain regions. More specifically, we predicted that OFC activity would be correlated with valence ratings and amygdala activity correlated with arousal ratings across all categories.

Method

Design & Participants

The experiment contained two training sessions and an fMRI scan session. The first training session occurred 24 to 48 hours before the second training session, followed immediately by the scan session. The fMRI design was optimized to examine brain activity during six critical conditions created by the factors of emotion category (fear, happiness, sadness) and dimension (valence, arousal). Two trial types existed in all conditions. In 144 *complete* trials, participants first imagined a fear, happiness, or sadness scenario, and then focused on and rated the valence or arousal quality of the feeling. In 36 *partial* trials, participants only imagined a scenario. Partial trials were included so scenario events could be mathematically separated from subsequent focus-rate events when complete trials were analyzed (Ollinger, Corbetta, & Shulman, 2001; Ollinger, Shulman, & Corbetta, 2001; Ruge, Goschke, & Braver, 2009). Separation of these events was critical for isolating activity during valence and arousal focus-rate events *after* the fear, happiness, or sadness scenario had been mentally constructed. All the activations reported here occurred during focus-rate events. As necessary for this type of catch trial design, partial trials were unpredictable and accounted for 20% of all trials.

Six imaging runs consisted of one arousal block and one valence block, with block order counterbalanced across runs and trials as events within blocks. In arousal blocks, participants focused on and rated the arousal quality of their feeling state during complete trials. In valence blocks, participants focused on and rated the valence quality. Within each block, four complete trials and one partial trial for each category were presented amidst jittered no-sound baseline periods (ranging from 3-15 s in increments of 3; average ISI = 6.3 s) in a pseudo-random order optimized for deconvolution analysis using optseq2 software.

Across the six runs in the experiment, each fear, happiness, and sadness scenario was presented twice, once in an arousal block and once in a valence block. To create the first version of the experiment, the 24 critical scenarios presented in complete trials and the 6 scenarios presented in partial trials for each category were randomly assigned to the valence blocks of the six runs. Scenarios in the valence block of run one, two, and three were repeated in arousal block of run four, five, and six respectively. Scenarios in the valence block of run four, five, and six were first encountered in arousal block of run one, two, and three. To control for repetition order, a second version was created in which scenarios in valence and arousal blocks were flipped (i.e., scenarios initially rated first for valence were rated first for arousal in the second version and vice versa). Finally, two additional versions were created paralleling the first two in which the order of the runs was simply reversed to control for general stimulus order effects.

Sixteen right-handed, native English speakers from the Emory community, ranging in age from 19-30 (8 female), were randomly assigned to one of the four versions. Each participant received $100 in compensation.

Materials

Scenarios ranging in valence and arousal were written and recorded to induce fear, sadness, and happiness states. To create range within the valence dimension for each category, we developed scenarios of atypical valence (e.g., pleasant sadness of nostalgia, pleasant fear of risk-taking, unpleasant happiness of relief; see Appendix for examples). A full and core form of each scenario was developed, the core being a subset of the full form. The full form provided a richly detailed and affectively compelling description of a fear, sadness, or happiness episode. The core form served to minimize presentation time in the scanner so the number of trials necessary for a powerful design could be implemented. In both forms, scenarios were explicitly categorized as fear, sadness, or happiness to avoid ambiguity. More details on the construction and selection of scenarios can be found in the *Supplemental Materials*.

Procedure

In the first training session, participants provided informed consent and were screened for problems that could arise in the MRI environment. Participants had no history of psychiatric illness and were not taking psychotropic medication. The purpose of the first session was to familiarize participants with the scenarios by having them actively practice a) vividly imagining the full versions of the scenarios they would hear later in the scanner or during practice trials b) reinstating the rich imagery of each full scenario upon hearing the core version c) focusing on and rating the valence or arousal quality of the feeling state induced by a scenario. Before performing any task, participants received an overview of the scenarios and were encouraged to imagine each scenario from a first-person perspective as vividly as possible. They were also informed of the use of full and core scenario versions, and the relation between them. Participants then listened to the full versions of the scenarios, immersing themselves with eyes closed, and rated their personal familiarity with each imagined emotion. After a short break, they listened to the core versions of the same scenarios, reinstated imagined details from the full versions, and then rated the internal, external, and thought imagery experienced (further encouraging immersion in the imagined scenarios). Finally, participants practiced focusing on and rating the arousal or valence quality of the emotion induced in several practice scenarios.

When participants returned to the lab 24-48 hours later, they began the second session by listening to and vividly imagining each full scenario again. Participants provided one rating of how much they experienced being immersed in the feeling of fear, happiness, or sadness described in the scenario. Imagining the full versions in the second session ensured that participants were reacquainted with all the details just prior to hearing the core versions in the scanner. Participants were then instructed on and practiced the task they would perform in the scanner with scenarios not used in imaging runs. Participants were first informed that they would complete one block of valence trials and one block of arousal trials in each imaging run and that the cue word ‘valence’ or ‘arousal’ would be repeated three times at the beginning of each block. They then practiced the different trial types, beginning with 15 s complete trials. During complete trials, participants were instructed to immerse themselves fully as they listened to a core version of a scenario lasting no longer than 8 s. When they next heard a 1 s 'beep, beep, beep,' they were to continue imagining the emotion and begin centering in on the valence or arousal quality of the feeling (depending on the block), maintaining focus for 3 s. Finally, a 1 s cowbell cued participants to rate their introspective sense of valence or arousal within the next 2 s using the appropriate scale. At this point, participants had received much practice using the 5-point valence and arousal scales with their eyes closed. The points on the valence scale from left to right were very unpleasant, somewhat unpleasant, neutral, somewhat pleasant, and very pleasant; the points on the arousal scale were low, medium-low, medium, medium-high, and high. During 9 s partial trials, participants heard a 1 s 'whoosh' sound when the 8 s scenario concluded, which signified the end of the trial. During baseline rest trials, participants cleared their mind during the 3-15 s period of no sound as they waited to hear the next scenario begin. After practicing each trial type separately, participants engaged in several short arousal and valence blocks with all trial types intermixed like in each block of the imaging experiment. Participants were then informed that although scanner blocks would be longer, the task itself would be exactly the same with all trial types occurring in a random, unpredictable order. A more detailed description of the training sessions can be found in the *Supplemental Materials*.

Following practice, the experimenter and participant walked to the MRI facility. Once the participant was situated comfortably in the scanner, an initial anatomical scan was collected. The participant was then reminded of the task using brief instructions and of the valence and arousal scales by pressing the appropriate button as the experimenter named each point on the scale. When the participant was ready, the experimenter initiated the first functional run and then continued with the next five runs, pausing for a short break between runs. A second anatomical scan was collected last. Total time spent in the scanner was a little over an hour.

Imaging and Analysis

Images were collected at the Emory Biomedical Imaging Technology Center on a 3T Siemens Trio scanner and preprocessed using standard methods in AFNI (Cox, 1996; see *Supplemental Materials* for details). Two critical regression analyses were performed on each participant's preprocessed data using canonical Gamma functions to model the hemodynamic response. In the first analysis, the onset times were specified for five conditions: cues beginning each block, scenario events during valence blocks, scenario events during arousal blocks, focus-rate events during valence blocks, and focus-rate events during arousal blocks. Scenario events included the 9 s during which participants heard the scenario and the brief auditory cue that followed in complete and partial trials. Modeling the scenarios in complete and partial trials as a single condition allowed for the mathematical separation of the scenario period from the focus-rate events in complete trials. The focus-rate period included the 6 s during which participants focused on and rated the valence or arousal quality of the feeling. Because the scenario and focus-rate conditions were each several seconds long, the gamma function used to model the hemodynamic response was convolved with a boxcar function reflecting the duration of the event.

Participant valence ratings were specified for each trial in the valence focus-rate condition and arousal ratings specified for each trail in the arousal focus-rate condition. The following numerical codes were used for valence (1-very unpleasant 2-somewhat unpleasant 3-neutral 4-somewhat pleasant 5-very pleasant) and arousal (1-low 2-medium-low 3-medium 4-medium-high 5-high). Any missing rating was replaced with the mean rating (1% of trials on average). For the focus-rate conditions, both the onset times and ratings were entered into the regression using the amplitude modulation option in AFNI. This option specified two regressors for each focus-rate condition, which were used to detect: 1) voxels in which activity is correlated with the ratings (also known as a parametric regressor); 2) voxels in which activity is constant for the condition and show no correlation with the ratings. Each participant's betas produced from the first parametric regressor for focus-rate conditions (i.e., indicating the strength of the correlation with valence or arousal ratings) were next entered into a random effects group analysis. In this analysis, the critical statistic for each condition was a *t* test indicating if the mean across subjects was significantly different from zero (zero indicating no correlation between brain activity and the ratings). To test our regional hypotheses, the group analysis was computed within anatomical masks for medial OFC, bilateral OFC, and bilateral amygdalae (see *Supplemental Materials* for more details). Within each mask, a voxel-wise threshold of $p < .005$ was used in conjunction with an extent threshold determined by AFNI clustsim as the $p < .05$ corrected threshold (12 voxels for medial OFC, 9 voxels for lateral OFC, 3 voxels for amygdala).

Any significant cluster identified in the first analysis was used to mask a second analysis, which analyzed the emotion categories separately. The critical difference from the first analysis was that each scenario and focus-rate condition was split into three conditions for the emotion categories fear, happiness, and sadness. Otherwise the analysis was exactly the same. Participant betas produced from the parametric regressors for the six category focus-rate conditions (i.e., fear-valence, happiness-valence, sadness-valence, fear-arousal, happiness-arousal, sadness-arousal) were then entered into a random effects group analysis in an identical manner to the first analysis. At the group level, voxel-wise *t* statistics representing significant correlations with either valence or arousal for each category ($p < .05$) were entered into a conjunction analysis. The conjunction was only computed within clusters identified in the first analysis to determine if these voxels were significantly correlated with valence or arousal in each emotion category. This key analysis allowed us to examine whether each voxel correlated with valence or arousal in the first analysis, which was conducted across categories, was correlated with valence or arousal in one or more emotion categories when each category was modeled separately.

Results

Valence

Activity in a medial OFC cluster was correlated with valence ratings, supporting our first prediction (peak -2 38 -13, 24 voxels; see Figure 1A). In this cluster, activity increased as unpleasantness ratings decreased and pleasantness ratings increased on the bipolar scale. Furthermore, 92% of the voxels within this cluster showed a significant correlation with one or more of the emotion categories when each category was modeled independently. As shown in Figure 1A, 50% of these voxels were correlated with valence ratings in multiple categories.

[Insert Figure 1 about here]

Interestingly, voxels correlated with valence ratings made during sadness were more inferior (z = -12 or below) than voxels correlated with valence ratings made during happiness (z = -11 or above). No voxel showed sensitivity to valence ratings made during sadness and during happiness. Many voxels, however, were correlated with valence ratings made during fear and happiness or during fear and sadness. Although the ratings for each category displayed a range that spanned both sides of the bipolar scale (see Table S1 for descriptive statistics), the mean valence rating for fear was closest to the scale mid-point of 3 (M = 2.81) with sadness falling below (M = 2.48) and happiness above (M = 3.57). To further investigate if voxels most sensitive to unpleasantness were located inferior to those most sensitive to pleasantness, we re-coded the valence ratings and recomputed the first analysis. First, we coded valence ratings with a unipolar focus on the unpleasant end of the scale (2-very unpleasant 1-somewhat unpleasant 0-neutral 0-somewhat pleasant 0-very pleasant), and found a cluster in medial OFC that increased in activity as unpleasantness decreased (peak -2 32 -16; 19 voxels).

Second, we coded valence ratings with unipolar focus on the pleasant side of the scale (0-very unpleasant 0-somewhat unpleasant 0-neutral 1-somewhat pleasant 2-very pleasant), and found a cluster in medial OFC that increased in activity as pleasantness increased (peak -2 44 -4; 36 voxels). We then computed a conjunction to compare the spatial extents of the bipolar and unipolar clusters. As can be seen in Figure 1B, the cluster sensitive to unpleasantness was most inferior in medial OFC, and the cluster sensitive to pleasantness was most superior in medial OFC, with the cluster sensitive to the bipolar ratings in between the two.

Arousal

In support of our second prediction, activity in left amygdala was correlated with arousal ratings (peak -23 -2 -10, 6 voxels; see Figure 2A). In this cluster, activity increased as arousal ratings increased from low to high on the scale. As shown in Figure 2A, voxels within this cluster correlated with arousal ratings made during happiness, sadness, or both happiness and sadness when each category was modeled independently. Because the mean fear arousal rating was higher (M = 4.13) than happiness (M = 3.40) and sadness (M = 3.38) and fear ratings varied less than the other categories (Levene's test $p < .05$), restricted range was a potential limitation for fear. Thus we performed an additional analysis in which each category was split into a high and low arousal condition and modeled using standard methods (see *Supplemental Materials*). Participant betas for the high and low condition of each category were extracted from the left amygdala cluster identified above and the means plotted in Figure 2B. In each category, the activity was significantly greater ($p < .05$) in the high arousal condition than the low arousal condition, including fear. As shown in Figure 2B, mean activity in the low arousal conditions was higher in fear than the other categories (marginally significant), consistent with the behavioral ratings.

[Insert Figure 2 about here]

An unexpected result was that clusters in left and right lateral OFC displayed the opposite arousal pattern, with activity increasing as arousal ratings decreased from high to low (left peak -44 47 -1, 19 voxels; right peak 38 47 -7, 33 voxels; see Figure S1). In the left OFC cluster, activity in 15 voxels was modulated by happiness, sadness, and/or fear whereas in right OFC cluster, activity in 31 voxels was modulated by sadness and/or fear.

Discussion

The results supported our predictions that valence ratings made during experiences of fear, happiness, and sadness would correlate with OFC activity, and that arousal ratings would correlate with amygdala activity. Our findings suggest that basic neural systems exist to continually evaluate experience and mobilize mental and physical resources across emotions, at times producing conscious feelings of valence and arousal. These systems are arguably the primitive foundation upon which emotion experience is built. This view stands in contrast to basic emotion theories, which propose that evolved affect programs are the foundational mechanisms supporting emotion experience.

Our finding that activity in medial OFC increased as rated unpleasantness decreased and pleasantness increased is consistent with proposals that medial OFC represents and monitors reward values (for reviews see Kringelbach, 2004;

Kringelbach & Rolls, 2004). Interestingly, we found that the voxels most sensitive to decreasing unpleasantness were more inferior in neural space than those most sensitive to increasing pleasantness, with voxels sensitive to the bipolar scaling overlapping in the space between. Although qualitatively different systems supporting positive and negative evaluation are often emphasized (e.g., Cacioppo, Bernston, & Gardner, 1999), theorists have also pointed out that values must be compared for action selection (Cabanac, 2010). The gradient we observed may reflect this integrative processing, which is also consistent with anatomical data suggesting medial OFC supports valuation to guide behavior (Ongur & Price, 2000). To our knowledge, this is the first time such an inferior-superior gradient has been identified.

Although lateral OFC has been suggested to play a role coding punishment leading to behavior change (Kringelbach, 2004; Kringelbach & Rolls, 2004), we did not find correlations with unpleasant states in this region, even when valence was coded with a unipolar focus on the unpleasant side of the scale. Instead, clusters in lateral OFC showed more activity as arousal decreased. Several alternative hypotheses have been proposed for lateral OFC, several of which involve attention, awareness, or modulation of affective states (Eisenberger & Lieberman, 2004; Beer, Knight, & D'Esposito, 2006). Consistent with these accounts, low arousal states may have required greater attentional processing and awareness of internal states.

Increased amygdala activity during high arousal states is consistent with many previous studies (for a review see Costafreda, Brammer, David, & Fu, 2007). The left-lateralization may be associated with the linguistic nature of the scenarios or possibly the more conscious cognitive processing of arousal (Baas, Aleman, & Kahn, 2004; Costafreda et al., 2007; Glascher & Adophs, 2003). The amygdala correlation with arousal across a variety of situations, pleasant and unpleasant, fits with the idea that this structure responds to motivationally salient events that require attention and learning (Whalen, 1998; Whalen et al., 2009). Whereas specific physiological, motor, and mental responses may be engaged on a situational basis, the more generalized function of detecting emotional salience would be a common feature of most high arousal states, important for mobilizing cognitive and physical resources.

In conclusion, the neural evidence presented supports the view that valence and arousal are basic properties of core affect, fundamental to emotional experience. Much remains to be learned about these neural systems and how they interact with other neural circuitry to produce the variety of complex emotional states we experience.

References

Anderson, A.K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D.G., Glover, G., Gabrieli, J.D.E., & Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience, 6*(2), 196-202.

Barrett, L. F. (2006a). Emotions as natural kinds? *Perspectives on Psychological Science, 1*, 28-58.

Barrett, L. F. (2006b). Valence as a basic building block of emotional life. *Journal of Research in Personality, 40*, 35-55.

Barrett, L. F. (2009). The future of psychology: Connecting Mind to Brain. *Perspectives on Psychological Science, 4*, 326-339.

Barrett, L.F. & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology, 41*, 167-218.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology, 58*, 373-403.

Beer, J.S., Knight, R.T., & D'Esposito (2006). Controlling the integration of emotion and cognition. *Psychological Science, 17*(5), 448-453.

Berridge, K.C., & Kringelbach, M.L. (2008). Affective neuroscience of pleasure: reward in humans and animals*.* *Psychopharmacology, 199*, 457-480.

Cabanac, M. (2010). The dialectics of pleasure. In M.L. Kringebach & K.C. Berridge (Eds). *Pleasures of the Brain* (pp. 113-124). USA: Oxford University Press.

Cacioppo, J.T., Bernston, G.G., Gardner, W.L. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*(5), 839-855.

Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., Zhu, H., Kangarlu, A., & Duan, Y. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion, 10*(3), 377-389.

Cox, R.W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*, 162-173.

Eisenberger, N.I., & Lieberman, M.D. (2004). Why rejection hurts: A common neural alarm system for physical and social pain. *Trends in Cognitive Sciences, 8*(7), 294-300.

Eisenberger, N.I., & Lieberman, M.D., & Williams, K.D. (2003). Does rejection hurt? An fMRI study of social exclustion. *Science, 302*, 290-292.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*(3/4), 169-200.

Glascher, J., & Adolphs, R. (2003). Processing of the arousal of subliminal and supraliminal emotional stimuli by the human amygdala. *The Journal of Neuroscience, 23*(32), 10274-10282.

Hamann, S. B., Ely, T.D., Hoffman, J.M., & Kilts, C.D. (2002). Activation of the human amygdala in positive and negative emotion. *Psychological Science, 13*(2), 135-141.

Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T.D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *NeuroImage, 42*, 998-1031.

Kensinger, E.A., & Schacter, D.L. (2006). Processing emotional pictures and words:

Effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience, 6*(2), 110-126.

Kringelbach, M.L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. *Nature Reviews Neuroscience, 6*, 691-702.

Kringelbach, M.L., O'Doherty, J., Rolls, E.T., & Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral Cortex, 13*, 1064-1071.

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

Izard, C.E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science, 2*(3), 260-280.

Lewis, P.A., Critchley, H.D., Rotshtein, P., & Dolan, R.J. (2007). Neural correlates of processing valence and arousal in affective words. *Cerebral cortex, 17*, 742-748.

Lindquist, K., & Barrett, L.F. (2008). Constructing emotion: The experience of fear as a conceptual act. *Psychological Science, 19*, 898-903.

Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., & Barrett, L.F. (submitted). The brain basis of emotion: A meta-analytic review.

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

Ollinger, J.M., Corbetta, M., Shulman, G.L. (2001). Separating Processes within a trial in event-related functional MRI: II. Analysis. *NeuroImage, 13*, 218-229.

Ollinger, J.M., Shulman, G.L., Corbetta, M. (2001). Separating processes within a trial in event-related functional MRI: I. Method. *NeuroImage, 13*, 210-217.

Ongur, D., & Price, J.L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys, and humans. *Cerebral Cortex, 10*, 206-219.

Ortony, A. & Turner, T.J. (1990). What's basic about basic emotions. *Psychological Review, 97*(3), 315-331.

Panksepp, J. (2000). Emotions as natural kinds within the mammalian brain. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 137–156). New York: Guilford.

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

Rolls, E.T., Kringelbach, M.L., & de Araujo, I.E.T. (2003). Different represenations of pleasant and unpleasant odours in the human brain. *European Journal of Neuroscience, 18*, 695-703.

Rolls, E.T., Grabenhorst, F. & Parris, B.A. (2008). Warm pleasant feelings in the brain. *Neuroimage, 41*, 1504-1513.

Rolls, E.T., O'Doherty, J., Kringelback, M.L., Francis, S., Bowtell, R., & McGlone, F. (2003). Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cerebral Cortex, 13*, 308-317.

Ruge, H., Goschke, T., Braver, T.S. (2009). Separating event-related BOLD components within trials: The partial-trial design revisited. *Neuroimage, 47*, 501-513.

Russell, J.A. (2003). Core affect of the psychological construction of emotion. *Psychological Review, 110*(1), 145-172.

Russell, J.A., & Barrett, L.F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*(5), 805-819.

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of Cognitive Neuroscience, 22*(12), 2864-2885.

Whalen, P.J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current directions in Psychological Science, 7*(6), 177-188.

Whalen, P.J., Davis, F.C., Oler, J.A., Kim, H., Kim, M.J., & Neta, M. (2009). Human amygdala responses to facial expressions of emotion. In P.J. Whalen & E.A. Phelps (Eds). *The Human Amygdala* (pp. 265-288). New York: Guilford Press.

Wilson-Mendenhall, C.D., Barrett, L.F., Simmons, W.K., & Barsalou, L.W. (submitted). Grounding emotion in situated conceptualization.

Wintson, J.S., Gottfried, J.A., Kilner, J.M., & Dolan, R.J. (2005). Integrated neural representations of odor intensity and affective valence in human amygdala. *The Journal of Neuroscience, 25*(39), 8903-8907.

Winston, J.S., O'Doherty, J., & Dolan, R.J. (2003). Common and distinct neural responses during direct and incidental processing multiple facial emotions. *NeuroImage, 20*, 84-97

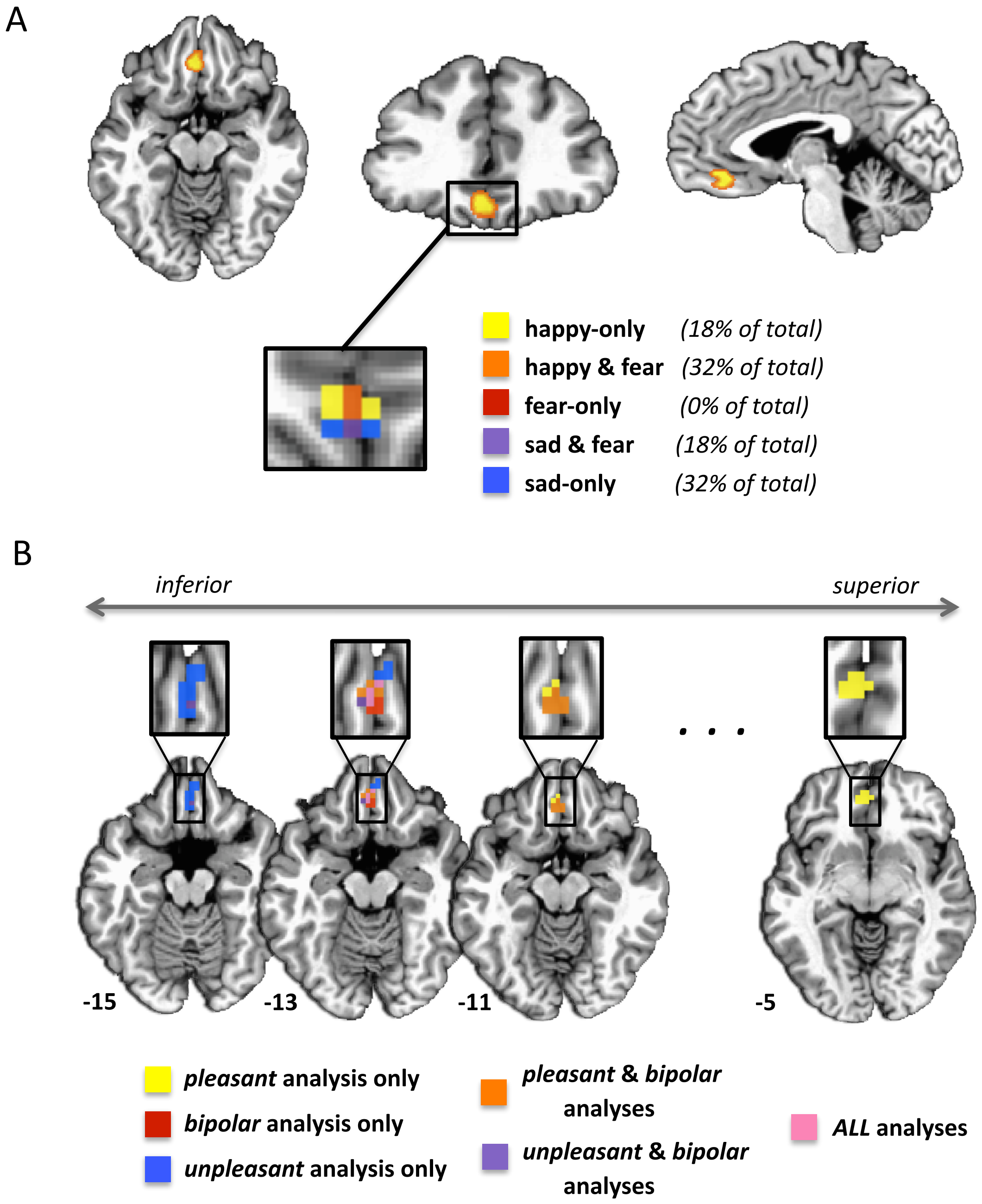


Figure 1. In panel A, the medial OFC cluster positively correlated with bipolar valence ratings across categories is shown in three views, an axial (z = -11), coronal (y = 34), and sagittal slice (x = -2). The magnified box at the coronal slice shows the category results with percentages indicating proportions across the 3D cluster (not the 2D slice). In panel B, the conjunction analysis of the three medial OFC clusters observed in the unipolar-focused and bipolar analyses (across categories) is shown on axial slices moving from inferior to superior. The ellipsis signifies that axial slices between -11 and -5 look very similar to the -11 slice shown.

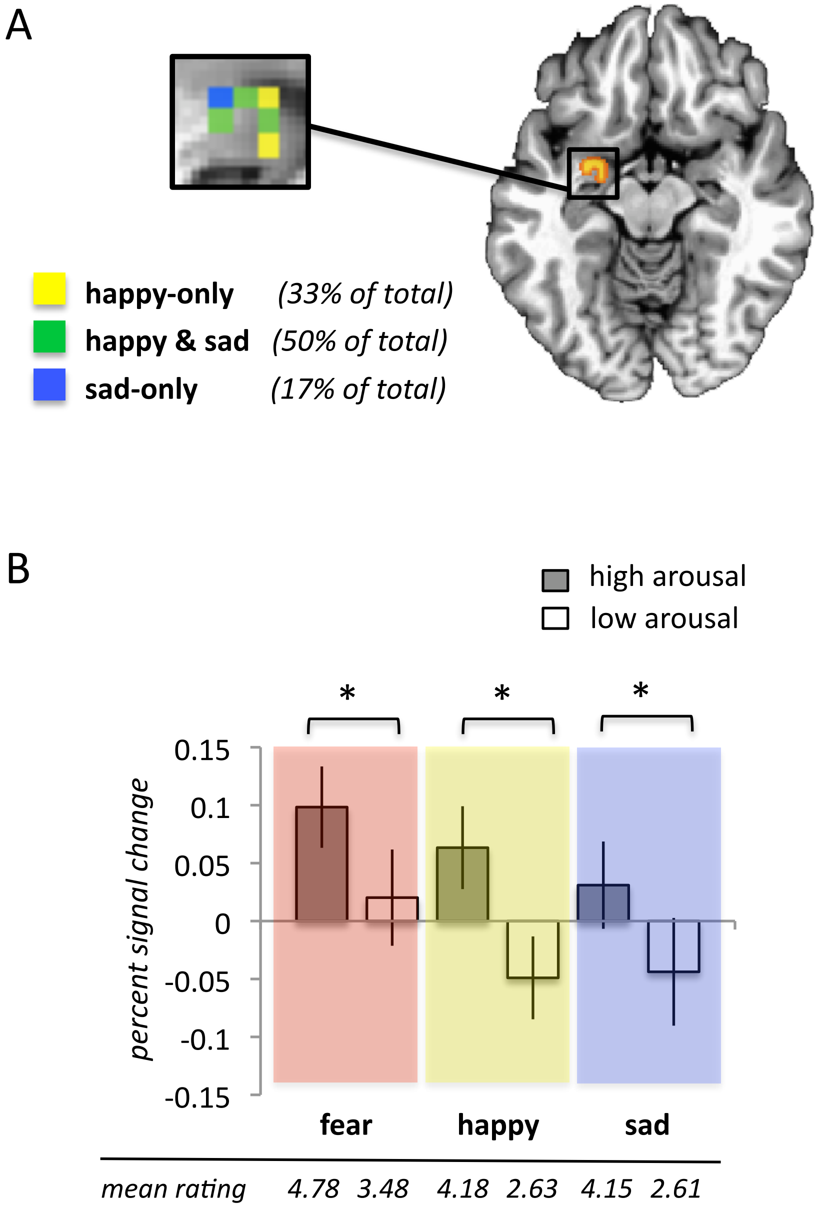


Figure 2. In panel A, the left amygdala cluster positively correlated with arousal ratings across categories is shown (*z* = -10). The magnified box shows the category results with percentages indicating proportions across the 3D cluster (not the 2D slice). In panel B, the mean betas (in percent signal change) are shown for high and low conditions of each category. The * indicates a significant difference. The mean behavioral rating for high and low conditions of each category is also shown below each bar for comparison purposes.

Appendix

Example Scenarios Ranging in Valence and Arousal for Each Emotion Studied

The core version of each scenario is indicated in italics.

FEAR

You are jogging onto the soccer field, your cleats digging into the firm ground. *You hear a booming voice welcome the crowd to the state championship*. You jump in place to shake off the restlessness in your stomach. Looking around at your team, a rushing excitement deepens your competitive fire. *You feel an energizing fear.*

You are sipping punch at a school reunion, scanning the growing crowd. *You notice your high school crush from across the room returning your gaze*. Your crush looks away and you smile to yourself in the private moment. A soft amusement begins to arise as your mind becomes lost in a familiar fantasy. *You feel a lovely fear.*

You are walking to your car alone, the city parking deck dimly lit. *You hear an explosive bang and see a man running with a pointed gun.* You quickly drop behind a car and attempt to control your shallow breathing. You try to dismiss the horrendous vision of what will happen if he finds you. *You feel a perilous fear.*

You are sitting down after lunch out, your desktop reappearing at your touch. *You notice a pressing e-mail from your boss that you forgot to address.* Taking a deep breath, you lengthen your spine in an attempt to reenergize. You slowly re-read the message with the burden of responding quickly. *You feel an inconvenient fear.*

HAPPINESS

You are performing a challenging piano solo, your fingers working the keys. *You finish the piece and receive thunderous applause as you rise.* You bend at the waist into a deep bow and sense your heart thumping rapidly. Glowing with satisfaction, you continue to feed off the crowd's energy. *You feel a proud happiness.*

You are lounging on a cushy floor pillow, opening a new magazine. *You glance up as your puppy trots over and wiggles into your lap.* As her small body relaxes, you sense both your hearts beating evenly. Tenderly petting her soft fur cultivates a lovely sense of ease. *You feel an affectionate happiness.*

You are walking down the hall, trying to get to a meeting on time. *You run into a difficult colleague and end a tense exchange with a biting remark.* Your stomach tightens the moment the last sarcastic jab escapes your lips. The cutting retort echoes poisonously in your head as your colleague sulks away. *You feel a disturbing happiness.*

You are rocking in your favorite chair, gently flipping your cell phone open and closed. *You want to share a recent promotion with your brother who is unavailable overseas.* Wishing you could call him, you close your eyes and release a held breath. You continue fiddling with your phone, a tender solitude clouding your mind. *You feel a lonely happiness.*

SADNESS

You are standing on your college quad, dressed in a smart looking cap and gown. *You listen for the graduation decree and upon hearing it fling your cap upwards.* Following energetic classmates, you sweat lightly as you march away a graduate. You catch a friend's eye and flashback to your delightfully lively freshman dorm. *You feel a spirited sadness.*

You are inching under the sheets, slowing getting settled at the late hour. *You long for a good night's sleep after spending all your waking hours working.* You sense your stiff neck relax as you rest your head on a pillow. You curl up and let go of the day, finally a moment of lovely calm. *You feel a peaceful sadness.*

You are walking into a friend's house, dropping by to return a movie. *You witness your significant other in an intimate embrace with your friend.* Your stomach is nauseated, the shocking infidelity settling into your body. Your mind is spinning trying to understand the terrible betrayal of trust. *You feel a devastating sadness.*

You are sitting at the table, spooning a heap of food on your plate. *You taste the casserole made from a new recipe and are disappointed.* Setting down your fork momentarily, you hear your stomach quietly rumbling. You look at your plate and avoid taking another bite of disagreeable blandness. *You feel a dissatisfied sadness.*

Supplementary Materials

Scenarios

Full scenarios 50-60 words in length were constructed using a five-sentence template. Written as second-person narratives, the language expressed a richly affective experience using engaging, gender-neutral prose. The first sentence established a bodily state and setting, always containing two clauses. A first clause described the state of the body (e.g., you are walking, sitting, lounging, running, awaking etc.) with the second clause elaborating the action and/or setting. The second sentence described the primary emotional stimulus or event using active construction and a present tense verb to encourage participants to fully engage with the imagined scenario. The third sentence developed the level of arousal using physiological references (e.g., heart rate, sweating, breathing, muscular or stomach states) and continued elaborating the event as it unfolded in time. The fourth sentence developed valence using pleasant (e.g., charming, lovely, delightful, etc.) and unpleasant descriptors (e.g., disagreeable, horrible, repulsive, etc.) to depict states of mind, developing the unfolding event even further. The terms pleasant and unpleasant were never used because they anchored the rating scale. The final sentence categorized the experience with the sentence "You feel a(n) [adjective] [fear, happiness, or sadness]." The adjective was used to reiterate valence (e.g., warm sadness, raging sadness, etc.). The core versions of each scenario contained the first clause of sentence one, sentence two, and sentence five.

In order to vary valence and arousal, scenarios were written to fall into four quadrants of affect space created from crossing unpleasant and pleasant valence with high and low arousal (see Appendix for examples). For each emotion category, 12 scenarios were written for each quadrant, producing 48 scenarios for each category and 144 scenarios overall. Quality audio recordings were then created for the 48 full and core versions developed for each category, spoken by an adult woman. Her prosody was tailored to the intended arousal and valence of each scenario to facilitate induction of those states in participants. Scenarios were recorded using the freely available Audacity software at a 44.1 khz sampling rate, 16 bits, and converted to MP3 format. The maximum amplitude was also normalized across scenarios to equate the volume.

To verify that the scenarios elicited the intended valence and arousal, an independent set of participants rated the 144 full scenarios, presented in a random order, for valence, arousal, or ease of experiencing the named emotion (39 participants overall; 13 each for valence, arousal, ease). Of the 48 scenarios developed for a given category, 24 were selected for the complete trials in the imaging experiment, 6 from each quadrant. Any scenario receiving a mean ease rating below 3 on a 6-pt scale ranging from 1-impossible to 6-extremely easy to imagine was excluded from the selection process. The critical scenarios selected for each category displayed the following properties. Scenarios written for high arousal quadrants were rated higher in arousal than those written for low arousal quadrants ($p < .05$). Scenarios written for pleasant quadrants were rated significantly higher on a bipolar valence scale (with the mean located on pleasant side of the scale) than scenarios written for unpleasant quadrants (with the mean located on the unpleasant side of the scale; $p < .05$). Furthermore, item valence and arousal ratings were not significantly correlated ($p > .05$).

To ensure that a range of scenarios was heard in each run of the imaging experiment, the four scenarios in complete trials for a given category were sampled equally from the four valence x arousal quadrants. In other words, the four complete trials per category in every run contained one scenario from each quadrant. Six additional scenarios were selected from each category for use in the partial trials of the experiment. One scenario from each valence x arousal quadrant, a scenario from the quadrant that would be considered typical of the emotion (e.g., high arousal, unpleasant fear; low arousal, pleasant happiness), and its opposite in affective space (e.g., low arousal, pleasant fear) made up the six scenarios selected for partial trials. Although affective space could not be sampled evenly to select six scenarios, the sampling procedure used ensured that there were equal numbers of pleasant and unpleasant valence, high and low arousal scenarios across the complete set of partial trials. A scenario from each quadrant not sampled twice for partial trials was selected for practice trials (i.e., there were two practice trials per category). During training, participants listened to the 72 scenarios that would later occur in complete trials of the imaging experiment, the 18 scenarios that would later occur in partial trials of the imaging experiment, and the 6 scenarios that would later occur in practice trials. In total, these 96 scenarios had equal numbers of fear, happiness, and sadness scenarios from each quadrant.

Training

A written script was used to instruct participants during both training sessions. In the first training session, which typically lasted about 2.5 hours, participants received an overview of the scenarios and instruction on how to imagine them. Participants were encouraged to immerse themselves in each scenario as they listened with eyes closed, and experience the scenario as if it was actually happening to them, in as much vivid detail as possible. Participants were also instructed that some scenarios might fit in more than one emotion category because emotional situations can often be construed in a number of ways. It was emphasized that the participants' task was to imagine the fear, happiness, or sadness described in the context of the scenario. In all training exercises, a few silent seconds followed each scenario so that participants could engage fully in the feeling that occurred when the situation was construed as the named emotion. Participants were told to focus on the feeling in their body and mind during the few silent seconds after hearing the emotion category (nothing was said at this point about valence or arousal). Participants practiced imagining two happiness scenarios (one being a more typical pleasant, low arousal example and the other being a more atypical unpleasant, high arousal example) during the instructional period. Before beginning any task, participants were also introduced to the relation between the full and core scenarios, and encouraged to reinstate the full scenario whenever they heard a core scenario.

Following the initial instruction period, participants listened over computer headphones to the full versions of the 96 scenarios that would later be presented in the scanner and in practice trials prior to the scan session. Upon hearing a bell two seconds after a scenario ended, participants opened their eyes and judged how personally familiar the feeling was on a 1-5 scale (1=not familiar, 3=somewhat familiar, 5=very familiar). It was explained that this rating should index how familiar the feeling of fear, happiness, or sadness induced by the scenario felt to feelings of fear, happiness, or sadness they had experienced in their life, and that the details of the scenario event did not have to match their experience exactly. At the start of all training tasks, participants performed two practice trials in which they rated the two happiness examples from the instructions. They then listened to the 96 scanner and practice scenarios in a random order (with no blocking of category).

After taking a short break, participants listened to the core versions of the same scenarios, again in a random order. During three silent seconds following the scenario, participants were encouraged to elaborate the imagined experience, immersing themselves fully into the scenario as it became enriched and developed from memory. Upon hearing a bell, they opened their eyes and rated three kinds of mental imagery using a 1-5 scale (1=none, 3=moderate, 5=high). In the following order, participants rated the vividness of internal imagery experienced within the body, external imagery experienced of the outside world, and the thought imagery experienced within their mind. The goal of the imagery ratings was to encourage participants to generate rich simulations as they listened to the core version of each scenario.

After another short break, participants were introduced to the concepts of valence and arousal, and practiced focusing on and rating their feeling state. Valence was explained as a basic sense of feeling good or bad as something is happening to you. Arousal was explained as a basic sense of how much the body and mind are being stirred as something is happening to you (i.e., the degree to which one feels awake and reactive). Low, medium, and high arousal states were described in further detail to ground out the definition of arousal (adjusted based on piloting). After listening to the definitions, participants were asked to define valence and arousal to the experimenter. If a participant provided an incorrect or vague definition, the experimenter explained the concept further and again asked the participant to generate the meaning.

Participants were next introduced to the rating scales used for valence and arousal, and asked to rate seven short sentences first for valence and then for arousal. Six sentences varying in their mean valence and arousal rating were selected from a published study (Colibazzi et al., 2010), along with a seventh neutral sentence we constructed about brushing one's teeth. If a participant's rating clearly deviated from the published means, the experimenter probed the participant for understanding of valence or arousal, again re-explaining if the participant misunderstood the concept. After rating the sentences, participants were introduced to the event sequence of complete trials. Participants practiced imagining the scenario, centering in on the valence quality of the feeling, and rating it using the appropriate scale. Valence ratings were practiced initially, followed by arousal ratings. The two happiness examples from initial instruction were used as the first practice trials. When ready, participants then engaged in a series of six complete trials. In these trials, they heard practice scenarios that had been included in the prior training tasks but were not used in the critical scanner runs.

Participants returned for the second session 24-48 hours later. At the start of this session, participants listened over headphones to the 96 full scenarios in a random order and vividly imagined the scenario as if it were happening to them. When they heard a bell sound two seconds after the scenario had finished, participants opened their eyes and rated how much they experienced 'being there' in the feeling using a 1-5 scale (1=not at all, 3=somewhat, 5=very much). It was emphasized during instruction that participants should rate their ability to imagine themselves experiencing the feeling of fear, happiness, or sadness described in the scenario. This first phase, which typically lasted about an hour, ensured that participants were reacquainted with the details of the full scenarios just before hearing the core versions in the scanner.

Participants were then instructed on and practiced the task they would perform in the scanner. The definitions of arousal and valence were refreshed, and participants were again asked to generate the meaning of each concept to confirm their understanding. Next, participants practiced using each scale with their eyes closed. Beginning with the valence scale, participants pressed the button corresponding to each point on the scale as the experimenter named them aloud moving from left to right: very unpleasant, somewhat unpleasant, neutral, somewhat pleasant, and very pleasant. The experimenter would then name a point on the scale at random, and the participant would press the corresponding button. The experimenter cycled through two random orders of the scale points in this exercise. Finally, the participants used the scale to rate the sentences from the first training session once more with eyes closed. This process was repeated with the arousal scale, which contained the following points from left to right: low, medium-low, medium, medium-high, and high. To make responses, participants rested the three middle fingers of their left and right hands on six response box buttons. Simultaneously pressing the middle two buttons with the right and left index fingers was the response for the mid-point of each scale. The middle and ring fingers were used to make responses on the left and right side of the scale. Participants practiced making responses on e-prime button boxes in the lab and used Current Designs fiber optic button boxes designed for high magnetic field environments in the scanner. In the scanner, two response boxes stabilized in a foam pad lay on the participant's legs so their fingers could rest comfortably on the six buttons.

Following practice with the scales, participants were told that they would complete one block of valence trials and one block of arousal trials in each imaging run. The cue word 'valence' or 'arousal' was repeated three times at the beginning and mid-point of the run to indicate the rating to be made during the first and second block respectively. The cueing period lasted 3 s and was always followed by 9 s of no sound so that the cueing period could be separated from the critical events in the experiment. The experimenter also informed the participant of the block order prior to each imaging run so they knew which cues to expect at the beginning and middle of the run. Debriefing after the scan session confirmed that every participant heard the cues in each run, and never lost track of which rating they were making.

Participants then practiced the different trial types, beginning with 15 s complete trials. The happiness examples from the instructions were used to practice each trial type separately before practicing blocks of intermixed trials. During complete trials, participants were instructed to immerse themselves fully as they listened to a core version of a scenario lasting no longer than 8 s. Whey they next heard a 1 s 'beep, beep, beep,' they were to continue imagining the emotion and begin centering in on the valence or arousal quality of the feeling (depending on the block), maintaining focus for three seconds. Finally, a 1 s cowbell cued participants to rate their introspective sense of valence or arousal within the next 2 s using the appropriate scale. During 9 s partial trials, participants heard a 1 s 'whoosh' sound when the 8 s scenario concluded, which signified the end of the trial. During baseline rest trials, participants cleared their mind during the 3-15 s period of no sound as they waited to hear the next scenario begin.

After practicing each trial type separately, participants practiced several short arousal and valence blocks with all trial types intermixed, as during each block of the imaging experiment. The practice blocks used the six practice scenarios not included in the imaging experiment, with each block containing equal numbers of fear, happiness, and sadness scenarios. Participants started with a short valence block that contained the valence cue followed by two complete trials, one partial trial, and no-sound baseline jitter presented in a pseudo-random order. A comparable arousal block was then performed. All six practice scenarios were heard during the first two practice blocks. In the next two practice blocks, practice scenarios heard initially in a valence block were presented in the arousal block and vice versa. The next two practice blocks, which were similar in length to the first two, were performed as a set so that the participant could become used to switching to the other rating upon hearing the cue at the beginning of the second block. In this set, the valence block came first followed immediately by the arousal block. In the last practice set, an arousal block came first followed immediately by a valence block. These blocks were longer so that the participant could gain a better sense of what the imaging runs would be like, containing five complete trials and one partial trial. The six practice scenarios were repeated in each block of the final practice set. Participants were informed that the repeated scenarios were for practice purposes and that scenarios would not repeat in this manner during the imaging runs. They were also informed that although scanner blocks would be longer, the task itself would be exactly the same with all trial types occurring in a random, unpredictable order.

Image Analysis

Anatomical MPRAGE scans were collected at the beginning and end of the session (192 sagittal slices, TR=2300 ms, TE=4ms, flip angle = 8°, FOV = 256 mm, matrix = 256, bandwidth = 130 Hz/Px, voxel size = 1 mm × 1 mm × 1 mm). In each intervening 9 min 3 s functional run, 163 T2*-weighted echo planar image volumes depicting BOLD contrast were collected using a Siemens 12-channel head coil and parallel imaging with an iPAT acceleration factor of 2 (56 2 mm axial slices, TR = 3000 ms, TE = 30 ms, flip angle = 90°, bandwidth = 2442 Hz/Px, FOV = 220 mm, matrix = 64, voxel size = 3.44 mm × 3.44 mm × 2 mm). These parameters were selected to minimize susceptibility artifact in OFC and amygdala while maintaining satisfactory signal-to-noise ratios.

The second anatomical scan was registered to the first anatomical scan and the two datasets averaged to produce a single high-quality anatomical. Slice-time correction was performed on the functional volumes followed by motion correction and transformation to Talairach space, which were performed in a single step to reduce error that occurs when the functional data are independently warped multiple times. The transformation matrix for motion correction was generated in which all functional volumes were registered to a volume near the beginning of the first run. To generate the Talairach transformation matrix, the averaged anatomical was first skull-stripped and aligned to the same functional volume used as the registration base for motion correction. The anatomical was then transformed to Talairach space using an automated procedure employing the TT_N27 template (also known as the Colin brain, an averaged dataset from one person scanned 27 times). The matrices generated from the Talairach transformation of the anatomical dataset were concatenated with the motion correction matrix and applied in one step to the functional volumes. At this point, the voxel dimensions of the functional volumes were also resampled from 3.44 mm × 3.44 mm × 2 mm to 3 x 3 x 3 mm. The functional data were next smoothed using an isotropic 6 mm full-width-half-maximum Gaussian kernel. Finally, the signal intensities in each volume were divided by the mean signal value for the respective run and multiplied by 100 to produce percent signal change from the run mean. The first four volumes of each run, which often contained more outliers, were not included in the calculation of the run mean because these volumes were later censored in regression analyses. All later analyses were performed on the percent signal change data. In all later regressions, the six regressors obtained from motion correction during preprocessing were included to remove any residual signal changes correlated with movement (translation in the X, Y, and Z planes; rotation around the X, Y, Z axes). Scanner drift was removed by finding the best-fitting polynomial function correlated with time in the preprocessed time course data.

Anatomical Regions of Interest

Anatomical masks for bilateral amygdalae, lateral OFC, and medial OFC were created using the AFNI CA_N27_ML atlas (Eickoff et al., 2005). The lateral OFC mask was defined as the left and right middle orbital gyrus and inferior frontal gyrus par orbitalis. The medial OFC mask was defined as the mid orbital gyrus, rectal gyrus, and superior orbital gyrus.

Additional Arousal Analysis

An additional regression analysis was run at the subject level with no amplitude modulation. In this regression, the arousal focus-rate conditions for each category were split into a low and high arousal condition, creating six arousal focus-rate conditions (i.e., fear-low, fear-high, happiness-low, happiness-high, sadness-low, sadness-high). The high and low conditions for each category contained equal numbers of trials and were defined prior to the imaging experiment using arousal ratings from an independent set of participants (see scenarios section). The ratings of participants in the imaging experiment were consistent with the earlier ratings; the high condition was rated as significantly higher in arousal than the low condition in each category ($p < .001$). The scenarios were modeled as six conditions representing category and rating block (i.e., fear-arousal, fear-valence, etc.). The valence focus-rate conditions were modeled as six conditions, with each category split into pleasant and unpleasant conditions (e.g., fear-pleasant, fear-unpleasant, etc.). As for all other regressions, conditions were modeled using Gamma functions convolved with a boxcar function that represented the duration of events in the condition

References

Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage, 25*, 1325-1335.

Table S1. Descriptive statistics for valence and arousal ratings and the correlation between valence and arousal ratings for each emotion category.

	Valence				Arousal				Correlation
Category	M	SE	Min	Max	M	SE	Min	Max	r
fear	2.81	0.05	1.00	5.00	4.13	0.06	1.88	5.00	0.01
happiness	3.57	0.04	1.38	5.00	3.40	0.05	1.19	5.00	-0.01
sadness	2.48	0.04	1.00	4.94	3.38	0.05	1.13	4.94	-0.30

Note. The standard error of the mean (SE) was computed on participant means. Min and max represent the mean minimum and maximum computed across individual participant minimums and maximums. Similarly, the correlation reported is the mean of individual correlations. Seven participants showed a moderate, significant correlation between valence and arousal for sadness. One participant showed a moderate, significant correlation for fear.

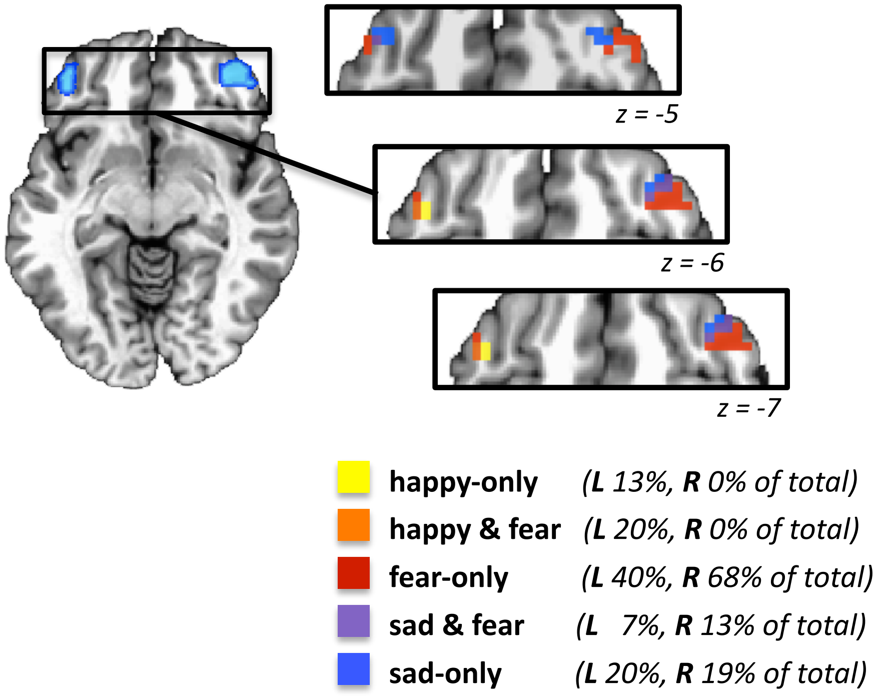


Figure S1. The lateral OFC clusters that correlated negatively with arousal ratings across categories is shown on the left (*z* = -6). The magnified boxes to the right show the category results with percentages indicating proportions across the 3D cluster (not the 2D slices) in the left (L) and right (R) cluster. The three boxes show variations in the category results at different axial slices, especially in the left hemisphere.

Running head: ATYPICAL EMOTIONS

Taking the Plunge:
An Investigation of Pleasant Fear and Other Atypical Emotions

Christine D. Wilson-Mendenhall
Emory University

Lisa Feldman Barrett
Northeastern University

Lawrence W. Barsalou
Emory University

Abstract

Distinct large-scale networks have been identified in the brain that appear to underlie a ‘default mode’ of internal socio-emotional simulation, and task-oriented attention operations. In this experiment, we examined if familiar emotions that have atypical valence (e.g., pleasant fear) are processed differently in these networks than familiar emotions that have typical valence (e.g., unpleasant fear). In a novel fMRI paradigm, participants immersed themselves in an atypical or typical fear, sadness, or happiness emotion scenario, and then focused on the valence of the feeling and rated it. The experiment was designed so that the initial scenario immersion event could be separated from the later valence focus event. Our first prediction was that greater activation in the default network would be observed during scenario immersion and valence focus for atypical emotions (compared to typical), because the atypical emotions would require more interpretive social processing of the scenario. Our second prediction was that attention and control networks would be more active for atypical emotions when participants focused on valence, the aspect of the emotion that made it atypical. The results supported our predictions, suggesting that default and attention networks support immersing oneself in an atypical emotion and focusing on its complex valence. Future research is necessary to understand the properties of atypical emotions (e.g., complexity, ambiguity, etc.) that underlie the heightened activations in these networks when emotions are atypical.

Taking the Plunge: An Investigation of Pleasant Fear and Other Atypical Emotions

Social cognition is the fundamental ability to infer others' intentions, feelings, and thoughts, as well as one's own intentions, feelings, and thoughts (Adolphs, 2009). Interpretation and inference, at many levels, are basic principles of the human mind. The brain adaptively uses prior experience to dynamically interpret ongoing internal and external sensations and mental states, making predictions (inferences) that guide actions and interactions (Barsalou, 1999, 2003, 2009; Barrett, 2009). This process of conceptualization, which often proceeds outside of awareness, makes our world meaningful. Without it, we would not be able to infer, for example, that a steaming cup of coffee is hot, tastes better with sugar, or that ingesting it might wake us up, at least for a little while. We also wouldn't be able to infer that the new employee at the local coffee means well but is inefficient, that the person in line talking on their cell phone is self-involved, or that we, ourselves, are reacting in an overly-sensitive manner.

Although social cognition and emotion have, to some extent, been considered separate research domains, neuroimaging research implicates similar brain regions (Olsson & Ochsner, 2007; Kober et al., 2008). Consistent with these findings, a new constructivist view of emotion has emerged in which grounded conceptualization, as a fundamental cognitive process, is also at the heart of emotion (Barrett, 2006; 2009). Unlike basic emotion or appraisal emotion views, this view suggests that conceptualizing a situation in a grounded manner causes it to be experienced as an emotion (Barrett, Barsalou, Lindquist, & Wilson-Mendenhall, in prep; Wilson-Mendenhall, Barrett, Simmons, & Barsalou, submitted).

Social Cognition, Emotion, and a Default Mode of Processing

Recent emphasis on the critical role of grounded conceptualization in emotion suggests a new way of thinking about our rich emotional lives. Similarly, the discovery of several global brain networks has initiated new thinking about the way cognition works (Bressler & Menon, 2010). Brain regions implicated in social cognition and emotions are now being interpreted in terms of their dynamic interactions in large-scale networks. In particular, the so-called default network appears to play a central role in socio-emotional processing.

The default mode was initially discovered as a network of regions active during experimental blocks when participants were ‘resting,’ not engaging in a cognitive task (Gusnard & Raichle, 2001). This network includes two midline hubs in medial prefrontal cortex and posterior cingulate, as well as inferior parietal regions and to lesser extent the hippocampal formation and lateral temporal cortex (Buckner, Andrews-Hanna, & Schacter, 2008). Named the default mode for it task-negative profile, this network was initially thought to be a passive monitoring system (Shulman et al., 1997). Recent evidence, however, suggests that this system underlies active interpretation of the social world, and that it is implicated in many facets of social cognition, including autobiographical memory, envisioning the future, theory of mind, and moral decision-making (Buckner et al., 2008; Harrison et al., 2008). Furthermore, this network has been associated with spontaneous mind wandering during undemanding external tasks (Christoff, Gordon, Smallwood, Smoth, & Schooler, 2009).

Default mode activity in various forms of social cognition has led to proposals that it is generally involved in inner-oriented (Golland, Golland, Bentin, & Malach, 2008) and self-related processing (Northoff & Panksepp, 2008), used for flexible self-projection (Buckner & Carroll, 2007) and for generating predictions from contextual representations (Bar, 2007). These accounts share the idea that this large-scale brain network is dedicated to interpreting (often social) situations to guide behavior, which is a dramatic departure from the idea that the brain is primarily reactive, responding in a bottom-up manner to sensory stimuli (Raichle, 2010). Notably, the idea of reactive responding has permeated much theorizing on emotion.

Attention, Control, and Saliency

Using functional connectivity, data reduction techniques, and clustering algorithms, researchers have distinguished the default network from lateral fronto-parietal networks thought to subserve goal-directed attention and cognitive control, and also from primary and secondary sensory-motor cortices (Golland et al., 2007; Golland et al., 2008; Fox et al., 2005). Attention is a multi-faceted construct that involves operations for maintaining an alert state, orienting to sensory events, and regulating thoughts and behaviors (Posner, 2008). These operations have primarily been studied in externally oriented tasks such as detecting visual targets. Interestingly, the networks identified in visual attention tasks include regions traditionally associated with affect, such as ventral fronto-insular cortex (Corbetta, Patel, & Shulman, 2008). Furthermore, frontal regions in these networks have been implicated in the cognitive control of emotion, suggesting that attention networks also operate on internal events (Ochsner & Gross, 2005).

The executive functions of dorsolateral prefrontal regions to select and manipulate information and to maintain goals within working memory are well established (Duncan & Owen, 2000; Curtis & D'Esposito, 2003). The roles of ventrolateral prefrontal regions and anterior cingulate regions are less clear, and are especially interesting because of their involvement in both cognitive and emotion tasks. Recently, it has been suggested that dorsal anterior cingulate and orbital fronto-insular cortices form a salience network that is distinct from dorsal fronto-parietal executive control networks (Seeley et al., 2007). In this work, salience is described as a process in which highly processed sensory information is integrated with visceral, autonomic, and hedonic information to inform decision-making. The primary function of this network appears to be coding important internal or extra-personal stimuli for further processing and to initiate control signals, integrating bottom-up attention switching with top-down control (Menon & Uddin, 2010). Consistent with this idea, a recent review revealed that anterior insula and anterior cingulate were active in a wide range of tasks in which the only common denominator appeared to be awareness (Craig, 2009).

Typicality of Emotions

The framework of large-scale networks offers a new way of thinking about socio-emotional processing. In addition to sensory-motor processing of the external world, large-scale systems in the brain actively interpret an inherently social world, direct attention, and exert cognitive control to guide behavior. The purpose of the experiment reported here was to investigate the involvement of these networks in a unique kind of socio-emotional processing, specifically, experiencing atypical instances of emotion categories. Until recently, research has primarily focused on identifying the neural substrates that distinguish various types of social cognition (e.g., beliefs, intentions, emotions). We know relatively little about how the brain processes instances within a type of social cognition that vary in frequency, ambiguity, or typicality (Jenkins & Mitchell, 2010).

A long-standing idea is that some instances of a category are more typical than others (Barsalou, 1985, 1987; Rosch & Mervis, 1975). Typicality refers to how good of an example a particular instance is of its category, with some instances being more typical than others. The resultant graded structure is a central and ubiquitous property of categories, including emotion categories (Dube & Le Bel, 2003; Fehr & Russell, 1984; Fehr & Russell, 1991; Russell, 1991). In Fehr and Russell (1991), for example, maternal love was rated as a more typical example of the emotion category love than was patriotic love or puppy love.

One clear way that instances within an emotion category appear to vary in their typicality is through valence. As described in dimensional approaches to emotion, emotion categories can be organized around a circumplex according to their typical valence and arousal (e.g., Russell, 2003). For example, fear is typically unpleasant, and many fear situations are indeed experienced as unpleasant (e.g., being diagnosed with a major illness, encountering a snake, or losing control of a car). Interestingly, however, there are also familiar situations in which fear is predominantly a pleasant experience (e.g., riding a rollercoaster, performing in front of a crowd, engaging in a competitive sport). Intuitively, these situations appear less typical of the emotion category fear than the unpleasant situations.

The phenomenon of a mixed emotion often appears related to experiencing an emotion with atypical valence (e.g., the pleasant fear of thrill-seeking, the pleasant sadness of nostalgia). Several studies have shown that people report experiencing complex 'mixed' emotions in certain situations such as graduating or gambling (Larsen, McGraw, & Cacioppo, 2001; Larsen, McGraw, Mellers, & Cacioppo, 2001; Miyamoto, Uchida, & Ellsworth, 2010). In these cases, it appears that multiple conceptualizations of the situation exist. In a similar fashion, atypical instances of an emotion category may also contain situational elements that are flexibly conceptualized in multiple ways.

Methodological Overview and Predictions

The goal of this experiment was to examine if atypical emotion states are processed differently in the brain than typical emotion states. We defined atypical emotions as fear, sadness, or happiness in which valence was atypical for the category (i.e., pleasant fear, pleasant sadness, unpleasant happiness). We compared the atypical emotions to fear, sadness, or happiness in which valence was typical for the category (i.e., unpleasant fear, unpleasant sadness, pleasant happiness). Table 1 presents an example of an atypical and typical emotion scenario for each category.

[Insert Table 1 about here]

In a novel fMRI task, participants engaged in two distinct events, one immediately following the other. First, in scenario immersion events, participants listened to and immersed themselves in a fear, happiness, or sadness scenario.

Second, in subsequent valence focus events, participants focused on and then rated the valence of the induced emotion. As described shortly, this design allowed us to separate these consecutive events and examine neutral activations in each event. Thus, we examined relative neural differences in atypical and typical emotions during immersion in the emotion, and during focus on the affective feature (valence) that defined it as atypical or typical.

To facilitate immersion, participants became familiar with fear, happiness, and sadness scenarios in training sessions prior to fMRI scanning. In two separate training sessions before the critical scans, participants listened to the scenarios and rated the respective scenario for familiarity, imagery, and their ability to 'be there' immersed in the scenario. As participants listened to the scenario, they were instructed to immerse themselves in the situation it described as deeply as possible. As Table 1 illustrates, the scenarios were written in a second person narrative and contained various details designed to induce immersion.

The training versions of the scenarios were longer in duration than was optimal for use in a scanner. For this reason, shorter core versions only contained critical components of the longer full versions (see Table 1 for examples). During training, participants were told about the relation between the full and core version of each scenario, and they practiced generating the full version while listening to the core version. This ensured that participants were prepared to imagine the full version of each scenario as they listened to the core version later in the scanner.

On critical trials in the scanner, participants immersed themselves in equal numbers of pleasant and unpleasant fear, happiness, and sadness scenarios. In complete trials, two critical events occurred on each trial: scenario immersion followed by valence focus. During scenario immersion, participants listened to a scenario and immersed themselves in the experience of it. Following the scenario, participants were instructed to focus on the valence of the emotion in the scenario and then rate it. In partial trials, participants only engaged in scenario immersion and did not judge valence subsequently. Including both complete and partial trials allowed us to separate brain activity during scenario immersion and valence focus, and thus to examine these two time periods separately.[5]

Of primary interest was comparing the brain areas active for atypical (i.e., pleasant fear, pleasant sadness, unpleasant happiness) vs. typical (i.e., unpleasant fear, unpleasant sadness, pleasant happiness) emotions. In a series of analyses, we examined activity for atypical vs. typical emotion during both scenario immersion events and valence focus events. Our first prediction was that greater activation in the default network would be observed during scenario immersion and valence focus for atypical emotion states (compared to typical), because the atypical emotions would require more interpretive social processing of the scenario. Our second prediction was that attention and control networks would also be more active for atypical emotions than for typical emotions when participants focused on valence, the aspect of the emotion that made it atypical.

---

[5] This paradigm was initially developed to examine brain regions correlating with valence and arousal in each category, which are reported in Wilson-Mendenhall et al. (in prep).

Methods

Design & Participants

The experiment contained two training sessions and an fMRI scan session. The first training session occurred 24 to 48 hours before the second training session, followed immediately by the scan session. The fMRI design was constructed to examine six critical conditions created by the factors of emotion category (fear, sadness, happiness) and valence (pleasant, unpleasant). Crossing these factors created the conditions that would later make up the atypical (pleasant fear, pleasant sadness, unpleasant happiness) and typical emotion conditions (unpleasant fear, unpleasant sadness, pleasant happiness).

Two trial types existed in each condition. In 72 *complete* trials, participants immersed themselves in a fear, happiness, or sadness scenario, and then focused on and rated the valence quality of the feeling. In 18 *partial* trials, participants only imagined a scenario. Partial trials were included so scenario events could be mathematically separated from subsequent focus-rate events when complete trials were analyzed (Ollinger, Corbetta, & Shulman, 2001; Ollinger, Shulman, & Corbetta, 2001; Ruge, Goschke, & Braver, 2009). This design was critical for separating neural activity during scenario immersion from activity during valence focus. As necessary for this type of catch trial design, partial trials were unpredictable and accounted for 20% of all trials.

Six imaging runs consisted of one valence block and one arousal block, with block order counterbalanced across runs and trials as events within blocks. Arousal blocks were analyzed elsewhere (Wilson-Mendenhall, Barrett, & Barsalou, in prep) and will only be mentioned when critical for describing methods. Within each valence block, four complete trials and one partial trial for each category were presented amidst jittered no-sound baseline periods (ranging from 3-15 s in increments of 3; average ISI = 6.3 s) in a pseudo-random order optimized for deconvolution analysis using optseq2 software. The four complete trials per category in a block consisted of two pleasant and two unpleasant scenarios; thus there were a total of 12 pleasant and 12 unpleasant complete trials in the experiment for each category.

Across the six runs in the experiment, each fear, sadness, and happiness scenario was presented twice, once in an arousal block and once in a valence block. To create the first version of the experiment, the 24 critical scenarios presented in complete trials and the 6 scenarios presented in partial trials for each category were randomly assigned to the valence blocks of the six runs. Scenarios in the valence block of run one, two, and three were repeated in arousal block of run four, five, and six respectively. Scenarios in the valence block of run four, five, and six were first encountered in the arousal block of run one, two, and three. To control for repetition order, a second version was created in which scenarios in valence and arousal blocks were flipped (i.e., scenarios initially rated first for valence were rated first for arousal in the second version and vice versa). Finally, two additional versions were created paralleling the first two in which the order of the runs was simply reversed to control for general stimulus order effects.

Sixteen right-handed, native English speakers from the Emory community, ranging in age from 19-30 (8 female), were randomly assigned to one of the four versions. Each participant received $100 in compensation.

Materials

Scenarios were written and recorded to induce states of fear, sadness, and happiness. A full and core form of each scenario was developed, the core being a subset of the full form (see Table 1). The full form provided a richly detailed and affectively compelling description of a fear, sadness, or happiness episode. The core form served to minimize presentation time in the scanner so the number of trials necessary for a sufficiently powerful design could be implemented. In both forms, scenarios were explicitly categorized as fear, sadness, or happiness to avoid ambiguity.

In each category, half of the scenarios had typical valence (e.g., unpleasant fear and sadness; pleasant happiness), and the other half had atypical valence (pleasant fear and sadness; unpleasant happiness). To verify that scenarios written to be pleasant or unpleasant were experienced as such, valence ratings were collected from an independent set of 13 participants. In each category, scenarios written to be pleasant were rated significantly higher on a bipolar valence scale (with the mean located on the pleasant side of the scale) than scenarios written to be unpleasant (with the mean located on the unpleasant side of the scale; $p < .05$). Both pleasant and unpleasant scenarios systematically varied in their arousal properties. More details on the construction and selection of scenarios, as well as their arousal properties, can be found in Wilson-Mendenhall et al. (in prep).

Procedure

In the first training session, participants provided informed consent and were screened for problems that could arise in the MRI environment. Participants had no history of psychiatric illness and were not taking psychotropic medication. The purpose of the first session was to familiarize participants with the scenarios by having them actively practice a) vividly imagining the full versions of the scenarios they would hear later in the scanner or during practice trials b) reinstating the rich imagery of each full scenario upon hearing the core version c) focusing on and rating the valence or arousal quality of the feeling state induced by a scenario. Before performing any task, participants received an overview of the scenarios and were encouraged to imagine each scenario from a first-person perspective as vividly as possible. They were also informed of the use of full and core scenario versions, and the relation between them. Participants then listened to the full versions of the scenarios, immersing themselves with eyes closed, and rated their personal familiarity with each imagined emotion. After a short break, they listened to the core versions of the same scenarios, reinstated imagined details from the full versions, and then rated the internal, external, and thought imagery experienced (further encouraging immersion in the imagined scenarios). Finally, participants practiced focusing on and rating the arousal or valence quality of the emotion induced in several practice scenarios.

When participants returned to the lab 24-48 hours later, they began the second session by listening to and vividly imagining each full scenario again. Participants provided one rating of how much they experienced being immersed in the feeling of fear, happiness, or sadness described in the scenario. Imagining the full versions in the second session ensured that participants were reacquainted with all the details just prior to hearing the core versions in the scanner. Participants were then instructed on and practiced the task that they would perform in the scanner with scenarios not used in imaging runs. Participants were first informed that they would complete one block of valence trials and one block of arousal trials in each imaging run and that the cue word 'valence' or 'arousal' would be repeated three times at the beginning of each block. They then practiced the different trial types, beginning with 15 s complete trials.

During complete trials, participants were instructed to immerse themselves fully as they listened to a core version of a scenario lasting no longer than 8 s. When they next heard a 1 s 'beep, beep, beep,' they were to continue imagining the emotion and to begin centering in on the valence or arousal quality of the feeling (depending on the block), maintaining focus for 3 s. Finally, a 1 s cowbell cued participants to rate their introspective sense of valence or arousal within the next 2 s using the appropriate scale. At this point, participants had received much practice using the 5-point valence and arousal scales with their eyes closed. The points on the valence scale from left to right were very unpleasant, somewhat unpleasant, neutral, somewhat pleasant, and very pleasant; the points on the arousal scale were low, medium-low, medium, medium-high, and high.

During 9 s partial trials, participants heard a 1 s 'whoosh' sound when the 8 s scenario concluded, which signified the end of the trial. During the silent baseline periods of rest (3-15 s), participants cleared their mind as they waited to hear the next scenario begin. After practicing each trial type separately, participants engaged in several short arousal and valence blocks with all trial types intermixed as in each later block of the imaging experiment. Participants were then informed that, although scanner blocks would be longer, the task itself would be exactly the same with all trial types occurring in a random, unpredictable order. A more detailed description of the training session procedures can be found in Wilson-Mendenhall et al. (in prep).

Following practice, the experimenter and participant walked to the MRI facility. Once the participant was situated comfortably in the scanner, an initial anatomical scan was collected. The participant was then reminded of the task using brief instructions and of the valence and arousal scales by pressing the appropriate button as the experimenter named each point on the scale. When the participant was ready, the experimenter initiated the first functional run and then continued with the next five runs, pausing for a short break between runs. A second anatomical scan was collected last. Total time spent in the scanner was a little over an hour.

Imaging and Analysis

Images were collected at the Emory Biomedical Imaging Technology Center on a 3T Siemens Trio scanner. Anatomical MPRAGE scans were collected at the beginning and end of the session (192 sagittal slices, TR=2300 ms, TE=4ms, flip angle = 8°, FOV = 256 mm, matrix = 256, bandwidth = 130 Hz/Px, voxel size = 1 mm × 1 mm × 1 mm). In each intervening 9 min 3 s functional run, 163 T2*-weighted echo planar image volumes depicting BOLD contrast were collected using a Siemens 12-channel head coil and parallel imaging with an iPAT acceleration factor of 2 (56 2 mm axial slices, TR = 3000 ms, TE = 30 ms, flip angle = 90°, bandwidth = 2442 Hz/Px, FOV = 220 mm, matrix = 64, voxel size = 3.44 mm × 3.44 mm × 2 mm). These parameters were selected to minimize susceptibility artifact in orbitofrontal cortex, amygdala, and the temporal poles while maintaining satisfactory signal-to-noise ratios.

All preprocessing and statistical analyses were conducted in AFNI (Cox, 1996). The second anatomical scan was registered to the first anatomical scan and the two datasets averaged to produce a single high-quality anatomical. Slice-time correction was performed on the functional volumes followed by motion correction and transformation to Talairach space, which were performed in a single step to reduce error that occurs when the functional data are independently warped multiple times. The transformation matrix for motion correction was generated in which all functional volumes were registered to a volume near the beginning of the first run. To generate the Talairach transformation matrix, the averaged anatomical was skull-stripped and aligned to the same functional volume used as the registration base for motion correction. The anatomical was then transformed to Talairach space using an automated procedure employing the TT_N27 template (also known as the Colin brain, an averaged dataset from one person scanned 27 times). The matrices generated from the Talairach transformation of the anatomical dataset were concatenated with the motion correction matrix and applied in one step to the functional volumes. At this point, the voxel dimensions of the functional volumes were also resampled from 3.44 mm × 3.44 mm × 2 mm to 3 x 3 x 3 mm.

The functional data were next smoothed using an isotropic 6 mm full-width-half-maximum Gaussian kernel. Finally, the signal intensities in each volume were divided by the mean signal value for the respective run and multiplied by 100 to produce percent signal change from the run mean. The first four volumes of each run, which often contained more outliers, were not included in the calculation of the run mean because these volumes were later censored in regression analyses. All later analyses were performed on the percent signal change data.

Regression analyses were computed at the individual subject level in which the hemodynamic response was modeled using Gamma variate functions. The trial onset times for 25 conditions were specified in this analysis. For valence blocks, the onsets of the 15 pleasant and 15 unpleasant scenario immersion events and the onsets of the 12 pleasant and 12 unpleasant valence focus events were specified for each category, creating twelve conditions. As described further below, scenario immersion events from the complete and partial trials were modeled as a single condition, which is why these conditions contained 15 trial onsets and the valence focus conditions contained 12 trial onsets. For arousal blocks, twelve parallel conditions were created for high and low arousal scenario immersion and arousal focus events in each category. Finally, a cue condition specified the onsets for the cues beginning each valence and arousal block. Only analyses of the valence conditions are reported here; analyses of the arousal conditions can be found in Wilson-Mendenhall et al. (in prep).

Scenario immersion events included the 9 s during which participants heard the scenario and the brief auditory cue in the final second of this period for both complete and partial trials. Modeling the scenarios in complete and partial trials as a single condition allowed for the mathematical separation of the scenario immersion period from the valence focus period in complete trials. The valence focus events included the 6 s during which participants focused on and rated the valence quality of the feeling. Because the scenario immersion and focus conditions were each several seconds long, the gamma function used to model the hemodynamic response was convolved with a boxcar function reflecting the duration of the event. Six regressors obtained from motion correction during preprocessing were also included to remove any residual signal changes correlated with movement, and scanner drift was removed by finding the best-fitting polynomial function correlated with time in the preprocessed time course data.

The betas resulting from the each participant's regression analyses were then entered into a second-level random effects analysis. In this group analysis, two key contrasts were computed. In the first contrast analysis, the typical scenario immersion conditions (unpleasant fear, unpleasant sadness, pleasant happiness) were compared to the atypical scenario immersion conditions (pleasant fear, pleasant sadness, unpleasant happiness). In the second contrast analysis, the typical valence focus conditions (unpleasant fear, unpleasant sadness, pleasant happiness) were compared to the atypical valence focus conditions (pleasant fear, pleasant sadness, unpleasant happiness). A voxel-wise threshold of $p < .005$ was used in conjunction with a 36-voxel extent threshold determined by AFNI Clustsim to produce a corrected threshold of $p < .05$.

Finally, the brain circuits involved in making motor responses during valence focus were removed from the contrasts described above. To remove these activations, a second regression analysis was computed in which the trial onsets for six conditions were modeled using Gamma variate functions. Scenario immersion events were modeled in two conditions, one for valence and one for arousal blocks. Valence and arousal focus events, however, were split into conditions based on the hand used to make the response. Participants responded with a right-hand finger, left-hand finger, or fingers from both hands (mid-point of the scale) to rate valence or arousal. Across valence and arousal focus events, conditions were created for left-hand responses, right-hand responses, and both-hand responses. Missing responses were included in the both-hand response condition because there were too few trials to create a separate condition. The sixth and final condition included the cues beginning each block. The betas resulting from each participant's regression analyses were then entered into a second-level random effects analysis. At the group level, a contrast was computed that compared focus left-hand responses to focus right-hand responses (threshold $p < .005$; 15 contiguous voxels). We then created a mask that included regions showing greater activity for left than right-handed responses and those regions showing greater activity for right- than left-handed responses. Activations in bilateral motor and somatosensory cortex, middle cingulate, putamen, thalamus, and cerebellum were removed from all group contrasts using this mask.

Results

Atypical > Typical Scenario Immersion

The results are first reported here for scenario immersion, with the results for the valence focus periods being reported later. As Figure 1 and Table 2 illustrate, several regions were more active when participants immersed themselves in atypical fear, sadness, and happiness emotions than when they immersed themselves in typical emotions. In support of our first prediction, we observed activity in dorsomedial prefrontal cortex (dmPFC) and posterior cingulate, default network regions associated with inner-oriented processing, context, and prediction (Bar, 2007; Buckner & Carroll, 2007; Golland et al., 2008). Also more active for immersion in the atypical emotions were lateral frontal regions associated with working memory and cognitive control, including and bilateral middle frontal gyrus (MFG) and left inferior frontal gyrus (IFG; Curtis & D'Esposito, 2003; Duncan & Owen, 2000; Miller & Cohen, 2001; Thompson-Schill, Bedny, & Goldberg, 2005), and supplementary motor area (SMA), associated with planning and sequencing internally generated actions (Nachev, Kennard, & Husain, 2008; Picard & Strick, 1996). In lateral parietal cortex, bilateral angular gyrus was also active, in a region often referred to as temporoparietal junction (TPJ). It has been suggested that the TPJ region is necessary for inferring the goals of others, and is generally important for shifting out of the present situation (such as imagining social situations; Van Overwalle, 2009). Finally, left temporal pole was also more active during atypical scenario immersion. Recent evidence suggests that the temporal poles represent domain-specific social knowledge that may be specific to the processing of individuals (Damasio et al., 2004; Drane et al., 2008; Simmons & Martin, 2009;

Tranel, 2006). No regions were found to display the opposite pattern, that is, more activity during typical scenario immersion.

[Insert Figure 1 about here]

Atypical > Typical Valence Focus

In this section, we report the results for the valence focus period that followed immersion in the scenarios. Many of the same regions were more active when participants focused on the valence of atypical relative to typical emotions. As Figure 1 and Table 2 illustrate, activity in frontal and parietal regions was much more extensive than related activations during the scenario period. In support of our first prediction, default network regions were more active for atypical relative to typical valence focus, including ventral and dorsal aspects of medial prefrontal cortex and posterior cingulate, extending into precuneus. In support of our second prediction, robust activation was observed in lateral fronto-parietal networks thought to support attention and cognitive control processes (Curtis & D'Esposito, 2003; Corbetta, Patel, & Shulman, 2008; Duncan & Owen, 2000; Fox et al., 2005). In lateral prefrontal cortex, activation spread bilaterally from superior frontal gyrus down through the middle and inferior frontal gyri into the anterior insula and orbitofrontal cortex, extending anteriorly from the orbital sector of the inferior frontal gyrus into middle orbital gyrus. In lateral parietal cortex, bilateral activation occurred in angular gyrus and in inferior and superior parietal cortex. Whereas activity in right lateral parietal was largely centered in the angular gyrus, activity in the left hemisphere was more evenly distributed across lateral parietal regions. Both regions in the proposed salience network, bilateral anterior insula and dorsal anterior cingulate, were also more active for atypical valence focus (Seeley et al., 2007).

In addition to default and lateral fronto-parietal activations, bilateral regions of middle temporal gyrus nearest to auditory regions, along with bilateral regions of the basal ganglia, thalamus, and brainstem, were also more active during atypical valence focus. Again, no region was more active during typical valence focus relative to atypical valence focus.

[Insert Table 2 about here]

Conjunction of Scenario Immersion and Valence Focus Contrasts

We performed a conjunction analysis to establish the unique and overlapping activations in the scenario immersion and valence focus contrasts. The circled areas in Figure 2 indicate overlap in the default network (dmPFC, posterior cingulate), areas associated with internally generated action (SMA), areas representing alternative situations (bilateral TPJ), and areas associated with working memory and cognitive control (left IFG, bilateral MFG). Figure 2 also illustrates that activity during valence focus was generally much more extensive in regions of overlap.

Only one region was active for the atypical emotions in the scenario immersion contrast that was not also present in the valence focus contrast. This area was the left temporal pole, which has been associated with domain-specific social knowledge and the processing of individuals.

In contrast, several regions associated with attention, executive function, and salience were observed for atypical emotions in the valence focus contrast that were not present in the scenario immersion contrast. These regions included ventrolateral prefrontal cortex (bilateral orbital frontal cortex and right inferior frontal gyrus), anterior insula, dorsal anterior cingulate, and bilateral inferior and superior parietal cortex. Ventromedial prefrontal cortex (part of the default network), middle temporal gyrus, basal ganglia, thalamus, and brainstem regions were also unique to the valence focus contrast.

[Insert Figure 2 about here]

Contrasts Analyses of Scenario Immersion and Valence Focus Conditions

The conjunction just discussed identified overlapping and unique regions in the scenario immersion and valence focus atypical vs. typical contrasts, but it did not directly compare activity occurring during these two events. To further understand neural changes occurring when participants shifted from scenario immersion to valence focus in typical and atypical conditions, we computed a series of contrast analyses across the typical and atypical conditions. For these analyses, we first extracted each participant's condition betas from all regions that were significantly more active during atypical valence focus than during typical valence focus (see Table 2). We chose to initially interrogate regions identified in the valence focus contrast because virtually all of the clusters identified in the scenario immersion contrast were subsets of with the widespread activations observed during valence focus. We later used the same procedure to examine clusters identified in the scenario immersion contrast, which largely showed the same patterns. We discuss these results after presenting the clusters from the valence focus contrast.

Because the significant clusters identified in the valence focus contrast often spanned multiple cortical regions, we used anatomical masks from the AFNI anatomical toolbox to partition these extensive activations (Eichoff et al., 2005). Using these masks, condition betas were extracted and averaged for significant voxels within each anatomical region listed in Table 3.[6] This procedure allowed us to examine the results of the contrast analyses with greater anatomical precision than if we had not broken down these large clusters that spanned diverse brain areas.

The four condition betas extracted in each region were for the atypical and typical valence focus conditions (VF-typ, VF-atyp) and for the typical and atypical scenario immersion conditions (SI-typ, SI-atyp).[7] Because the regions were defined by the valence focus contrast, the atypical valence focus condition was always significantly higher than the typical valence focus condition. Of primary interest was how the activity in the two scenario immersion conditions compared to the activity in the two valence focus conditions.

As illustrated by the schematic graphs in Figure 3, five contrasts were tested to identify different activation profiles occurring for typical and atypical emotions during scenario immersion and the valence focus that followed. These contrasts were developed to quantify a set of relations between conditions, indicating how activity during atypical and typical scenario immersion could be similar or different

---

[6] dmPFC in Table 3 corresponds to the anatomical mask for superior medial frontal gyrus from the toolbox.

[7] The condition betas described here are averages of the extracted category betas that made up each condition (e.g., betas from the regression for scenario immersion during pleasant fear, pleasant sadness, and unpleasant happiness were averaged to produce the atypical scenario immersion beta referred to above).

to activity during atypical and typical valence focus. Again, all contrasts specified greater activity for atypical valence focus than for typical valence focus because the regions examined were identified based on this contrast. The first contrast (C1) measured whether there was greater activity for atypical relative to typical conditions during both scenario immersion and valence focus. The second contrast (C2) measured whether there was greater activity during atypical valence focus than during all other conditions. The third contrast (C3) measured whether activity during both scenario immersion conditions was similar to atypical valence focus and greater than typical valence focus. The fourth contrast (C4) measured whether activity during both scenario immersion conditions was greater than the valence focus conditions, which still differed. The fifth contrast (C5) measured whether activity during the both scenario immersion conditions was less than the valence focus conditions, which still differed.

[Insert Figure 3 about here]

As the weights along the X axes of the graphs in Figure 3 illustrate, the five contrasts were computed by weighting the betas of each condition within the brain region of interest. Prior to these calculations, a constant integer (1) was added to all betas so all values were positive (i.e., the contrasts do not work properly when values are negative). This linear transformation of the data was independent of the activation pattern in the brain area, and thus had no effect on the contrast tested. For every participant, each condition beta (SI-typ, SI-atyp, VF-typ, VF-atyp) was multiplied by its numerical weight, with the sum of these products then being tested for significance from zero. Take the first contrast in Figure 3 as an example. The SI-typ beta was multiplied by -1, the SI-atyp beta multiplied by 1, the VF-typ beta multiplied by -1, and the VF-atyp beta multiplied by 1. Again, this contrast was designed to identify regions in which significantly greater activity occurred for atypical relative to typical emotion conditions during scenario immersion and valence focus. To compute significance, the weighted condition betas for each participant were summed, and then a one-sample, one-tailed t test was conducted on the mean of the sum to determine if it was significantly above zero. Due to the large number of tests, we considered a contrast significant for a region if it survived a threshold of $p < .0005$ ($t > 4.073$).

In most regions, two or three contrasts were significant. As Table 3 illustrates, we sorted the brain regions into five groups based on the pattern of significant contrasts, with the t statistic for significant contrasts shown for each region. To interpret the results for groups in which multiple contrasts were significant, we visually inspected graphs of the condition betas for the brain regions within a group. These graphs can be viewed in the appendix. As described in more detail shortly, the graphs indicated that regions within a group often contained elements of more than one contrast, which is why multiple contrasts were significant for many regions. Furthermore, in a few cases, regions were identified during visual inspection that showed patterns which appeared to be closer to a different group. When this occurred, we examined the contrast that would have shifted the region into the group that appeared most appropriate from visual inspection. If the contrast was marginally significant by our criteria ($p < .005$; $t > 2.95$), we moved the region into the group identified during visual inspection. For these regions, the *t* for the marginally significant contrast is italicized in Table 3, so that it is clear when these assignments were made. In the sections to follow, we describe the patterns of activation characterizing each group. In the Discussion, we then integrate these results with the results from the general contrast analyses presented earlier.

[Insert Table 3 about here]

The first group of brain areas was defined by the presence of a significant C1 contrast only. This group contained only one region, ventromedial prefrontal cortex (vmPFC). In this region, greater activation was observed for atypical relative to the typical conditions during both scenario immersion and valence focus.[8] Thus this region was important for processing atypical emotions during scenario immersion and valence focus in a similar manner.

The second group of brain areas was defined by the simultaneous significance of C1 and C2 contrasts. This group included dmPFC, SMA, right inferior frontal gyrus (pars orbitalis), posterior cingulate, and left-lateralized superior frontal gyrus, middle frontal gyrus, lateral parietal, and precuneus. Visual inspection of the condition means for these regions showed patterns that involved elements of both C1 and C2. As C1 specifies, greater activation was observed during atypical relative to typical scenario immersion and also during valence focus. Visual inspection further revealed, however, that activity in the atypical condition

---

[8] This result may seem surprising because vmPFC was not identified in the whole-brain scenario immersion atypical > typical contrast. Further inspection revealed that a vmPFC cluster of 15 voxels was active in the whole-brain analysis that did not pass the 36-voxel cluster threshold in the whole-brain analysis. Because a cluster of 15 voxels was identified in the whole-brain contrast, we discuss this region as being active during both atypical scenario immersion and valence focus in the Discussion.

increased substantially during valence focus, making it significantly greater than all other conditions, as C2 specifies. This shift also resulted in a much larger difference between the atypical and typical conditions during valence focus than during the preceding scenario immersion. Thus these regions appear to be driven most by atypical valence focus but also discriminate atypical relative to typical emotion states during scenario immersion.

The third group of brain areas was defined by the simultaneous significance of C1, C2, and C5 contrasts. This group included dorsal anterior cingulate, bilateral insula, bilateral middle orbital gyrus (i.e., anterior lateral orbitfrontal cortex), right superior and middle frontal gyrus, right lateral parietal and precuneus, and areas of the basal ganglia. Visual inspection indicated that the patterns of activation across conditions were most like C2 and C5. All regions showed significantly greater activity during atypical valence focus than all other conditions as C2 specifies, again suggesting that these regions were driven mostly by atypical valence focus. Additionally, many regions also showed strong trends towards typical valence focus being higher than both scenario immersion conditions as C5 specifies. In particular, dorsal anterior cingulate, anterior insula (especially on the right), right middle orbital gyrus, right inferior parietal, as well as the basal ganglia and thalamus, showed robust patterns corresponding to C5, in which atypical valence focus was greater than typical valence focus, which was greater than both preceding scenario immersion conditions. This pattern suggests that these regions were generally important during valence focus, and more so in the atypical condition.

Within this third group, the significant C1 contrast often appeared to be an artifact of the large differences between atypical and typical conditions during valence focus. Inspection of the beta patterns indicates that very few regions exhibited clear differences between atypical and typical emotions during scenario immersion. The exception was right lateral parietal regions, especially angular gyrus, and right middle frontal, where small differences occurred for the deactivations of atypical and typical conditions during scenario immersion.

The fourth group of brain areas was defined by the simultaneous significance of C1 and C3 contrasts, with visual inspection confirming elements of both contrasts. In the all sub-regions of left inferior frontal gyrus and in the more dorsal sub-regions of right inferior frontal gyrus (pars triangularis and opercularis), activity during typical valence focus was less than activity during both scenario immersion conditions and during atypical valence focus, suggesting that these regions were important during scenario immersion generally but then persisted during atypical valence focus only. Consistent with C1, atypical scenario immersion was often slightly greater than typical scenario immersion. Atypical valence focus usually did not differ from atypical scenario immersion.

The fifth group of brain areas was defined by the simultaneous significance of C1, C4, and C5 contrasts. Visual inspection of activations in bilateral MTG (the only region active in this group) exhibited a pattern that was clearly consistent with C4. The scenario immersion conditions did not differ from each other and were both greater than the valence focus conditions, which differed from each other. This pattern suggests that these regions were important for scenario immersion generally and to a lesser degree for atypical valence focus (relative to typical valence focus). The significant C1 and C5 contrasts appear to have resulted from elements of C4, including the large atypical-typical difference during valence focus (consistent with C1) and the significant activity during both scenario immersion conditions (consistent with C5).

As mentioned earlier, we performed the same contrast analyses on the clusters identified in the atypical vs. typical scenario immersion contrast. In the results for the conjunction shown in Figure 2, eight out of nine clusters in the scenario immersion contrast overlapped with clusters in the valence focus contrast. Of these overlapping regions, the smaller posterior cingulate and angular gyrus clusters identified in the scenario immersion contrast were the only regions that showed different contrast patterns from those identified in the same regions (defined by the valence focus contrast) in Table 3. Unlike the patterns described for these regions above, these clusters only showed a significant C1 contrast, consistent with the patterns in the first group shown in Table 3. Interestingly, this result suggests that in these particular parts of posterior cingulate and angular gyrus, the activity during atypical scenario immersion and valence focus was similar, with both conditions being greater than they were in the typical conditions.

Discussion

The results supported our two predictions. First, more default mode activity was observed in medial prefrontal cortex and posterior cingulate for atypical emotions during scenario immersion and valence focus. Second, more activity in lateral fronto-parietal attention and control networks was observed for atypical emotions during valence focus. Moreover, increased activity was also observed for atypical valence focus in the proposed salience network made up of anterior insular and cingulate regions.

Socio-Emotional Interpretation in the Default Network

Core hubs of the default network in medial prefrontal cortex and posterior cingulate were more active for atypical emotion during both scenario immersion and valence focus, consistent with the idea that experiencing atypical emotions involved greater interpretive processing related to self-projection into an unusual situation (Bar, 2007; Buckner & Carroll, 2007; Golland et al., 2008). Although the precise functions of regions in the default mode remain unclear, it has been suggested that medial prefrontal cortex underlies the processing of internal mental and affective states, whereas posterior cingulate underlies the processing of external social context (Johnson et al., 2006). To be elucidated in future work is exactly what elements of the atypical scenarios are related to increased activation in these regions. One possibility is that the atypical scenarios were more situationally complex, with more uncertainty surrounding the interpretation or conceptualization of the situation. Recent evidence suggests that medial prefrontal cortex responds more when another person's beliefs or preferences are ambiguous (Jenkins & Mitchell, 2010). Medial prefrontal cortex activation has also been observed in emotion regulation studies in which participants reappraise or reinterpret a negative stimulus (Ochsner & Gross, 2005).

In general, medial prefrontal cortex covers much brain territory and has broadly been associated with the processing of social inference, including evaluation and attitudes, self and trait knowledge, and understanding the minds of others (for reviews see Amodio & Frith, 2006; Mitchell, 2009; Van Overwalle, 2009). Furthermore, it has been proposed that medial prefrontal cortex supports different socio-emotional functions along a dorso-ventral axis (Amodio & Frith, 2006; Krueger, Barbey, & Grafman, 2009). Goals central to action monitoring are thought to be represented more dorsally, whereas the values of action outcomes are thought to be represented more ventrally. Consistent with this idea, ventromedial and dorsomedial prefrontal cortex displayed different patterns in the contrast analyses we performed. In a ventromedial prefrontal cluster, located in the frontal pole, activity during atypical scenario immersion and atypical valence focus was comparable. This frontal pole cluster is in the middle of the proposed dorso-ventral axis, in a late-developing region proposed to support the integration of dorsal goal and ventral value outcomes into more abstract representations (Krueger, Barbey, & Grafman, 2009). It appears that this integrated representation was of similar importance for atypical emotions during both scenario immersion and valence focus. Conversely, in the dorsomedial prefrontal cluster, activity during atypical valence focus was greater than during atypical scenario immersion. Because participants made a motor response to rate valence during valence focus periods, dorsomedial prefrontal regions representing goals for action may have become increasingly active. As Figure 2 illustrates, activity in this region was also much more extensive during valence focus.

Comparatively little is known about the functions that posterior cingulate supports. One proposal is that this region represents self-referential information in a temporal and spatial context (Johnson et al., 2006; Northoff et al., 2006). The posterior cingulate cluster defined by the valence focus contrast showed a similar pattern to dorsomedial prefrontal cortex, with activity increasing from atypical scenario immersion to atypical valence focus. Interestingly, however, the posterior cingulate cluster defined by the scenario immersion contrast showed a similar pattern to ventromedial prefrontal cortex, namely, comparable activation during atypical scenario immersion and valence focus. As shown in Figure 2, these two clusters overlapped, but also contained unique voxels. The significance of these patterns remains unclear.

Attention, Control, and Saliency during Internal States

As displayed in Figures 1 and 2, lateral fronto-parietal attention networks were robustly active during atypical valence focus, more so than during atypical scenario immersion. The contrast analyses further showed that the lateral fronto-parietal regions in groups two and three showed more activation during atypical valence focus than during the other three conditions. This pattern suggests that attentional and executive resources were most engaged when the internal valence feeling on which participants focused was atypical of the induced emotion (i.e., for fear during pleasant fear, for happiness during unpleasant happiness, and for sadness during pleasant sadness). Notably, no external stimulus was present during the valence focus period, providing additional evidence that attention systems act on internal states.

One possible interpretation of these fronto-parietal activation patterns is that focusing on the atypical feature of the emotion (i.e., valence) was generally more difficult or effortful than focusing on more typical valences. Another, non-mutually exclusive possibility is that this perhaps more complex state involved fast fluctuations between pleasant or unpleasant states that were compared to produce a valence estimate for guiding action.

Several fronto-parietal regions, which tended to be left lateralized, clustered in the second group that emerged from the contrast analyses. Visual inspection of these regions revealed that they often showed greater activation during atypical scenario immersion relative to typical scenario immersion, and showed a similar, more robust difference during valence focus. Of all the regions identified in group two, only a few showed an atypical vs. typical difference during scenario immersion that was robust enough to also be observed in the initial whole-brain atypical vs. typical scenario immersion contrast. To be conservative, we restrict our discussion to the regions showing these robust patterns, which were located in bilateral dorsolateral prefrontal cortex. Dorsolateral prefrontal cortex has been associated with executive functions related to maintaining goals and manipulating information in working memory (Duncan & Owen, 2000; Curtis & D'Esposito, 2003). Thus, one possibility is that the atypical scenarios may have been more situationally complex or ambiguous and required more integration in working memory.

The only frontal regions in which activity during atypical valence focus was not significantly greater than activity during scenario immersion was in bilateral inferior frontal cortex. In these regions, shown for group four of the contrast analyses, the activity levels observed during both scenario immersion conditions were maintained during atypical valence focus but not during typical valence focus.

It has been suggested that these regions play a role in selecting competing representations, and that in this way they may also support language processing in the left hemisphere (Thompson-Schill, Bedny, & Goldberg, 2005). Thus, the patterns in these regions suggest processes important for language processing during scenario immersion may have continued during valence focus only for atypical emotions.

A final pattern emerging from the contrast analyses was observed for a subset of the regions in group three. These regions showed a profile in which atypical valence focus was greater than typical valence focus, which was greater than both scenario immersion conditions. These were the only regions in which typical valence focus was more active than the scenario immersion conditions. Right anterior insula, left anterior insula (marginally), dorsal anterior cingulate, right middle orbital gyrus (anterior lateral orbitofrontal cortex), right inferior parietal, and bilateral basal ganglia and thalamus all exhibited this pattern.

Orbital fronto-insular and anterior cingulate regions have recently been implicated in a salience network that identifies important internal or extrapersonal stimuli for decision-making and action initiation (Seeley et al., 2007; Menon & Uddin, 2010). These regions have also been proposed to underlie awareness of how one feels (Craig, 2009). In our paradigm, it seems likely that participants may have focused on an internal feeling to make a decision (i.e., the valence rating) during valence focus. In other words, participants became aware of their affective state to initiate an action. Valence focus, relative to scenario immersion, may have tapped into the process of detecting a salient feature of experience in order to attend and/or act on it. Consistent with this proposal, inferior parietal cortex and basal ganglia, regions known to support goal-directed action, showed similar patterns (Buxbaum, Kyle, Grossman, & Coslett, 2007; Grahn, Parkinson, & Owen, 2008; Gross & Grossman, 2008; Tunik, Lo, Adamovich, 2008).

In addition to being generally more active during valence focus than during scenario immersion, these regions were also more active during atypical valence focus than during typical valence focus. As discussed above, this could indicate that general effort or difficulty were associated with processing the atypical emotions. Much work has similarly implicated the anterior cingulate in conflict monitoring, such as response competition (Botvinick, Cohen, & Carter, 2004). Perhaps participants experienced greater competition in processing the salience of unpleasant and pleasant feelings during atypical emotions than during typical emotions.

Integration of the Processes Underlying Atypical Emotions

During scenario immersion, atypical emotions recruited the default network, dorsolateral prefrontal regions in the executive network, as well as the temporoparietal junction and temporal pole, regions related to social cognition. This profile suggests that participants processed and interpreted complex socio-emotional information as they immersed themselves in atypical emotions. Interestingly, this pattern of default and executive activations is very similar to activations recently observed during mind wandering in a concurrent task (Christoff et al., 2009).

During atypical valence focus, default and executive regions showed similar or heighted activity relative to atypical scenario immersion. Bilateral angular gyrus was also similarly active during atypical scenario immersion and atypical valence focus, but temporal pole was not, only being active during scenario immersion. This activation profile suggests that the complex, internal socio-emotional simulations generated during atypical scenario immersion were largely maintained during atypical valence focus. In addition, a few regions that were active in both atypical and typical scenario immersion remained active to some degree during atypical valence focus (but not typical valence focus). These activations in bilateral inferior frontal and middle temporal gyri may be related to language and auditory processing generally important during scenario immersion that continued only during atypical valence focus.

Finally, atypical valence focus recruited fronto-parietal networks involved in attention and control, as well as regions involved in salience detection and goal-directed action, which were not observed during atypical scenario immersion. Presumably, these networks came online to support focusing on valence, the atypical feature of the emotion. The regions associated with salience detection and goal-directed actions were also active, albeit to a lesser degree, for typical valence focus. Thus the anterior cingulate and insular regions that make up the proposed salience network appear to be central to becoming aware of one's affective state and responding based on it.

References

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology, 60*, 693-716.

Amodio, D.M., & Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7*, 268-277.

Bar, M. (2007). The Proactive Brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280-289.

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

Barrett, L.F., Barsalou, L.W., Lindquist, K., & Wilson-Mendenhall, C.D. (in preparation). The Conceptual Act Theory of Emotion.

Barsalou, L.W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 629-654.

Barsalou, L.W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101-140). Cambridge: Cambridge University Press.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-609.

Barsalou, L.W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes, 18*, 513-562.

Barsalou, L.W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London: Biological Sciences, 364*, 1281-1289.

Barrett, L. F. (2009). Variety is the spice of life: A psychologist constructionst approach to understanding variability in emotion. *Cognition and Emotion*, *23*, 1284-1306.

Bohlhalter, S., Hattori, N., Wheaton, L., Fridman, E., Shamim, E.A., Garraux, G., & Hallett, M. (2009). Gesture subtype–dependent left lateralization of praxis planning: An event-related fMRI study. *Cerebral Cortex, 19*, 1256-1262.

Botvinick, M.M., Cohen, J.D., & Carter, C.S. (2004). Conflict monitoring and anterior cingulated cortex: an update. *Trends in Cognitive Sciences, 8*(12), 539-54.

Bressler, S.L., & Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends in Cognitive Sciences, 14*, 277-290.

Buckner, R.L., Andrews-Hanna, J.R., & Schacter, D.L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences, 1124*, 1-38.

Buckner, R.L., & Carroll, D.C. (2007). Self-projection and the brain. *Trends in Cognitive Science, 2*, 49-57.

Christoff, K., Gordon, A.M., Smallwood, J., Smith, R., & Schooler, J. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences, 106*(21), 8719-8724.

Corbetta, M., Patel, G., & Shulman, G.L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron, 58*, 306-324.

Cox, R.W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*, 162-173.

Craig, B. (2009). How do you feel--now? Interception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience, 3*, 655-666.

Curtis, C.E., & D'Esposito (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences, 7*(9), 415-423.

Damasio , H., Tranel , D., Grabowski , T., Adolphs , R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition, 92*, 179-229.

Drane, D.L., Ojemann, G.A., Aylward E., Ojemann J.G. Johnson, L.C., Silbergeld D.L., et al. (2008). Category-specific naming and recognition deficits in temporal lobe epilepsy surgical patients. *Neuropsychologia, 46*, 1242-1255.

Dube, L. & Le Bel, J.L. (2003). The content and structure of laypeople's concept of pleasure. *Cognition and Emotion, 17*(2), 263-295.

Duncan, J., & Owen, A.M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neuroscience, 23*(10), 475-83.

Fehr, B. & Russell, J. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113*(3), 464-486.

Fehr, B. & Russell, J. (1991). The concept of love viewed from a prototype perspective. *Journal of Personality and Social Psychology, 60*(3), 425-438.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C. & Raichle, M.E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Processing of the National Academy of Sciences, 102*(27), 9673-9678.

Golland, Y., Bentin, S., Gelbard, H., Benjamini, Y., Heller, R., Nir, Y., Hasson,U., & Malach, R. (2007). Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. *Cerebral Cortex, 17*, 766-777.

Golland, Y., Golland, P., Bentin, S., & Malach, R. (2008). Data-driven clustering reveals a fundamental subdivision of the human cortex into two global systems. *Neuropsychologia, 46*, 540-553.

Grahn, J.A., Parkinson, J.A., & Owen, A.M. (2008). The cognitive functions of the caudate nucleus. *Progress in Neurobiology, 86*, 141-155.

Gross, R.G., & Grossman, M. (2008). Update on apraxia. *Current Neurology and Neuroscience Reports, 8*, 490-496.

Gusnard, D.A., & Raichle, M.E. (2001). Searching for a baseline: Functional imaginging and the resting human brain. *Nature Reviews Neuroscience, 2*, 685-694.

Harrison, B.J., Pujol, J., Lopez-Sola, M., Ribas-Hernandez, R., Deus, J., Ortiz, H., Soriano-Mas, C., Yucel, M., Pantelis, C., & Cardoner, N. (2008). Consistency and functional specialization in the default mode brain network. *Proceedings of the National Academy of Sciences, 105*(28), 9781-9786.

Jenkins, A.C., & Mitchell, J.P. (2010). Mentalizing under uncertainty: Dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cerebral Cortex, 20*, 404-410.

Johnson, M.K., Raye, C.L., Mitchell, K.J., Touryan, S.R., Greene, E.J., & Nolen-Hoeksema, S. (2006). Dissociating medial frontal and posterior cingulate activity during self-reflection. *Social Cognitive and Affective Neuroscience, 1*, 56-64.

Kober, H., Barrett, L.F., Joseph, H., Bliss-Moreau, E., Lindquist, K., & Wager, T.D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *NeuroImage, 42*, 998-1031.

Krueger, F., Barbey, A.K., & Grafman, J. (2008). The medial prefrontal cortex mediates social event knowledge. *Trends in Cognitive Sciences, 13*(3), 103-109.

Larsen, J.T., McGraw, A.P., & Cacioppo, J.T. (2001). Can people feel happy and sad at the same time. *Journal of Personality and Social Psychology, 81*(4), 684-696.

Larsen, J.T., McGraw, A.P., Mellers, B.A., & Cacioppo, J.T. (2004). The agony of victory and thrill of defeat. *Psychological Science, 15*(5), 325-330.

Menon, V., & Uddin, L.Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function, 214*, 655-667.

Miller, E.K., & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

Mitchell, J.P. (2009). Social psychology as a natural kind. *Trends in Cognitive Sciences, 13*, 246-251.

Miyaoto, Y., Uchida, Y., & Ellsworth, P.C. (2010). Culture and mixed emotions: Co-occurrence of positive and negative emotions in Japan and the United States. *Emotion, 3*, 404-415.

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre supplementary motor areas. *Nature Reviews Neuroscience, 9*, 856-869.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage, 31*, 440-457.

Picard, N., & Strick, P.L. (1996). Motor areas of the medial wall: A review of their location and functional activation. *Cerebral Cortex, 6*, 342-353.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, H.D., & Panksepp, J. (2006). Self referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage, 31*, 440-457.

Northoff, G., & Panksepp, J. (2008). The trans-species concept of self and the subcortical cortical midline system. *Trends in Cognitive Sciences, 12*(7), 259-264.

Ochsner, K.N., & Gross, J.J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences, 9*(5), 242-249.

Ollinger, J.M., Corbetta, M., Shulman, G.L. (2001). Separating Processes within a trial in event-related functional MRI: II. Analysis. *NeuroImage, 13*, 218-229.

Ollinger, J.M., Shulman, G.L., Corbetta, M. (2001). Separating processes within a trial in event-related functional MRI: I. Method. *NeuroImage, 13*, 210-217.

Olsson, A., & Ochsner, K.N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences, 12*(2), 65-71.

Posner, M.I. (2008). Measuring alertness. *Annals of the New York Academy of Sciences, 129*, 193-199.

Raichle, M.E. (2010). Two views of brain function. *Trends in Cognitive Sciences, 14*(4), 180-190.

Rosch, E., & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

Ruge, H., Goschke, T., Braver, T.S. (2009). Separating event-related BOLD components within trials: The partial-trial design revisited. *Neuroimage, 47*, 501-513.

Russell, J.A. (1991). In defense of a prototype approach to emotion concepts. *Journal of Personality and Social Psychology, 60*(1), 37-47.

Russell, J.A. (2003). Core affect of the psychological construction of emotion. *Psychological Review, 110*(1), 145-172.

Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., & Greicus, M.D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of Neuroscience, 27*(9), 2349-2356.

Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., & Peterson, S.E. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *Journal of Cognitive Neuroscience, 9*(5), 648-663.

Simmons, W.K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, *15* 645-649.

Thompson-Schill, S., Bedny, M., & Goldberg, R.F. (2005). The frontal lobes and the regulation of mental activity*. Current Opinion in Neurobiology, 15*, 219-224.

Tranel , D. (2006). Impaired naming of unique landmarks is associated with left temporal polar damage. *Neuropsychology, 20*, 1-10.

Tunik, E., Lo, O.Y., Adamovich, S.V. (2008). Transcranial magnetic stimulation to the frontal operculum and supramarginal gyrus disrupts planning of outcome-based hand–object interactions. *Journal of Neuroscience, 28,* 14422-14427.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping, 30*, 829-858.

Wilson-Mendenhall, C.D., Barrett, L.F., & Barsalou, L.W. (in preparation). Deconstructing emotion experience: Neural evidence for core affect.

Wilson-Mendenhall, C.D., Barrett, L.F., Simmons, W.K., & Barsalou, L.W. (submitted). Grounding emotion in situated conceptualization.

Table 1. Example typical and atypical full scenarios with the core version in italics.

Examples of atypical fear, happiness, and sadness scenarios

You are jogging onto the soccer field, your cleats digging into the firm ground. *You hear a booming voice welcome the crowd to the state championship*. You jump in place to shake off the restlessness in your stomach. Looking around at your team, a rushing excitement deepens your competitive fire. *You feel an energizing fear.*

You are standing on your college quad, dressed in a smart looking cap and gown. *You listen for the graduation decree and upon hearing it fling your cap upwards.* Following energetic classmates, you sweat lightly as you march away a graduate. You catch a friend's eye and flashback to your delightfully lively freshman dorm. *You feel a spirited sadness.*

You are walking down the hall, trying to get to a meeting on time. *You run into a difficult colleague and end a tense exchange with a biting remark.* Your stomach tightens the moment the last sarcastic jab escapes your lips. The cutting retort echoes poisonously in your head as your colleague sulks away. *You feel a disturbing happiness.*

Examples of typical fear, happiness, and sadness scenarios

You are walking to your car alone, the city parking deck dimly lit. *You hear an explosive bang and see a man running with a pointed gun.* You quickly drop behind a car and attempt to control your shallow breathing. You try to dismiss the horrendous vision of what will happen if he finds you. *You feel a perilous fear.*

You are walking into a friend's house, dropping by to return a movie. *You witness your significant other in an intimate embrace with your friend.* Your stomach is nauseated, the shocking infidelity settling into your body. Your mind is spinning trying to understand the terrible betrayal of trust. *You feel a devastating sadness.*

You are performing a challenging piano solo, your fingers working the keys. *You finish the piece and receive thunderous applause as you rise.* You bend at the waist into a deep bow and sense your heart thumping rapidly. Glowing with satisfaction, you continue to feed off the crowd's energy. *You feel a proud happiness.*

Note. Scenarios also systematically varied in arousal. For comparison purposes, all the examples shown here are high arousal.

Table 2. Coordinates, extent, and statistics for regions more active for atypical emotions during scenario immersion and valence focus. For large clusters, multiple local peaks a minimum of 10mm apart are reported, labeled in italics.

		Brodmann	Spatial	Peak			
Cluster	Brain Region	Area(s)	Extent	x	y	z	*t*
scenario immersion, atypical > typical							
1	dmPFC/SMA	8,6	167	2	35	45	6.59
2	dmPFC	9,10	109	8	44	30	6.21
3	R Angular gyrus	39	97	47	-59	24	6.54
4	L Angular gyrus	39	96	-41	-59	24	5.62
5	L Inferior Frontal gyrus	45	51	-44	20	9	5.53
6	L Middle Frontal gyrus	9,6,8	48	-44	14	48	5.98
7	L Temporal Pole	38	45	-41	11	-31	5.53
8	Post Cingulate/Precuneus	31,7	40	-8	-47	30	4.28
9	R Middle Frontal gyrus	9,8	38	35	17	39	5.38
valence focus, atypical > typical							
1	dmPFC/Ant Cing/SMA/ R Lateral PFC/R Ant Insula	32,8,9,6,46 45,44,47,11,10	2758				
	dmPFC			-8	20	48	11.98
	dmPFC			5	23	45	9.02
	R Middle Frontal gyrus			44	20	36	7.65
	R Inferior Frontal gyrus			50	20	12	7.35
	R Orbital Frontal			44	41	-4	7.17
	R Inferior Frontal gyrus			50	20	27	6.24
2	L Lateral PFC/L Ant Insula/ BG/Thalamus/Brainstem	8,9,6,46, 45,44,47,11,10	1737				
	L Inferior Frontal gyrus			-44	17	24	11.56
	L Ant Insula			-29	23	6	9.88
	L Orbital Frontal			-38	23	-4	8.61
	L Pallidum			-11	-2	9	7.05
	L Inferior Frontal gyrus			-44	32	9	6.53
3	L Lateral Parietal	39,40	375	-35	-59	54	5.89
4	R Lateral Parietal	39,40	319	41	-59	39	6.31
5	Post Cingulate/Precuneus	23,31,7	219	5	-41	18	5.59
6	R BG/Thalamus/Brainstem		183	11	-5	12	7.60
7	L Middle Temporal gyrus	21,22	163	-59	-38	-4	7.09
8	R Middle Temporal gyrus	21,22	134	56	-20	-7	5.47
9	vmPFC	10,11	80	-8	53	-10	6.00

Note. Spatial extent is the number of $27mm^3$ functional voxels. L is left and R is right. Post is posterior and Ant is anterior. dmPFC and vmPFC is dorsomedial and ventromedial prefrontal cortex. SMA is supplementary motor area. BG is basal ganglia.

Table 3. Groupings of brain regions identified in the ayptical > typical contrast (during valence focus) according to significant contrasts.

Group	Brain Region	Spatial Extent	C1	C2	C3	C4	C5
1	vmPFC	80	7.06				
2	dmPFC	573	9.10	6.34			
2	SMA	194	8.34	5.75			
2	L Superior Frontal g	236	8.98	5.09			
2	L Middle Frontal g	324	8.03	4.89			
2	R Inferior Frontal g *(p. orb)*	256	7.18	5.07			
2	Posterior Cingulate	45	4.81	*3.57*			
2	L Inferior Parietal	158	5.47	4.71			
2	L Angular g	96	7.11	4.13			
2	L Superior Parietal	83	5.35	4.24			
2	L Precuneus	19	4.74	4.23			
3	Anterior Cingulate	119	5.79	7.64			6.6
3	L Anterior Insula	53	7.17	5.84			4.15
3	R Anterior Insula	58	7.76	4.39			*3.40*
3	L Middle Orbital g	107	6.21	7.29			4.41
3	R Middle Orbital g	83	5.49	7.38			5.98
2	R Superior Frontal g	271	7.14	5.89			*3.95*
3	R Middle Frontal g	327	5.98	7.59			5.56
3	R Inferior Parietal	38	4.75	6.90			7.06
3	R Angular g	203	6.13	6.45			4.76
3	R Superior Parietal	46	4.51	4.83			*3.77*
3	R Precuneus	76	5.45	6.70			5.58
3	L Pallidum	34	6.20	5.89			4.93
3	R Pallidum	11	6.84	6.73			6.52
3	L Thalamus	53	7.03	6.61			5.44
3	R Thalamus	20	6.28	6.23			5.54
3	R Caudate	66	6.58	6.15			5.17
3	L Caudate	57	5.64	6.79			5.97
4	L Inferior Frontal g *(p. op)*	61	7.47		4.65		
4	R Inferior Frontal g *(p. op)*	75	8.02		*4.02*		
4	L Inferior Frontal g *(p. tri)*	464	8.83		7.70		
4	R Inferior Frontal g *(p. tri)*	260	10.45		5.17		
4	L Inferior Frontal g *(p. orb)*	228	8.74		4.16		
5	L Middle Temporal g	163	5.68		6.87	5.61	
5	R Middle Temporal g	134	5.76		7.32	6.58	

Note. Spatial extent in 27mm^3 functional voxels. Regions with an italicized *t*, indicating lesser signficance of .005, were moved into the specified group based on visual inspection of plotted betas. L is left and R is right; G is gyrus; dmPFC and vmPFC is dorsomedial and ventromedial prefrontal cortex; SMA is supplementary motor area. In parentheses, *p. op* is pars opercularis, *p. tri* is pars triangularis, and *p. orb* is pars orbitalis.

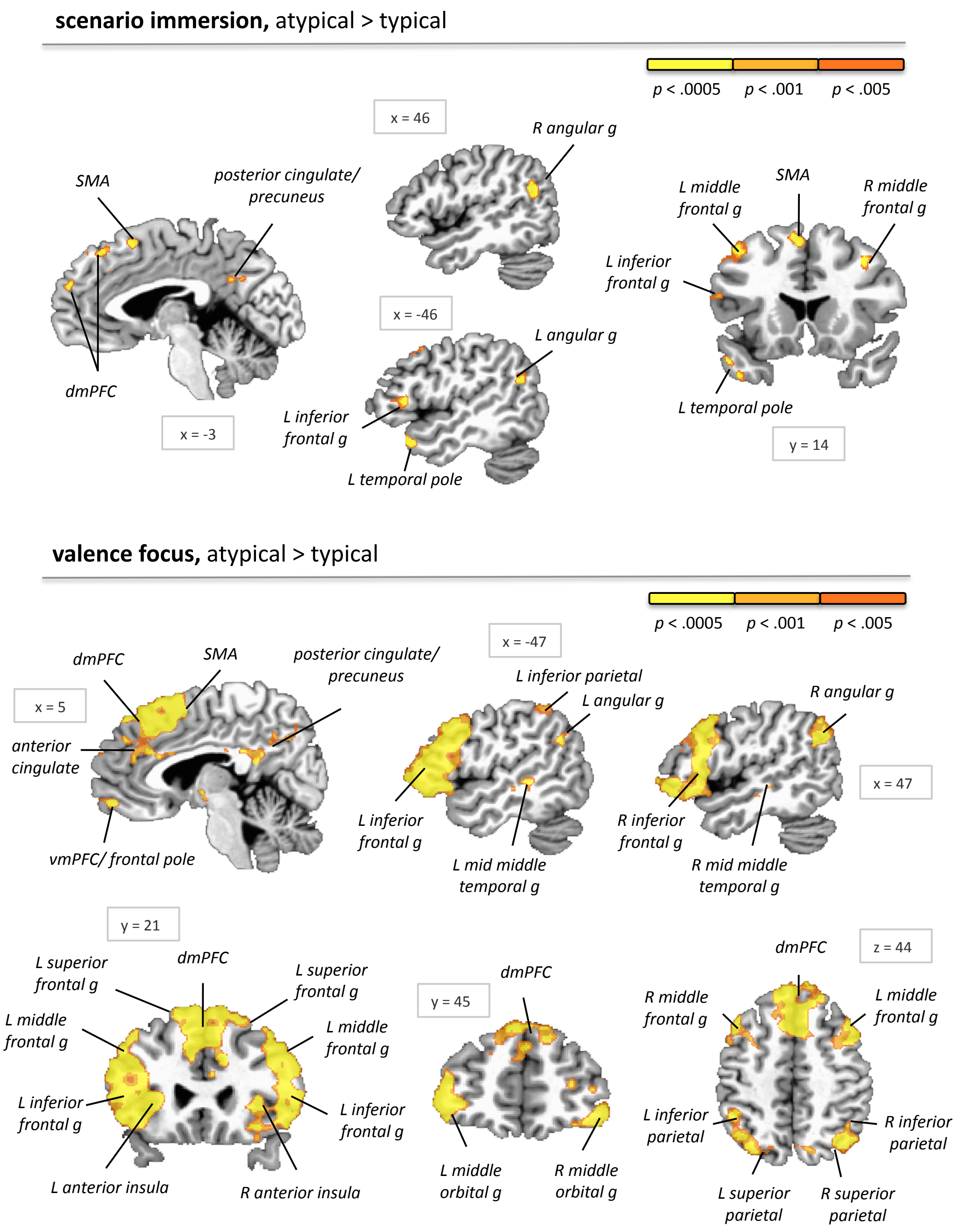


Figure 1. Brain regions more active for atypical vs. typical emotions during scenario immersion and valence focus.

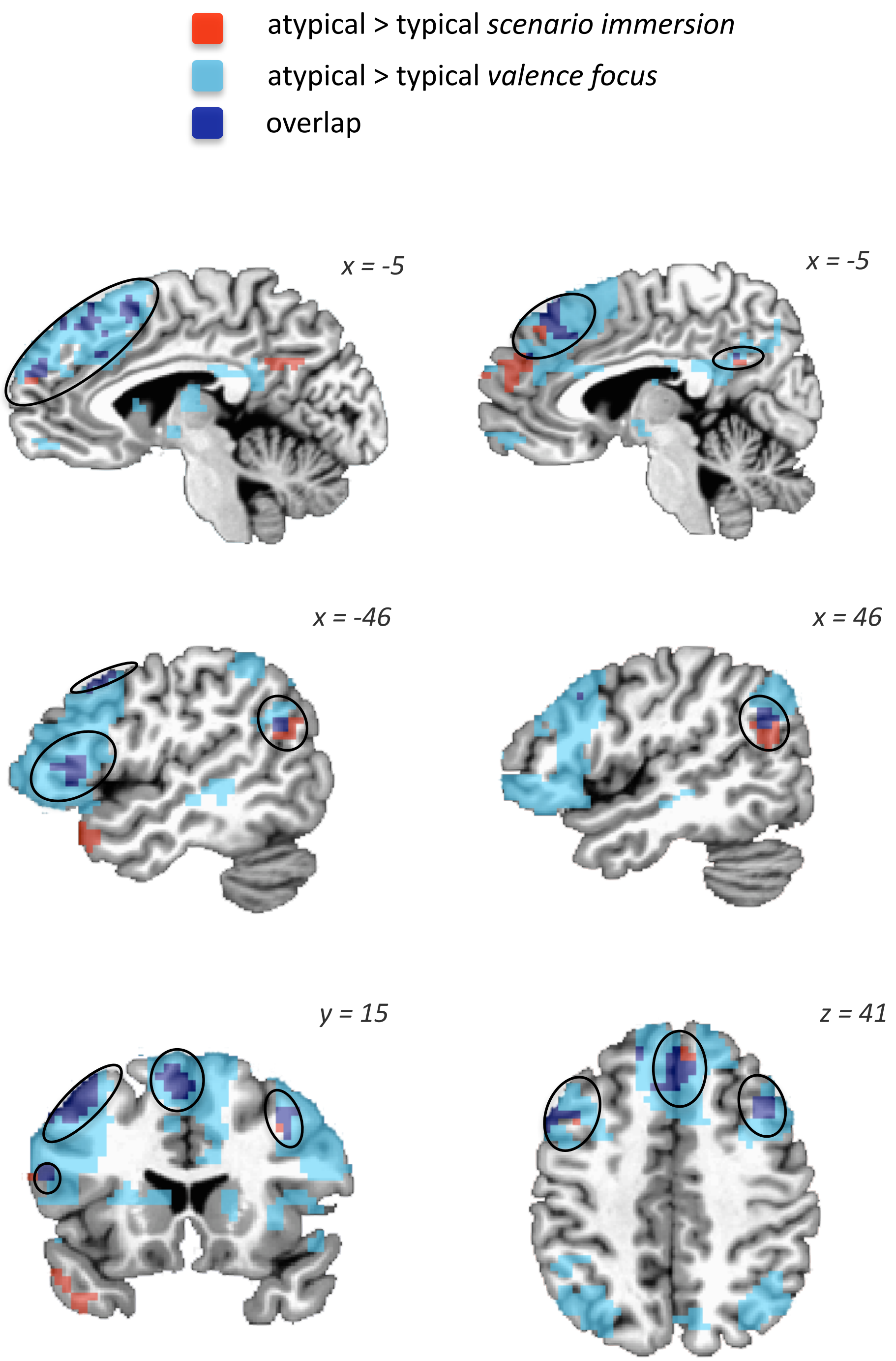


Figure 2. Conjunction of activity for atypical emotions during scenario immersion and during valence focus.

Contrast 1
Greater activity for atypical than typical valence focus which also occurs during preceding scenario immersion.

Contrast 2
Greater activity for atypical than typical valence focus. Also greater activity during atypical valence focus than during preceding scenario immersion, which is similar to typical valence focus.

Contrast 3
Greater activity for atypical than typical valence focus. Activity during preceding scenario immersion is also greater than typical valence focus, which is similar to atypical valence focus.

Contrast 4
Greater activity for atypical than typical valence focus. Activity during preceding scenario immersion is greater than typical and atypical valence focus.

Contrast 5
Greater activity for atypical than typical valence focus. Activity during preceding scenario immersion is less than typical and atypical valence focus.

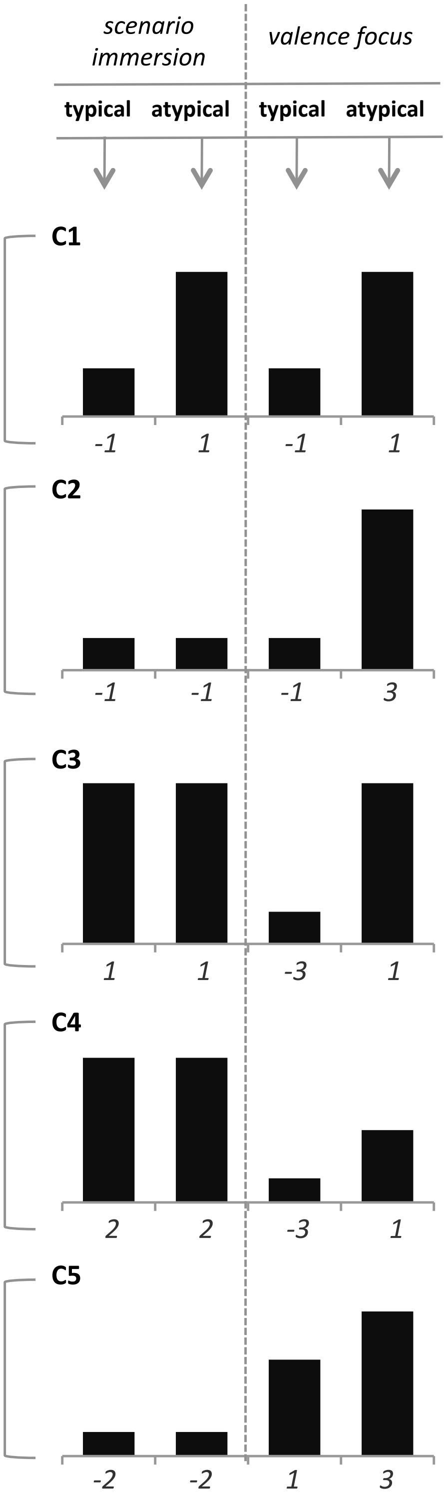
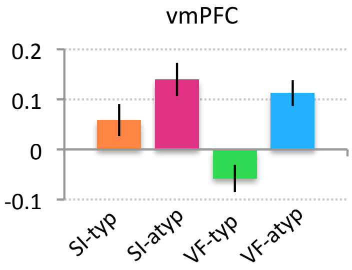


Figure 3. Description and schematic image corresponding to each contrast. Contrast weights used in each analysis are shown in gray on the x-axis.

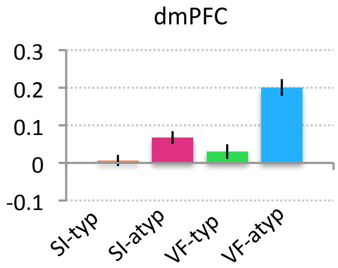
Appendix

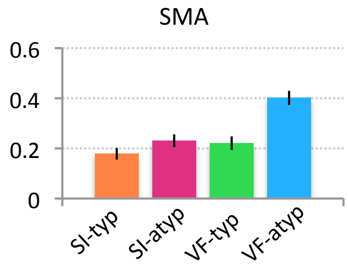
All histograms show the mean and standard error of participant condition betas in terms of percent signal change (the y-axis). A star (*) indicates the region was moved into its current group based on visual inspection.

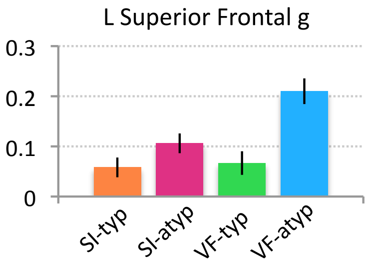
Group - C1 only

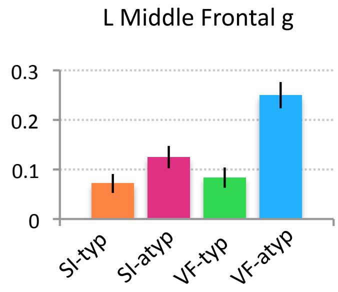


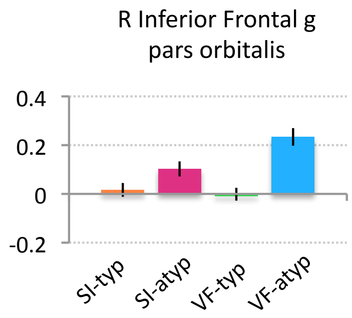
Group 2 – C1 & C2

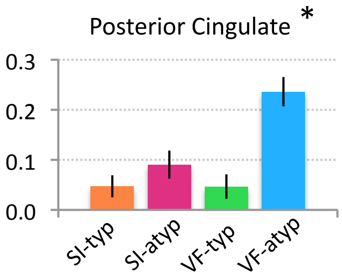


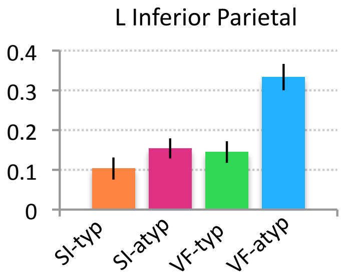


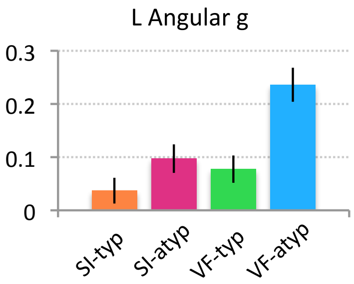


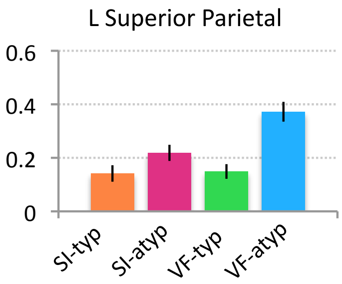




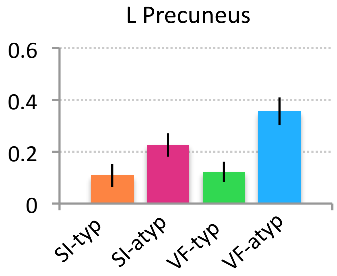




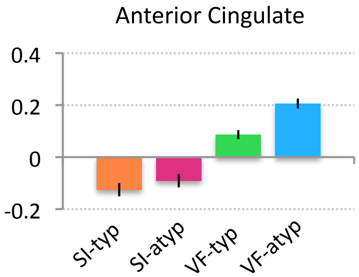


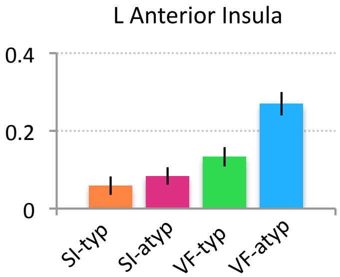


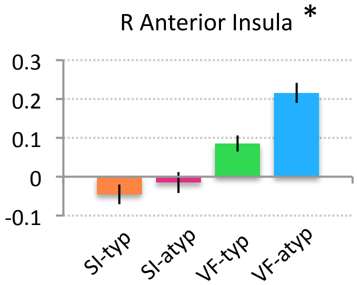
Group 2 (continued)

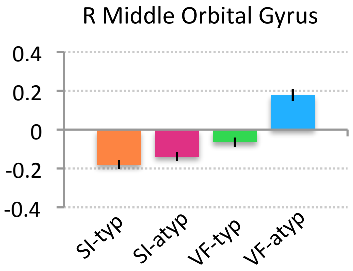

L Precuneus

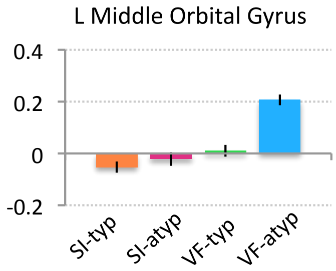
Group 3 – C1, C2, & C5

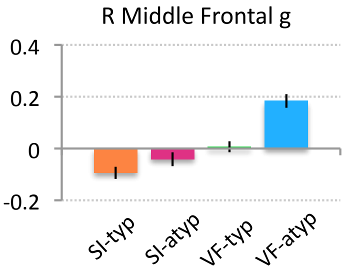

Anterior Cingulate

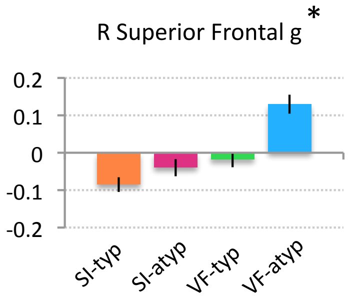

L Anterior Insula

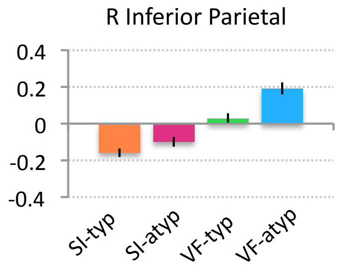

R Anterior Insula *

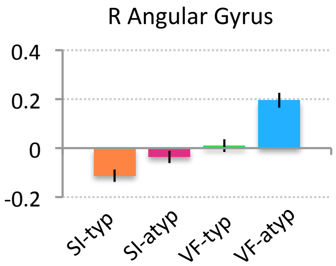

R Middle Orbital Gyrus


L Middle Orbital Gyrus

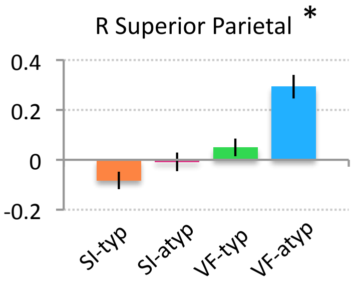

R Middle Frontal g

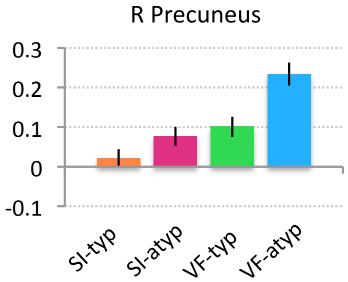

R Superior Frontal g *

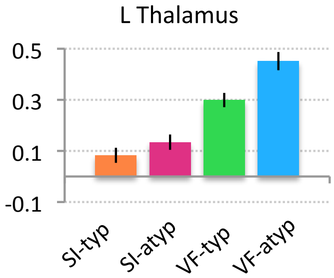

R Inferior Parietal

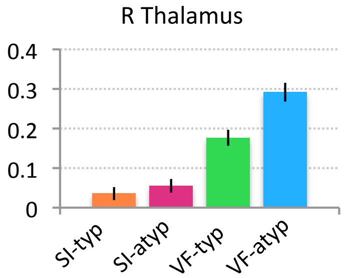

R Angular Gyrus

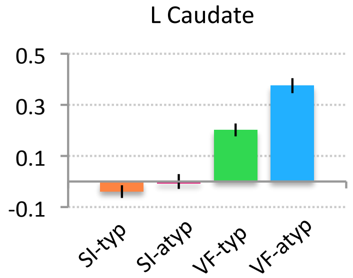
Group 3 (continued)

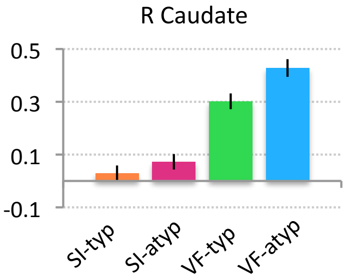

R Superior Parietal *

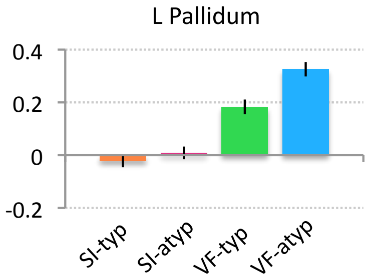

R Precuneus

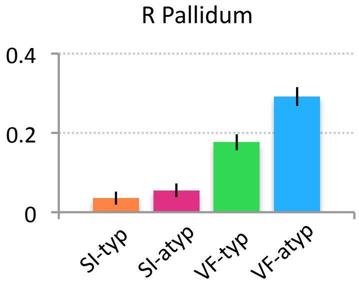

L Thalamus


R Thalamus

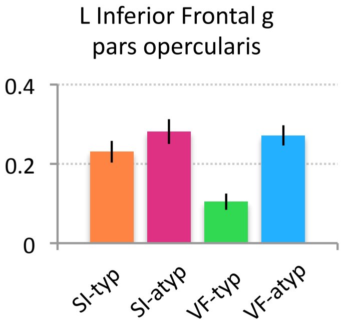

L Caudate

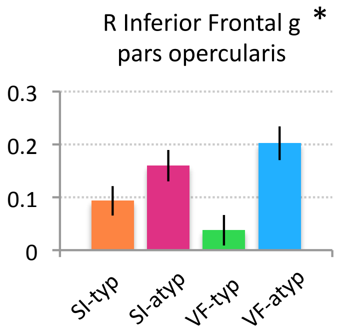

R Caudate

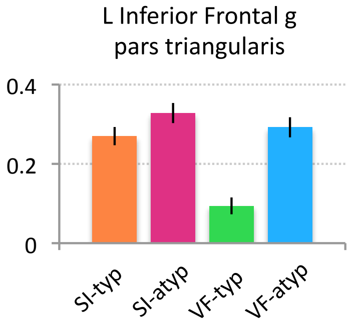

L Pallidum


R Pallidum

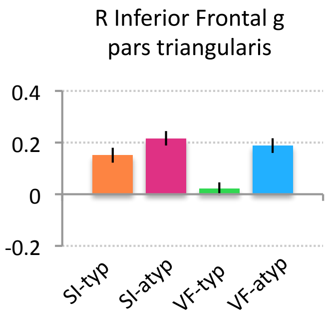
Group 4 – C1 & C3

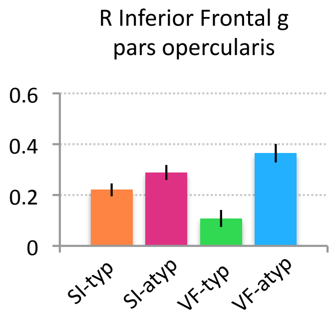

L Inferior Frontal g
pars opercularis


R Inferior Frontal g *
pars opercularis

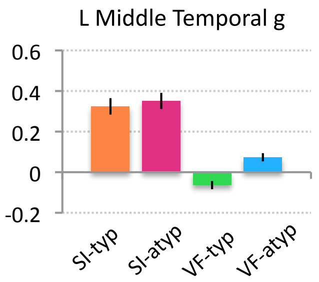

L Inferior Frontal g
pars triangularis

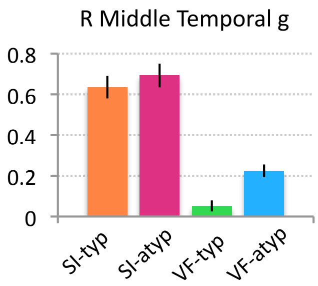
Group 4 (continued)


R Inferior Frontal g
pars triangularis


R Inferior Frontal g
pars opercularis

Group 5 – C1, C3, & C4


L Middle Temporal g


R Middle Temporal g

General Discussion

In the manuscripts just presented, we examined two distinct sets of findings from the same experiment. In the first manuscript, we presented neural evidence for the core affect properties of valence and arousal. We suggested that fundamental neural systems evolved to support core affect instead of category-specific emotion programs. Consistent with a psychological construction approach, we view these basic mechanisms as necessary for emotion experience, but not sufficient. The affect generated by these systems is integrated within a general interpretive, conceptual system to produce emotion (Barrett, 2006, 2009).

The second manuscript emerged from the methods used to test the core affect hypothesis, in which we created atypical scenarios for each category in order to vary valence. As described in this manuscript, there were robust differences between emotions when valence was typical relative to when it was atypical. Furthermore, these differences occurred in large-scale networks that have been the focus of much recent work. Atypical emotions appear to draw more on the default network to generate social-emotional simulations, perhaps more complex or ambiguous in nature. Atypical emotions also placed greater demand on attention, executive, and salience networks during focus on the atypical valence of these emotions. We believe this work also is also consistent with constructivist approaches that suggest emotions are generated from processing rich socio-emotional situations, and thus vary tremendously across the situations in which they are experienced (see also Wilson-Mendenhall, Barrett, Simmons, & Barsalou, submitted).

In this final discussion section, I address some emergent themes from these two pieces of work. The primary goal of this discussion is to address issues relevant to future research directions. In this way, it will be less tied to previous research and more speculative in nature.

Revisiting Basic Emotions

The idea of basic emotions is intuitively appealing. We have all felt a piercing moment of fear, a profound sadness, a glowing joy, a burning anger. In our work, however, we have argued against the idea that neural programs detect some number of basic emotions and respond in a stereotyped manner. Instead, we propose that emotion experience is much more flexible, emerging from the interactions of distributed brain systems.

This does not mean that these categories of experience are unimportant. One possibility is these categories tend to involve common relations between pleasure, pain, and self. For example, one key relation appears to be process of attachment that occurs when we experience pleasure and then 'crave' a pleasure producing object. This process describes the basic capacities of 'liking' and 'wanting' underlying reward, recently thought to be somewhat dissociable (Berridge & Kringelbach, 2008). Once we are attached to something, we can experience the loss of it, a fundamental part of sadness. Once we are attached to something, we can experience being blocked from having it, a fundamental part of anger. A recent study showed that more physical effort was expended to win objects subliminally paired with angry faces than those paired with neutral faces, suggesting reward may be associated with anger (Aarts et al., 2010). Furthermore, the idea that afflictive emotions are grounded in attachment is a central theme of Buddhist philosophies that have been around for centuries.

Other basic relations appear to revolve around pain. The threat or anticipation of pain seems central to fear, and also seems more likely to develop in situations during which pain was actually experienced (e.g., being afraid to ride a horse after falling off).

In general, these basic principles of pleasure and pain appear to be bodily in nature; initially children want food, comfort (touch), sensory stimulation (e.g., toys), and they initially avoid various forms of physical pain, which again is often embodied. Presumably evolutionary constraints exist as to what is pleasurable (and thus what we become attached to) and what may cause pain. As we develop more abstract concepts of self in relation to others, these basic pleasure and pain relations appear to operate increasingly in a social context. Attachment now operates on achievements and goals that are consistent with one's self concept, producing negative emotions when not met or blocked. Similarly, fears develop anticipating psychological pain that could be caused by situations like social rejection. In line with this thinking, it appears that social pain is experienced in similar ways to physical pain in the brain (Eisenberger & Lieberman, 2004; Eisenberger, Lieberman, & Williams, 2003).

As discussed in other recent work, we believe that these forms of core affect are just one aspect of situated conceptualizations that develop to represent emotions (Barrett, Barsalou, Lindquist, & Wilson-Mendenhall, in prep; Wilson-Mendenhall, Barrett, Simmons & Barsalou, submitted). Because affect is experienced in many different situations, situated conceptualizations develop to represent an emotion like fear in many different forms. These conceptualizations are constantly used to interpret experiences, producing inferences that can cause an emotion state. From this theoretical perspective, these mechanisms underlie our ability to experience a diverse range of emotion states that support diverse forms of physical and social interaction in the world. Thus the dynamic, situated operations of the conceptual systems seem critical for understanding emotion experience. Nevertheless, in the spirit of basic emotion approaches, it also seems important to consider fundamental mechanisms of core affect and how they may be related to common emotion categories like fear or anger.

Emotion Conceptualization Unfolds over Time

In the paradigm used in the experiment described in both manuscripts, brain activity was measured in blocks of several seconds. Presumably, core affect and conceptualizations of this affect within a situation occur much faster in time. Core affect is assumed to be a state that is constantly fluctuating. As core affect fluctuates, so does activity in many other brain systems that process perception, action, and internal states. In Conceptual Act Theory, a constructivist approach, occasional conceptual acts occur that classify these patterns of activity as emotions (Barrett, Barsalou, Lindquist, & Wilson-Mendenhall, in prep; Wilson-Mendenhall, Barrett, Simmons & Barsalou, submitted). When this happens, a situated conceptualization developed from prior experience classifies the pattern as an instance through pattern detection mechanisms. This process is typically not a conscious, deliberate event. As construal takes place, prior experience contained in the situated conceptualization is used to make inferences or predictions. Very quickly, this conceptualization can change core affect and other bodily responses that underlie the experience of an emotion.

One question, then, is how core affect fluctuates in a dynamic manner as a situation unfolds, and how it changes when a conceptualization occurs. In the paradigm discussed here, the conceptualization, including the feeling of the emotion, was very deliberately developed via a scenario. Participants reported their subjective experience of core affect after construing an emotion scenario as fear, happiness, or sadness. Furthermore, in many of our scenarios, the fear, happiness, or sadness situation was somewhat atypical. It would be interesting to measure core affect in a more dynamic manner moment to moment as the situation was unfolding and then after it was explicitly categorized, especially in atypical cases. Perhaps measures of body physiology, EEG, and MEG could be used to inform this issue. A related question is how valuations from many sources (e.g., pleasant or unpleasant tastes, touches, mental states, etc) become integrated over time to influence decision-making and action.

Integrating Emotion and Cognition in Large-Scale Networks

A final theme revolves around recent thinking about large-scale networks of dynamically interacting brain regions. As discussed in the second manuscript, the default network appears to be central for producing social and self-related inferences that we assume are also central to emotion experience. Furthermore, network approaches are generating new questions about the way affect, attention, and cognitive control are integrated. Recent theorizing about a salience network that integrates highly processed sensory information with visceral, autonomic, and hedonic information to select stimuli to attend to and/or act on illustrates this point (Seeley et al., 2007). The organization of these networks is consistent with recent arguments that it is counterproductive to think about cognition and emotion as separate processing domains (Barrett, 2009; Duncan & Barrett, 2007; Pessoa, 2008).

An interesting observation about the large-scale networks established in recent work is that they do not appear to include the two regions we found to be correlated with core affect in the first manuscript, medial orbitofrontal cortex and amygdala. A major distinction in the network literature is between task-positive and task-negative networks. In comparison to other networks, the default mode has been discussed as being task-negative because it deactivates during traditional cognitive tasks involving attention to the external world. Also, recent evidence suggests that attention and default networks are anti-correlated during rest states, so that when one is activated, the other is deactivated (Fox et al., 2005). In contrast to these networks, it may be that regions underlying the processing of core affect in medial orbitofrontal cortex and amygdala form a network that is operating constantly, producing affective information that is critical to both networks. Core affect is thought to play a role in many forms of decision-making and may even be a fundamental feature of consciousness (Barrett, 2009).

Concluding Remarks

Clearly, much remains to be learned about emotion experience. In the two manuscripts presented here, a central idea is that emotion emerges from the interactions of distributed circuits in the brain. This idea suggests a shift in our approach to studying emotion, one that is less focused on specific emotion categories and more focused on identifying the underlying processes central to many categories of emotion. At the same time, common emotion categories may still help us understand how primitive affective mechanisms work. A next step is to reconsider what we know about these categories, as we begin to deconstruct our subjective experience of them into neural mechanisms.

References

Aarts, H, Ruys, K.I., Veling, H., Renes, R.A., de Groot, J.H., van Nunen, A.M., & Geertjes S. (2010). The art of anger: reward context turns avoidance responses to anger-related objects into approach. *Psychological Science, 21*(10), 1406-1410.

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

Barrett, L. F. (2009). Variety is the spice of life: A psychologist constructionst approach to understanding variability in emotion. *Cognition and Emotion*, *23*, 1284-1306.

Barrett, L.F., Barsalou, L.W., Lindquist, K., & Wilson-Mendenhall, C.D. (in preparation). The Conceptual Act Theory of Emotion.

Berridge, K.C., & Kringelbach, M.L. (2008). Affective neuroscience of pleasure: reward in humans and animals*. Psychopharmacology, 199*, 457-480.

Duncan, S., & Barrett, L.F. (2007). Affect as a form of cognition: A neurobiological analysis. *Cognition and Emotion, 21*, 1184-1211.

Eisenberger, N.I., & Lieberman, M.D. (2004). Why rejection hurts: A common neural alarm system for physical and social pain. *Trends in Cognitive Sciences, 8*(7), 294-300.

Eisenberger, N.I., & Lieberman, M.D., & Williams, K.D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science, 302*, 290-292.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C. & Raichle, M.E.

(2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Processing of the National Academy of Sciences, 102*(27), 9673-9678.

Pessoa, L. (2008). On the relation between emotion and cognition. *Nature Reviews Neuroscience, 9,*148-158.

Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., & Greicus, M.D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of Neuroscience, 27*(9), 2349-2356.

Wilson-Mendenhall, C.D., Barrett, L.F., Simmons, W.K., & Barsalou, L.W. (submitted). Grounding emotion in situated conceptualization.