View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Carolina Digital Repository

OXFORD

Cerebral Cortex, May 2016;26: 1910–1922

doi: 10.1093/cercor/bhv001
Advance Access Publication Date: 28 January 2015
Original Article

ORIGINAL ARTICLE

The Brain Basis of Positive and Negative Affect: Evidence from a Meta-Analysis of the Human Neuroimaging Literature

Kristen A. Lindquist[1,2,†], Ajay B. Satpute[3,†], Tor D. Wager[4], Jochen Weber[5], and Lisa Feldman Barrett[6,7]

[1]Department of Psychology, University of North Carolina, Chapel Hill, NC, USA, [2]Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, NC, USA, [3]Pomona College, Claremont, CA, USA, [4]Department of Psychology and Neuroscience, University of Colorado, Boulder, CO, USA, [5]Department of Psychology, Columbia University, New York, NY, USA, [6]Department of Psychology, Northeastern University, Boston, MA, USA, and [7]Department of Psychiatry and Radiology, Massachusetts General Hospital/Harvard Medical School/Martinos Center for Biomedical Imaging, Boston, MA, USA

Address correspondence to Kristen A. Lindquist, Department of Psychology, University of North Carolina, Davie Hall 321, CB # 3270, Chapel Hill, NC 27510, USA. Email: kristen.lindquist@unc.edu

[†]K.A.L. and A.B.S. contributed equally to this work.

Abstract

The ability to experience pleasant or unpleasant feelings or to represent objects as "positive" or "negative" is known as representing hedonic "valence." Although scientists overwhelmingly agree that valence is a basic psychological phenomenon, debate continues about how to best conceptualize it scientifically. We used a meta-analysis of 397 functional magnetic resonance imaging (fMRI) and positron emission tomography studies (containing 914 experimental contrasts and 6827 participants) to test 3 competing hypotheses about the brain basis of valence: the bipolarity hypothesis that positive and negative affect are supported by a brain system that monotonically increases and/or decreases along the valence dimension, the bivalent hypothesis that positive and negative affect are supported by independent brain systems, and the affective workspace hypothesis that positive and negative affect are supported by a flexible set of valence-general regions. We found little evidence for the bipolar or bivalent hypotheses. Findings instead supported the hypothesis that, at the level of brain activity measurable by fMRI, valence is flexibly implemented across instances by a set of valence-general limbic and paralimbic brain regions.

Key words: affect, meta-analysis, neuroimaging, valence, value

There is the good and the bad, the great and the low, the just and the unjust. . . all that will never change.

-Albert Camus

Introduction

Every person on the planet (barring illness) can tell good from bad, positive from negative, pleasure from displeasure. The basic ability to experience pleasant or unpleasant feelings and represent objects as positive or negative, or as pleasing or displeasing, is known as hedonic "valence." Valence is thought to be a fundamental, universal property of human experience. Representations of valence contribute to emotional feelings, morality, and personality. They are at the core of judgments about strangers, partners, friends, and leaders. Representations of valence inform decision-making, attitudes and preferences,

© The Author 2015. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com

and the valuation of goods and services (for a review see Barrett and Bliss-Moreau 2009). Cultures all over the world differ in the specific types of emotions they experience and recognize, but no culture lacks the concepts for valence (Osgood 1952; Wierzbicka 1992). Even infants as young as a few days old feel pleasure and discomfort (Lewis 1993) and can differentiate between pleasant and unpleasant facial expressions in others (Farroni et al. 2007). Although scientists overwhelmingly agree that valence is an integral aspect of most psychological phenomena, debate continues about how to best conceptualize the nature of valence (Barrett and Bliss-Moreau 2009). In this paper, we use evidence from functional magnetic resonance imaging (fMRI) to investigate how valence is represented as distributed brain activity in healthy human adults.

To date, research on the nature of valence has received the most attention within the psychological literature using behavioral responses and self-reported feelings as data. Dimensional methods (e.g., multidimensional scaling, factor analysis, and structural equation modeling) consistently reveal that valence is a fundamental property of self-reported affect and emotion when people view movies, images, and during momentary life experiences (for reviews, see Watson and Tellegen 1985; Larsen and Diener 1992; Feldman 1995; Barrett and Russell 1998; Russell and Barrett 1999; Bradley et al. 2001), but scientists disagree on the relation of positivity to negativity (all studies also identify a second property known as arousal, which refers to the degree of activation vs. quiescence that a person is feeling at a given moment). In some models, arousal is explicitly represented as its own dimension, and in other models it is not. Although arousal is an important aspect of affect, we do not focus on it in the present paper for several reasons. First, valence and arousal are difficult to separate in the context of experiments because most stimuli used to induce positivity or negativity also induce some change in arousal (e.g., the International Affective Picture System; Lang et al. 2005) (for a discussion, also see Kuppens et al. 2013; Lang and Bradley 2010). Second, valence and arousal are separable properties in some, but not all, individuals' self-reports of emotional experiences (Feldman 1995). Third, the concept of arousal, as a psychological property, is vague and underspecified. The term "arousal" is varyingly used to refer to enhanced attention, behavioral engagement, intensity of feeling, physiological activation, and subjective feelings of activation, and measures of each operationalization tend not to correlate with each other. As such, it is difficult to accurately quantify the degree of arousal that a given study is inducing in the context of a meta-analysis. Finally, fewer hypotheses about the nature of arousal have been put forth in the psychological literature to date than have models of valence). We use existing models of valence to test hypotheses about the brain's representation of valence.

Psychological Models of Valence: Predictions for the Brain Basis of Valence

The first psychological model of valence was initially put forth by Wundt (1897/1998) and hypothesizes that positivity and negativity constitute opposite ends of a single dimension; this view is referred to as the "bipolarity" hypothesis. Factor analyses and multidimensional scaling studies of humans' subjective experiences, perceptions of other people's facial movements and vocalizations, the semantic structure of emotion words, and the mathematics of measurement theory all support the view that positivity and negativity are bipolar opposites (Larsen and Diener 1992; Barrett and Russell 1999; Carroll et al. 1999).

A second psychological model of valence developed in the mid-twentieth century, based on evidence that self-reports of positive and negative affective experience are often uncorrelated. These data were taken as evidence that positivity and negativity might be independent dimensions; this idea is referred to as the "bivalence" hypothesis to indicate that there are 2 distinct constructs of valence that range from positivity-neutral and negativity-neutral (e.g., Watson and Tellegen 1985; Cacioppo et al. 1999; Norris et al. 2010). Whereas bipolarity was merely a descriptive hypothesis about conscious affect in reports of experience and perception, the bivalence view went further to hypothesize that there are separate physical systems for generating and representing positivity and negativity; evidence for bivalence includes self-report ratings (e.g., people can report experiencing positivity and negativity in response to the same stimulus or over the course of the same experimental trial; Larsen et al. 2001; Larsen et al. 2004) as well as biological data (e.g., relative differences in neurotransmitter activity for appetitive versus aversive stimuli, and evidence that spinal neurons can simultaneously cause activation of flexion and extraction muscles; Norris et al. 2010).

There has been a long and tortured debate over the structure of affect, largely because behavioral studies to date have been unable to show clear evidence for one model or the other (Barrett and Bliss-Moreau 2009). Bipolar and bivalence hypotheses are relatively untested in the domain of neuroscience, but each model makes unique predictions for how valence might be represented in neuronal activity. Support for the bipolarity hypothesis would be found if a given network of regions responds monotonically as affect changes from negative, to neutral, to positive or vice versa. In this view, neurons associated with increased positive affect would also be associated with reduced negative affect, and vice versa. Support for the bivalence hypothesis would be found in separate and independent networks for positivity and negativity, such that across studies, the same regions show consistent increases in activity for positive but not negative affect, and other regions show consistent increases in activity for negative but not positive affect. The networks would be independent insofar as information from neural areas responsive during negative affect would not provide any information about the state of activity from neural areas responsive during positive affect—neurons representing negativity would not systematically change firing rates based on increases or decreases in positivity, and vice versa. Technically, the networks "could" be negatively correlated in a given study and such a correlation would mean the 2 systems are reciprocally activating. Reciprocal activation would remain evidence for bivalence as long as the networks were not spatially dependent; if the networks are truly independent, then it should be possible for co-activation to occur sometimes (Berntson et al. 1991).

Whereas the bipolar and bivalence hypotheses were primarily formulated based on behavioral and self-report data, the goal of this paper is to test a hypothesis about the structure of valence inspired by theory and research from a neuroscience perspective. For example, neuroscience findings in non-human animals suggest a more nuanced view than either the bipolar or bivalent model affords. On the one hand, there are studies demonstrating that specific neurons respond to positive affect and specific neurons respond to negative affect. For instance, using single-cell recordings, studies observed cells distributed throughout the monkey amygdala (Paton et al. 2006; Belova et al. 2008) and orbitofrontal cortex (area 13) (Morrison and Salzman 2009) that respond relatively more for stimuli associated with reward (e.g., juice) than stimuli associated with aversion (e.g., air puffs). Similar patterns are observed whether using stimuli that are primary reinforcers (e.g., juice and air puffs) or stimuli conditioned to be associated with primary reinforcers (e.g., neutral visual stimuli)

(Belova et al. 2008). In some studies, there is even evidence for inhibitory relationships between positive-encoding and negative-encoding cells, such that in the presence of a positive stimulus (e.g., juice), positive-encoding amygdala cells increase their rate of firing, whereas negative-encoding cells decrease their rate of firing (Belova et al. 2008).

Yet even amidst evidence for functional selectivity at the cellular level, there exists evidence for more flexible cellular encoding of valence, suggesting that a strict relationship between positivity or negativity and single cells is not ubiquitous throughout the brain. For instance, certain cells in the monkey OFC (area 13) respond equally to appetitive and aversive stimuli (Morrison and Salzman 2009). Even those cells that show a preference for stimuli associated with reward (e.g., juice) also sometimes respond to stimuli associated with aversion (e.g., an air puff) (and vice versa) (Morrison and Salzman 2009). Findings from rats, and even the nematode *Caenorhabditis elegans*, also demonstrate the existence of cells with a valence-general response profile. In rats that are in neutral, safe circumstances, glutamate disruptions to rostral portions of the nucleus accumbens shell generate appetitive behaviors, whereas glutamate disruptions to caudal portions of the shell generate aversive responses. Yet, when rats are in a threatening context, glutamate disruptions to the same rostral cells begin to generate aversive responses instead of appetitive responses (Reynolds and Berridge 2008). Whether neurons code for approach versus avoidance behavior shifts according to context, even in *C. elegans*, a nematode with only 302 neurons. For instance, the olfactory neuron AWCON directs approach and avoidance behavior toward the same exact odor depending on the presence or absence of other neurochemicals in the brain (Tsunozaki et al. 2008). These findings might suggest that it is not neurons, but neurochemicals, that are valence specific, but research shows that neurochemicals such as opiods and dopamine are general to both pleasure and pain in mammals (Leknes and Tracey 2008).

Taken together, the findings from non-human animals imply that a third hypothesis on the structure of valence is possible: A representation of positivity or negativity emerges at the population level, as a "brain state" (Salzman and Fusi 2010) but is not necessarily consistently associated with a specific brain region or set of regions. Using the logic of large-scale brain organization that is available from neuroimaging studies (Smith et al. 2009; Biswal et al. 2010; Poldrack 2010; Yeo et al. 2011), this third hypothesis is a valence-general "affective workspace" hypothesis (Barrett and Bliss-Moreau 2009). We use the term "affective workspace" in a manner akin to Edelman's "neural reference space" (Edelman 1989). According to Edelman, a neural reference space is a set of neurons that are probabilistically involved in realizing a class of mental events (such as positivity or negativity) (for a discussion, also see Lindquist, Wager, Kober et al. 2012)—it is the neuronal workspace in which a mental state is likely to be implemented when it is experienced. In this view, on any given occasion, voxels of neurons are functionally selective for positivity or negativity, even if they do not consistently show increased activation exclusively for one or the other. Different instances of positivity and negativity are implemented dynamically as flexible neuronal assemblies within the same neural reference space. A given neuron might participate in both instances of negativity and positivity across contexts, with its receptive field being determined by the neural context. Because neuronal assemblies are flexible, a given neuron need not participate in every brain state within a class (e.g., positivity), or even in the exact same mental state at 2 different points in time (e.g., positivity at seeing a friend at work vs. at a pub). Such valence-general flexibility can account for both the behavioral evidence of bivalence and bipolarity (Barrett and Bliss-Moreau 2009).

To test the 3 hypotheses about the structure of valence in humans, we summarize almost 20 years of human neuroimaging studies (using functional magnetic resonance imaging; fMRI and positron emission tomography; PET). Several early attempts (e.g., Murphy et al. 2003; Wager et al. 2003; Kringelbach and Rolls 2004) covering the first 10 years or so of neuroimaging data (87, 65, 106 studies, respectively) examined which brain areas respond more frequently during positive versus negative affect. More recently, several meta-analyses assessed the neural correlates of subjective pleasantness (Kuhn and Gallinat 2012; 40 studies), reward-related decision-making (Liu et al. 2011; 142 studies) and the experience of subjective value during economic decisions (Bartra et al. 2013; 206 studies; Clithero and Rangel 2013; 81 studies).

Our meta-analysis is distinct from these other meta-analyses in several ways. First, our database of 397 studies of affective (i.e., positive and negative) and discrete emotional (i.e., anger, disgust, fear, sadness, happiness, etc.) experiences and perceptions (spanning January 1993–December 2011) is the largest existing database of neuroimaging studies on valence. Second, our meta-analysis is distinct from at least the most recent meta-analyses because it does not include studies of reward/loss. Other recent meta-analyses summarized studies of liking and attractiveness (Kuhn and Gallinat 2012) or focused explicitly on reward (Liu et al. 2011; Bartra et al. 2013; Clithero and Rangel 2013). Our database explicitly excludes studies of reward since the motivational processes underlying reward are thought to be distinct from the experience of pleasure per se (Robinson and Berridge 2013).

Finally, our meta-analysis is distinct because no previous meta-analyses were designed to compare different theoretical formulations on the nature of valence. We used our database to test whether valence is supported by bipolar brain systems, bivalent systems, or whether the brain regions that represent valence do so in a flexible context-specific manner consistent with the idea of a valence-general affective workspace.

Materials and Methods

Database

The database included neuroimaging studies of affective (i.e., positive and negative) and discrete emotional (i.e., anger, disgust, fear, sadness, happiness, etc.) experiences and perceptions published between January 1993 and December 2011. We sampled potential papers for our database using search criteria that have been reported elsewhere (Kober et al. 2008; Wager et al. 2008). We then added papers by searching the tables of contents of journals publishing neuroimaging research and/or research on emotion and affect. Each study contrast was characterized based on a variety of features (e.g., sample size, gender of participants, PET or MRI imaging modality, stimulus modality, valence, analysis type, etc.; see Kober et al. 2008; Wager et al. 2008; Lindquist, Wager, Bliss-Moreau et al. 2012) by 2 researchers; any disagreements between researchers about the characterization of each contrast were resolved through discussion. Notably, some studies explicitly assessed valence (e.g., compared neural responses to positive and negative stimuli), whereas others assessed valence as part of a discrete emotional experience or perception (e.g., compared neural responses to happy vs. fearful stimuli). Discrete emotion categories were qualified as positive or negative based on their typical location on the valence dimension in the circumplex model of affect (Russell and Barrett 1999;

Table 1 Frequency of contrasts used in the database for the MKDA

Contrasts	Frequency
Positive versus neutral	110
Negative versus neutral	255
Positive versus negative	36
Negative versus positive	45

Wilson-Mendenhall et al. 2014). The resulting database includes 397 studies containing 6827 participants and reports peak coordinates from 914 contrasts that compared neural activity during presentation of positive or negative affective stimuli (e.g., images, film clips, imagery, facial/vocal/bodily expressions, sounds, odors, etc.) either against each other or against neutral stimuli.

We analyzed only those study contrasts that were relevant for examining brain regions associated with positive and negative affective responses (see Table 1). We limited our analysis to study contrasts that reported whole-brain analyses or a combination of whole-brain and region of interest (ROI) analyses. We excluded study contrasts that reported ROI analyses only (e.g., study contrasts in which researchers investigated only the amygdala and did not report whole-brain analyses). We further excluded contrasts in which the baseline condition involved merely fixation or in which the baseline condition involved an altogether different class of stimuli, such as comparing smiling (positive) or sneering (negative) faces to abstract shapes or to non-face stimuli.

Multilevel Peak Kernel Density Analysis

The Multilevel Peak Kernel Density Analysis (MKDA) (Wager et al. 2007) (software available from http://wagerlab.colorado.edu/tools) summarizes the spatial overlap of peak coordinates reported in individual studies to reveal voxels that consistently show increases in brain activity during a particular class of psychological events (e.g., positivity) relative to some baseline. For more information about the MKDA and its validation against other coordinate-based meta-analytic methods such as the Activation Likelihood Estimation technique or "image-based" methods that use statistical maps from individual studies, see Wager et al. (2007); Kober et al. (2008); Salimi-Khorshidi et al. (2009); Kober and Wager (2010).

Following the standard MKDA procedure used in other published studies, peak coordinates from each study contrast in the database were first convolved with 12-mm spheres to form binary indicator maps. Rather than treating peak coordinates from individual studies as the unit of analysis, the MKDA treats study contrasts as the unit of analysis to prevent a single study that reports many peaks (because of more liberal thresholding or differences in statistical power in the study) from unduly biasing the results (Wager et al. 2007). To control for quality of the data entering the meta-analysis, study contrast maps are further weighted by the square root of the sample size and studies using fixed-effects analyses are down-weighted by 0.75 so that less rigorous statistical analyses contribute less strongly to the meta-analytic results. The MKDA creates binary indicator maps from peak coordinates, rather than using voxel-level z scores, and weighs the contribution of coordinates by study-level variables (sample size and rigor of analysis) rather than voxel-level variables (z-scores) for several reasons. First, between-study factors such as preprocessing decisions or the analysis method used can impact z scores, making them incomparable between studies. Second, the sample size of a study impacts variance estimates, with the result that smaller, more variable studies can have inflated z-scores. Weighting voxels by z-score would thus allow smaller, high variance studies to weigh more heavily toward the meta-analysis results. Treating all peak coordinates equally but weighing them by study-level factors such as the sample size and rigor of the statistical methods ensures that studies more likely to be diagnostic of the population at large contribute more strongly to the meta-analytic results. For further discussion of the implications of using z-scores versus binary indicator maps based on peak coordinates, see Wager et al. (2007) and Kober et al. (2008).

The next step in the MKDA is to compute a point estimate of the proportion of study contrasts that reported increased activation near each voxel in the brain. The MKDA uses the statistic P for each voxel, or the proportion of study contrasts in the database reporting activation near that voxel. MKDA maps were generated for study contrasts comparing: 1) "positive affect" versus "neutral," (ii) "negative affect" versus "neutral," (iii) "positive affect" versus "negative affect," and (iv) "negative affect" versus "positive affect" task conditions.

Finally, we computed meta-analytic contrasts and conjunctions to test the bivalent, bipolar, and affective workspace hypotheses (described in more detail later). To determine significance for comparisons, a Monte Carlo simulation with 5000 iterations was performed to produce a null distribution of probabilities that a given study contrast activated near a voxel of the brain. To produce a null distribution with similar characteristics to the database, the Monte Carlo simulation randomly assigned the center of clusters to different locations in the brain (excluding the ventricles and white matter) while still preserving the total number of study contrasts and coordinates within those contrasts. For each iteration, the MKDA map was calculated and the probability of observing a given proportion of study contrasts activating near a given voxel was calculated. Using the null distribution produced by the Monte Carlo simulation, we then obtained results corrected for multiple comparisons across the whole brain by using 3 levels of correction: for voxel-wise analyses, we used a false discovery rate (FDR) threshold, and for cluster-level analyses, we used family-wise error rate (FWER) thresholds to assess the cluster-size required given voxels that were significant at $P < 0.01$ and $P < 0.05$.

For conjunction analyses, we first thresholded each map individually using the FWER-corrected thresholds and then examined the overlap using a "global null" conjunction analysis that defines the conjunction as the intersection of individually thresholded maps (as in Nichols et al. 2005). Our global null conjunction was constrained on whether the conjunction of the 2 maps contained at least 12 contiguous voxels.

Testing the Bipolarity Hypothesis

The bipolarity hypothesis can be operationalized as a set of brain regions that respond monotonically across affective valence. Areas exhibiting activity that correlate with either of the following 2 patterns: negative affect > neutral > positive affect or positive affect > neutral > negative affect would provide neural evidence consistent with the bipolarity hypothesis. Part of this ordinal relationship involves showing that areas with an increase in activity during negative affect also show a decrease in activity during positive affect (relative to neutral). However, we were unable to test whether "deactivations" for positive affect or negative affect relative to neutral were present because our meta-analytic

database did not include any neutral > negative or neutral > positive contrasts. Like other meta-analyses of neuroimaging studies (e.g., Vytal and Hamann 2010; Lindquist, Wager, Kober et al. 2012), we did not include contrasts reporting such neural "deactivations" in our database because they are reported infrequently across individual studies. We also did not include "deactivations" due to problems of interpretation; a "deactivation" is really a situation where the experimental condition shows less activity than the control condition and so is dependent upon the nature of the neural responses during the control (i.e., they might not represent "deactivation" per se; for a discussion, see Lindquist, Wager, Bliss-Moreau et al. 2012). We were thus unable to provide a complete test of the bipolarity hypothesis in the context of a meta-analysis because it required examining decreases in activity relative to neutral affect. Given this limitation (which is inherent to the studies in the literature themselves, and not to our meta-analytic methods per se), we opted to search for regions that would meet the bipolarity hypothesis as closely as was feasible. That is, we searched for regions that were consistent with an "ordinal relationship" with valence (e.g., positive > neutral > negative or negative> neutral > positive).

To do so, we first masked out voxels that were "valence general" (i.e., responding more frequently during both positive affect and negative affect relative to neutral affect), since these voxels already violate the monotonic relationship predicted by bipolarity. Next, we examined whether there were any brain regions in which there was greater activity when one type of valence was compared against another as opposed to against neutral (e.g., "positive affect" vs. "negative affect" > "positive affect" vs. "neutral"). The logic here was that if positive and negative are truly bipolar opposites, then activations would be more likely to be observed when contrasting them against one another than when, say, contrasting positive versus neutral. In essence, this sort of contrast assumes that the neural differences between positive and negative affect are greater than the neural differences between negative and neutral affect or positive and neutral affect, which would be consistent with bipolarity.

Testing the Bivalence Hypothesis

The bivalent hypothesis can be operationalized as a 2 sets of brain regions: one set that responds during negative affect and another set that responds during positive affect. Activity in these brain regions should, at least in principle, be independent of one another. Areas exhibiting activity that correlates with either of the following 2 patterns: negative affect > neutral and no difference in response during positive affect, or positive affect > neutral and no difference in response during negative affect, would provide neural evidence consistent with the bivalent hypothesis. Part of these patterns involves showing that a region that responds more during negative affect also has no increase during positive affect, and vice versa. In the context of a meta-analysis, there are 2 ways of testing this hypothesis. One approach would be to first exclude voxels that are valence general (i.e., respond more during positive affect vs. neutral and negative affect vs. neutral), since these would violate the bivalent hypothesis at the outset and then to search outside of these areas for voxels that are more frequently engaged during "positive affect" versus "neutral" contrasts, and other voxels that are more frequently engaged during "negative affect" versus "neutral" contrasts.

However, we opted for a more conservative approach that may provide a better test of the bivalence hypothesis. First, we excluded voxels that were valence general. Because it is particularly critical that a region that responds during negative affect certainly does not respond during positive affect, we then compared "negative affect" versus "positive affect" directly and further compared the probability of these activations to study contrasts of "positive affect" versus "neutral affect." This is a more conservative test of the bivalence hypothesis because a region that is truly selective for negative affect should be more likely to be significantly active during negative affect versus positive affect even controlling for any chance activations that occur during positive versus neutral. We also performed a complimentary analysis for positive affect—that is, we compared the probability that a voxel was active during "positive affect" versus "negative affect" contrasts beyond "negative affect" versus "neutral" contrasts. Voxels meeting these criteria could be considered to be uniquely sensitive to one type of valence and not the other.

Testing the Affective Workspace Hypothesis

Identifying Valence-General Voxels

Valence-general voxels were identified as those consistently activated during both positive and negative valence across studies (i.e., were significant in the global null conjunction of positive versus neutral and negative versus neutral contrasts; Nichols et al. 2005).

Identifying Voxels within the Valence-General Affective Workspace that Show Preference for Positive versus Negative or Negative versus Positive Affect

It is possible that even among voxels that respond more frequently to positive and negative affect than neutral affect, there exist voxels that show a "relative" preference for positive versus negative affect, or negative versus to positive affect. To assess this possibility, we addressed whether any voxels within the valence-general affective workspace (the conjunction of positive > neutral and negative > neutral) had more frequent activity during the "positive" versus "negative" study contrasts or "negative" versus "positive" study contrasts in our database. We note that these study contrasts were independent of the contrasts used to identify the valence-general workspace, providing a rigorous means of testing this hypothesis.

Identifying Distributed Patterns for Positivity and Negativity

Finally, we performed a classification analysis using a linear support vector machine (SVM) to test whether information about valence is contained within patterns of neural activation distributed across the whole-brain space (implemented with the libSVM toolbox: http://www.csie.ntu.edu.tw/~cjlin/libsvm) (Chang and Lin 2011). As in the MKDA, indicator maps were generated by diluting peak coordinates from a contrast with a 12-mm binary sphere. Matrices were transformed into vectors that served as the features for the SVM. The model was implemented (cost parameter = 1) to classify whether a given study contrast map involved a comparison of positive versus neutral affect or negative versus neutral affect. For training, a random selection of 80 "positive" versus "neutral" study contrasts and 80 "negative" versus "neutral" study contrasts were used, and for testing a random selection of 20 "positive" versus "neutral" and 20 "negative" versus "neutral" different study contrasts were used. The training and testing steps were repeated 100 times, and classification accuracy (i.e., recall) and precision were calculated for tests of both positive and negative contrasts. Mean recall and precision were then calculated. Recall refers to the proportion of contrasts that were correctly classified as positive (or negative) out of the group of truly positive (or negative) contrasts. In contrast, precision refers to the proportion of truly positive (or negative)

contrasts out of the group of contrasts that were classified as positive (or negative). Evidence for the above chance classification accuracy was tested using a binomial probability distribution.

Results

The Bipolar Hypothesis: Regions that Respond Monotonically along a Single Valence Dimension?

First, we assessed whether any clusters of voxels were more frequently engaged during "positive" versus "negative" study contrasts than during "positive" versus "neutral" study contrasts across studies (while excluding any voxels that were valence general; also see Materials and Methods for our operationalization of the bipolarity hypothesis). This analysis revealed a cluster in a ventral portion of the rostral anterior cingulate cortex (ACC) and medial prefrontal cortex (MPFC) (MNI = [9, 39, −9], $k = 178$) (Fig. 1; Table 2). We next tested for clusters of voxels that were more often engaged during "negative" versus "positive" than by "negative" versus "neutral" study contrasts (while excluding voxels that were valence general) but were unsuccessful in identifying any. These findings suggest that the ventral MPFC and ACC areas may be candidate regions of interest coding for valence along the lines specified by the bipolarity hypothesis.

The Bivalence Hypothesis: Two Unipolar Dimensions?

First, we tested for voxels that responded exclusively to positive affect (i.e., a unipolar dimension ranging from positive to neutral), by assessing whether any voxels were more frequently engaged during "positive" versus "negative" study contrasts than during "negative" versus "neutral" study contrasts. Contrary to the bivalence hypothesis, no voxels displayed a significant profile of increased activation exclusively for positivity across studies. Next, we performed a complimentary analysis to test for voxels that responded selectively to negative affect (i.e., a second unipolar dimension); we were again unsuccessful in identifying any. These findings suggest that the bivalence view that positivity and negativity correspond to spatially separable and distinct brain systems is not a viable framework for understanding the brain basis of valence.

The Affective Workspace Hypothesis

Valence-General Voxels

We found the conjunction of voxels that showed consistent increases in activation during study contrasts comparing "positive" versus "neutral" baselines and "negative" versus "neutral" baselines using the global null conjunction (Nichols et al. 2005). In essence, these are regions of the brain that respond more frequently to positive AND negative valence than to neutral valence. Figure 2 shows the maps for "positive" versus "neutral" and "negative" versus "neutral" contrasts as well as their conjunction (also see Table 2; see Supplementary Fig. 1 for sagittal, coronal, and transaxial views of valence-general regions). Consistent with the hypothesis that valence-general voxels make up the brain's affective workspace, our conjunction revealed valence general increases in activity in the bilateral anterior insula, bilateral lateral orbitofrontal cortex, bilateral amygdala, the ventral striatum, thalamus, dorsomedial prefrontal cortex (~BA 9), dorsal ACC, supplementary motor area (~BA 6), bilateral ventrolateral prefrontal cortex, and lateral portions of the right temporal/occipital cortex. This set of regions has been referred to as a "salience network," on the assumption that it is involved in representing the body's reaction to affective stimuli in the environment (Seeley et al. 2007). In our view, the name "salience network" does not imply that there is one process termed "salience" that is being performed by this network. Rather, based on its anatomical connections, we hypothesize that this network might serve as a body-based form of attention that contributes flexibly to a wide variety of self-relevant mental phenomena (including, positive and negative "emotions" but also "cognitions" such as goal-oriented visual attention; Lindquist and Barrett 2012).

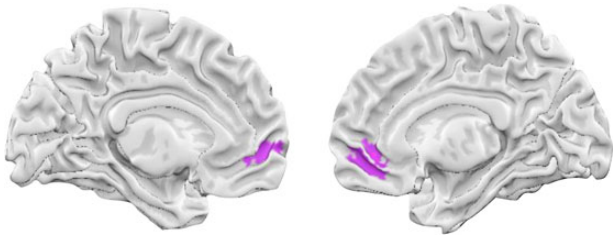

Figure 1. Neural activity in ventromedial prefrontal cortex. The figure illustrates portions of the ventromedial prefrontal cortex that responded as predicted by a bipolar model of affect. The bipolar model predicts the presence of neural regions that respond more frequently during positive versus negative contrasts than during positive versus neutral contrasts. No regions were observed that responded more during negative versus positive contrasts than during negative versus neutral. The activations shown are FWER corrected for clusters observed across the whole brain. The maximum voxel was also observed at voxel-wise FDR correction across the whole brain (MNI = [9, 39, −9]).

Voxels within the Valence-General Affective Workspace that Show Preference for Positive versus Negative Affect or Negative versus Positive Affect

Next, we searched within the valence-general affective workspace for voxels that, despite being valence-general, were relatively more likely across studies to show frequent activity for one type of valence than another. These voxels can be thought of as voxels that show a preference to one type of valence versus another, despite the fact that they on the whole respond to both positive and negative affect more frequently than neutral affect. To do so, we first sought voxels within the previously identified valence-general affective workspace (the conjunction of positive > neutral and negative > neutral) that had more frequent activity during the "positive" versus "negative" contrasts in our database. We note that these contrasts were independent of the contrasts used to identify the valence-general workspace, providing a rigorous means of testing this hypothesis.

No voxels within the valence-general affective workspace were relatively more likely to respond to positive than negative affect across studies in the literature. We next searched for voxels within the valence-general affective workspace that had more frequent activity during the "negative" versus "positive" contrasts in our database and observed activation in the left amygdala and both ventral and dorsal portions of the left anterior insula (Fig. 3 and Table 2). Voxels within the left amygdala and left anterior insula are therefore relatively more likely to show increased activation during negative than positive affect despite the fact that they respond to both positive and negative affect more so than neutral affect. Importantly, this relative distinction does not suggest an absolute distinction, insofar as the voxels identified consistently respond to both positive and negative affect more so than neutral affect across studies.

Distributed Patterns for Positivity and Negativity?

Even if individual brain areas are not functionally selective for positivity or negativity, it is possible that patterns of activity across the brain might reveal evidence for regions with a

Table 2 Meta-analytic results

Valence-general regions: intersection of (positive > neutral) and (negative > neutral) contrasts					Proportion of contrasts			
Region	x	y	z	FDR	Positive versus neutral		Negative versus neutral	
					Proportion	SE	Proportion	SE
Dorsomedial Prefrontal cortex	−9	51	33		0.10	0.028	0.12	0.020
	0	60	27		0.12	0.031	0.13	0.021
	−9	57	21		0.11	0.030	0.11	0.020
Ventromedial prefrontal cortex/rostral cingulate cortex	−3	39	0		0.15	0.034	0.09	0.018
Dorsal cingulate cortex	0	15	48		0.14	0.033	0.10	0.019
Supplementary motor area	0	15	57		0.15	0.034	0.12	0.021
Inferior frontal gyrus	48	12	30		0.13	0.032	0.12	0.021
	−48	21	18		0.10	0.029	0.12	0.020
	48	21	15		0.10	0.029	0.09	0.018
	48	30	3		0.13	0.032	0.13	0.021
Declive	42	−60	−21		0.10	0.028	0.13	0.021
Ventral anterior insula extending into lateral orbitofrontal cortex	39	33	−6		0.11	0.029	0.10	0.019
	33	21	−6		0.11	0.029	0.10	0.019
	42	24	−9		0.11	0.030	0.14	0.022
	−39	24	−12	Yes	0.21	0.039	0.20	0.025
	−27	15	−18		0.13	0.032	0.13	0.021
Insula/claustrum	33	9	−9		0.11	0.030	0.10	0.019
Middle temporal gyrus	48	−60	3		0.14	0.033	0.15	0.022
Superior temporal gyrus	48	15	−9		0.11	0.030	0.10	0.019
	48	6	−15		0.11	0.029	0.09	0.018
Inferior occipital cortex	42	−69	−9		0.12	0.031	0.12	0.020
Amygdala	24	3	−18	Yes	0.22	0.039	0.21	0.026
	−27	−6	−18	Yes	0.22	0.039	0.26	0.027
Midbrain	−9	−6	−12		0.10	0.029	0.12	0.021
Thalamus	6	−24	0		0.10	0.029	0.12	0.020
	−6	−9	−3		0.11	0.030	0.11	0.020
	−6	−24	−3		0.09	0.027	0.12	0.021
	3	−6	−6		0.10	0.029	0.10	0.019
Nucleus accumbens	−9	−12	−6		0.09	0.027	0.05	0.014
Valence-general regions with a preference for negativity: (negative > positive) masked within valence-general areas					Negative versus Positive			
					Proportion	SE		
Amygdala	−18	−3	−24	Yes	0.22	0.062		
Ventral anterior insula	−33	24	12	Yes	0.22	0.062		
Anterior insula	−30	21	−3		0.18	0.058		
Dorsal insula	−36	12	12		0.15	0.053		
Middle frontal gyrus	−36	33	9		0.13	0.049		
Bipolar regions (positive > negative) > (positive > neutral) masking out valence-general regions					(Positive versus negative) versus (positive versus neutral)			
					Proportion	SE		
Ventromedial prefrontal cortex	9	39	−9	Yes	0.18	0.00		
	−18	42	−12	Yes	0.16	0.00		

Note: the table lists selected activation peaks for global and local maxima. SE refers to standard error. The statistical map is also available for download upon request.

functional preference during positivity or negativity. To explore this possibility, we performed a classification analysis across the whole brain using an SVM (Chang and Lin 2011) on the study contrast maps investigating "negative" versus "neutral" and "positive" versus "neutral" conditions. Average classification accuracy was 52% (accuracy for positive contrasts = 52%, accuracy for negative contrasts = 52%), and average classification precision was 53% (precision for positive contrasts = 51%, precision for negative contrasts = 54%). Classification accuracy was not significantly greater than chance ($P > 0.6$), meaning that the classifier was unable to diagnose whether individual contrast maps involved either a "negative" versus "neutral" or a "positive" versus "neutral" comparison. Together, these findings suggest that the distribution of activity in voxels across the brain is also insufficient to classify valence. In combination with the other findings, the affective workspace may be best considered as valence general; brain regions may flexibly and interchangeably represent both positivity and negativity across different instances as we discuss later.

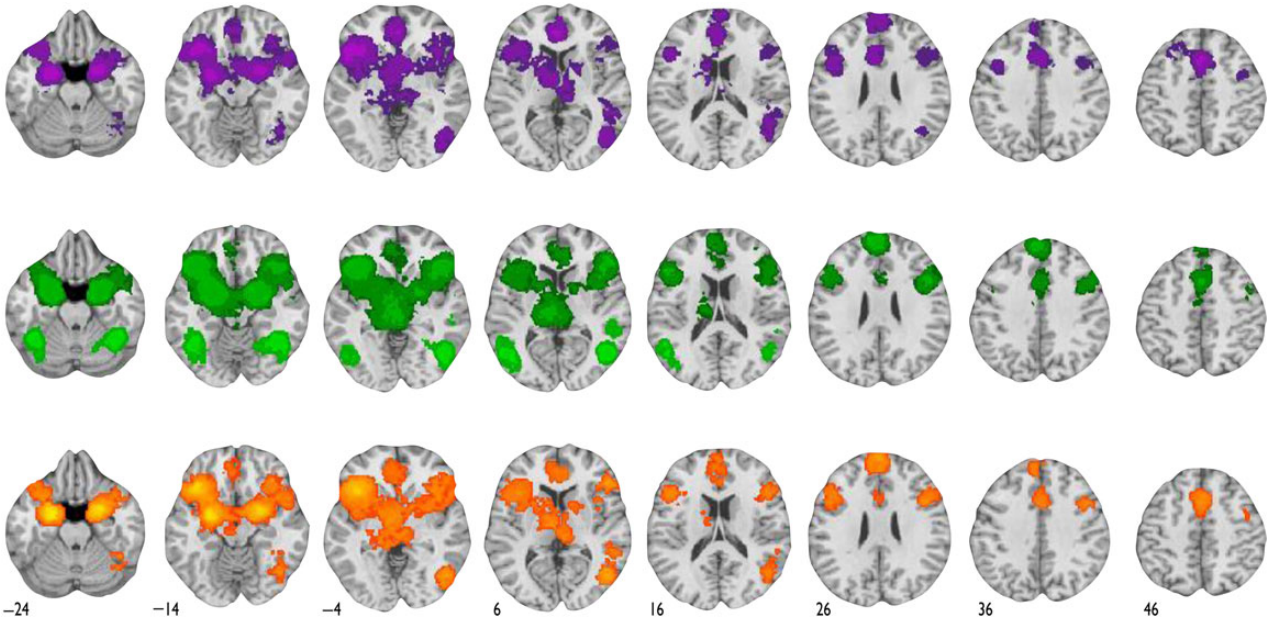

Figure 2. Neural regions consistently associated with valence-general activations. The top and middle rows of the figure illustrate regions that were frequently correlated with positive versus neutral study contrasts (illustrated in purple) and negative versus neutral study contrasts (illustrated in green), respectively. For these maps, the color codes reflect whole-brain statistical correction at the voxel level (bright purple and bright green) using FDR procedures, or at 2 different cluster-level thresholds (given voxel-level *P*-values of 0.05 and 0.01; see Materials and Methods) using FWER procedures. The bottom row of the figure illustrates valence-general neural regions. These regions were frequently engaged by both positive versus neutral study contrasts and by negative versus neutral study contrasts, as revealed by a conjunction analysis. In common, contrasts that compare positive or negative valence with neutral baselines engage several regions including the dorsomedial prefrontal cortex, supplementary motor area, ventrolateral prefrontal cortex, anterior insula, amygdala, ventral striatum, and thalamus. See Supplementary Figure 1 for additional views of the valence-general regions.

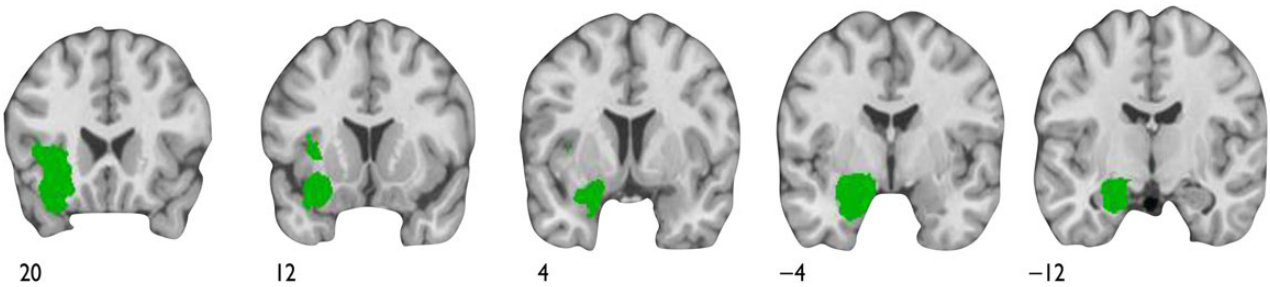

Figure 3. Neural regions exhibiting a preference for negative affect. The figure shows that portions of the left anterior insula and amygdala extending into the anterior hippocampus exhibited more frequent preference to negative affect among regions already shown to respond in a valence-general fashion. The analysis examined negative versus positive study contrasts while masking in regions that were already active in the intersection between negative versus neutral contrasts and positive versus neutral contrasts. The activations shown are FWER corrected for clusters observed across the whole brain. A voxel in the left amygdala was also observed at voxel-wise FDR correction across the whole brain (MNI = [−18, −3, −24]).

Discussion

Our meta-analysis of the neuroimaging literature on positive and negative affect is the most comprehensive to date and helps to answer questions about valence that have existed in the psychological literature since the 1960s. Although our findings can only speak directly to function at the organizational level of topographical regions across the human brain, it remains a possibility that function extends beyond this level of analysis.

Little Evidence for Bipolar or Bivalent Models of Neural Function

Our meta-analysis of neuroimaging studies was the first to assess bipolar versus bivalent models of affect and revealed little to no support for either model. Our findings suggest that portions of the ventromedial prefrontal cortex and ACC may serve as candidate ROIs for the bipolarity hypothesis but did not reveal any other brain regions that responded in a bipolar manner. VMPFC/ACC showed greater differences between "positive affect" versus "negative affect" than it did "positive affect" versus "neutral affect." This finding suggests there is more dissimilarity between brain responses to positive and negative affect than positive and neutral affect in this region, consistent with a bipolar view. However, we refer to these areas as candidate ROIs because we were only able to test whether this area increased more as positive affect increased (relative to different baselines), but we were unable to test whether it also showed decreasing activity during negative affect. Future studies that incorporate active, yet neutral, baselines would be helpful to fully test the bipolarity hypothesis. The need for active baseline conditions that provide a suitable

comparison to conditions of interest is not new in neuroimaging (Stark and Squire 2001).

The link of VMPFC to positive affect is consistent with other recent findings. Another recent meta-analysis found relatively more frequent activity in a similar region of VMPFC during positive as compared with negative feelings (but did not exclude valence-general regions; Roy et al. 2012). Regions on the medial orbital surface of the brain (which are slightly more ventral still to the region we observed) are also associated with the representation of reward (Kringelbach and Rolls 2004). A recent study observed that activity in medial OFC, slightly ventral still to the region we observed in our analysis, parametrically increased as participants' self-reported ratings of unpleasantness to evocative scenarios decreased and ratings of pleasure increased (Wilson-Mendenhall et al. 2014). However, another recent study observed that mean VMPFC/mOFC activity correlates with increasing ratings of both the pleasantness and unpleasantness of evocative pictures (Chikazoe et al. 2014).

In contrast, we found no neuroimaging support for the bivalence hypothesis. Of note, both the bipolar and bivalent hypotheses were developed based on human behavioral data and may not translate cleanly into hypotheses about large-scale functional brain activity. Indeed, alternative operational definitions of these hypotheses may yield different analytical tests. We caution, however, that it is certainly possible to observe greater activity in brain regions for either positive relative to negative affect or negative relative to positive affect (Liu et al. 2011; Bartra et al. 2013), but this is only evidence that a brain region has a "relative" preference for one type of valence over another and is not support for the bivalent hypothesis that independent brain systems support positivity and negativity.

A Valence-General Affective Workspace

The bulk of our meta-analytic evidence fell in support of a flexible affective workspace that correlates with both positive and negative valence across instances. These findings suggest that, at the level of regional brain activity, there is no single region or even voxel that uniquely represents positivity or negativity. Limbic tissue, including the anterior insula, rostral ACC/ventromedial prefrontal cortex, dorsal ACC, amygdala, ventral striatum, as well as several other regions including the thalamus and occipitotemporal cortex, appears to contain cells that are part of the brain's valence-general affective workspace or "affective neural reference space" (Barrett and Bliss-Moreau 2009). Consistent with our findings, these regions consistently show increased activity across neuroimaging studies when affective valence is being represented during emotion experience and perception, pain, aversion, and orgasm (for a review, see Lindquist and Barrett 2012). Importantly, regions in this workspace are involved in representing and regulating activity in the viscera (i.e., interoceptive cues; Craig 2009). Indeed, the affective workspace is routinely engaged not just when people experience an affective feeling but even in so-called cognitive states when internal sensations in the body, including afferent signals and central nervous system representations, are used to guide the allocation of attention (Corbetta et al. 2008). Given that interoceptive information is the basis of affective feeling and changes across a wide range of mental states, this finding is consistent with the hypothesis that every conscious moment has some affective tone (Wundt 1897/1998); circuitry within the affective workspace may infuse each and every conscious moment with some degree of positivity or negativity (for a review, see Craig 2009; Lindquist and Barrett 2012).

Although no other meta-analyses of neuroimaging data have explicitly compared different models of valence, our valence-general findings are largely consistent with other existing meta-analyses of valence. For instance, like our own meta-analysis, Murphy et al. (2003) did not find spatial differentiation between the brain activity correlated with experiencing and perceiving positive and negative emotions. The more recent meta-analysis of reward and loss performed by Liu et al. (2011) revealed a valence-general affective workspace that is largely similar to our own. In particular, Liu et al. (2011) observed activity in the dorsomedial prefrontal cortex, medial orbitofrontal cortex, amygdala, insula, ACC, ventral striatum, brainstem, and thalamus when participants were anticipating events involving reward/loss, experiencing reward/loss or evaluating a reward/loss. Where Liu et al.'s (2011) findings differ from our own is in their observation of activity within the posterior cingulate cortex. Posterior cingulate cortex is part of a network involved in projecting oneself into the future (Buckner and Carroll 2007) and so it is possible that this brain area is linked to the anticipation of reward across studies, but not valence per se. Other recent meta-analyses that examined valence in the context of reward-related tasks (Kuhn and Gallinat 2012; Bartra et al. 2013; Clithero and Rangel 2013) also identified a similar cluster within posterior cingulate cortex. Our meta-analysis may not have observed this cluster because our database intentionally did not include studies of reward or loss/punishment (on the basis that these are likely distinct phenomena from the representation of valence per se). The fact that we did not include studies of reward versus loss/punishment in our database could also explain why our findings do not replicate those of Kringelbach and Rolls (2004), who focused exclusively on the OFC and identified a medial to lateral spatial trend that corresponded to studies of monetary gain versus loss.

Our meta-analytic neuroimaging findings, which have both temporal and spatial limitations, are perhaps most importantly convergent with evidence from other methods including electroencephalography, lesion, electrical stimulation, and neurochemical studies in humans and non-human animals. For example, the human P300, an event-related potential originating in the ACC, responds to both positive and negative stimuli (Schupp et al. 2000; Junghofer et al. 2001; Keil et al. 2002; Moratti et al. 2004; Moratti et al. 2011); for a review see Olofsson et al. (2008). Similarly, human amygdala lesions are related to difficulties in perceiving both fearful (Adolphs et al. 1995) and happy faces (Kipps et al. 2007). Electrical stimulation of the human medial temporal lobe (Halgren et al. 1978) produces experiences of both pleasure and displeasure across instances. Indeed, a more recent review of human intracranial electrophysiological recordings reveals that when stimulated, limbic, paralimbic, and cortical regions produce both positive and negative affective responses across instances (Guillory and Bujarski 2014).

The findings in humans are further consistent with research in non-human animals, which shows that lesions to the amygdala disrupt behavioral reactions to both positive and negative stimuli (Bliss-Moreau et al. 2011). Even neurochemicals such as opiods and dopamine appear to be general to both pleasure and pain in non-human animals (Leknes and Tracey 2008). Together, these findings underscore the idea that valence-general responsivity might be a feature of large-scale brain activity. Furthermore, although single-cell recording studies demonstrate that some cells do respond more during positivity than negativity (and vice versa), questions remain as to whether this is the best level of analysis at which to map psychological function. If cells that encode positivity, negativity, and both, are distributed

throughout limbic cortex (e.g., the amygdala and orbitofrontal cortex; see Salzman and Fusi 2010) and have excitatory and inhibitory relationships to one another, then this renders inert the question of whether there truly exists a "positive system" and "negative system" in the brain.

Consistent with the interpretation that cells for positive and negative affect are spread throughout brain regions, Chikazoe et al. (2014) recently observed populations of neurons within mOFC, lOFC, anterior insula, ACC and ventral temporal cortex that showed population-level coding for positive versus negative affect within the brains of humans. To do so, they used a representational similarity analysis (RSA) of neuroimaging data to assess whether there are correlations between the spatial patterns of brain activity across a given event (e.g., viewing affective pictures) and patterns in the dimensions that characterize those events (e.g., valence ratings of pictures). RSA reveals the structure of representations in terms of distances between response vectors within a given representational space (Haxby et al. 2014). Rather than finding different regions or networks that separately support positivity versus negativity, Chikazoe et al.'s findings suggest that there are representations for positivity versus negativity both within the same brain region and spread across the brain. In essence, their findings, like ours, suggest that there are not brain systems for positivity or negativity per se, but populations of neurons within regions across the brain regions that support positive and negative valence.

The fact that Chikazoe et al. (2014) were able to classify patterns for positivity versus negativity within regions across the brain appears to be in conflict with our own pattern classification analysis. However, a closer look reveals that Chikazoe et al.'s findings are not really so different from our meta-analytic results. First, Chikazoe et al.'s univariate analyses using the general linear model replicate our own findings that brain areas such as the mOFC, ACC, insula, and ventral temporal cortex respond to both positive and negative valence across instances. Their findings, along with our meta-analytic findings, indicate that the majority of the BOLD signal within limbic and paralimbic regions of the brain is not in fact valence specific. Chikazoe et al. do show, however, that a small amount of variance at the level of the individual is accounted for by fine-grained spatial patterns within a given region of the brain (e.g., mOFC, mOFC, lOFC, anterior insula, ACC, and ventral temporal cortex). However, Chikazoe et al. are only able to predict patterns between individuals at 5.6% better than chance. What we do not know is how stable this small amount of variance is across studies of different methods. It is also possible that this relatively small proportion of the variance in brain activity actually reflects some other stimulus property that tracks with valence in the stimuli or experimental task that was not controlled for in Chikazoe et al.'s analysis (e.g., arousal level is confounded with valence in the International Pictures System; Bradley et al. 2001 that was used by Chikazoe et al.). Indeed, standard multivariate analyses are known to systematically admit certain confounds related to the individual or task (e.g., reaction time differences on different trial types) into analyses and as a result can find spurious patterns of brain activity that are not found with standard univariate analyses or multivariate analyses that properly control for sources of individual-level variation that are unrelated to the phenomenon of interest (Todd et al. 2013). Arousal was less likely to be a confound in our meta-analysis because we summed across many studies, and whereas some use stimuli that confound valence and arousal (e.g., the International Pictures System), others use other stimuli that may be less likely to confound valence and arousal (e.g., pictures of emotional facial expressions). Of course, these questions remain a topic of future research, which should assess whether the valence-specific similarity-based patterns revealed by Chikazoe et al. are replicable across people, studies, and paradigms. The benefit of the present meta-analysis is that it summarizes whether there is valence specificity across people, studies, and paradigms at the level of brain regions. We found that there are not classifiable patterns of brain activity in any region of the brain that corresponds to positivity versus negativity across people, studies, and paradigms. This finding is important because researchers still widely assume that valence specificity is present at the regional brain level (e.g., areas in the salience network form an "aversion network"; Hayes and Northoff 2011). Our meta-analysis thus suggests that assertions about valence specificity are not warranted; it is more appropriate to talk about regions as valence-general. If some regions show relatively more frequent activity for one type of valence over another, as we found, then this meta-analytic finding reflects that region's relative preference for one type of valence versus another but not valence specificity per se.

Brain Regions within the Affective Workspace with Relative Preference for Negativity

We found several regions that showed a relative preference for negativity over positivity. Within the brain's affective workspace, the left amygdala and both ventral and dorsal portions of the left anterior insula demonstrated relatively more consistent increases during negativity than during positivity—even though responding occurred to both positivity and negativity at levels consistently greater than chance in these regions across studies. No brain activity in any of these regions was classifiable using a pattern classification analysis, however, meaning that there were no identifiable patterns within the affective workspace that can definitively classify whether a study contrast specifically invoked positivity or negativity.

The relatively greater frequency of response for negativity that we observed within the left amygdala and insula might instead be related to the negativity bias (Baumeister et al. 2001) that is observed in the behavioral literature: negative valence has greater psychological impact than positive valence when it comes to reactions to life events, outcomes in relationships, and even fundamental psychological processes such as learning and attention. Since negative stimuli recruit relatively more attention than positive stimuli (Baumeister et al. 2001), our findings converge with literature demonstrating increases in dorsal aspects of the affective workspace (e.g., dorsal anterior insula and dorsal ACC). This network is also called the "ventral attention network" and is frequently identified during tasks involving visual orienting (for reviews, see Corbetta et al. 2008; Lagner and Eickhoff 2013). Fluctuations in connectivity strength within the ventral aspects of the affective workspace (ventral anterior insula and pregenual ACC) correlate with subjective reports of arousal while participants look at unpleasant images (Touroutoglou et al. 2012). It thus remains a possibility that the negative stimuli used in experiments tend, on average, to be more subjectively arousing than the positive stimuli, producing the observed negativity bias.

The Role of Arousal

Indeed, one interpretation of the valence-general brain activity we observed is that activity in this set of regions reflects a second psychological dimension that positive and negative states share in common: arousal—the degree of activation versus

deactivation associated with a valenced state (Larsen and Diener 1992; Salzman and Fusi 2010; Kuppens et al. 2013). This is often a possible interpretation when studies fail to identify differences in mean brain activity to positive versus negative affect (cf., Chikazoe et al. 2014). However, Chikazoe et al.'s (2014) RSA indicated that the positive and negative valence of pictures correlated with brain patterns that were maximally distinct from one another, suggesting that evidence for valence-general mean brain activity in univariate analyses does not merely reflect arousal (although as we note earlier, it is still possible that the arousal properties of affect could have contributed to Chikazoe et al.'s multivariate findings).

Nonetheless, the question of whether valence-general activations represent arousal is complex, both theoretically and empirically. Statistically, changes in arousal and valence are impossible to separate because affective responses are arrayed as a circumplex, such that changes in valence always accompany some change in arousal (Barrett & Bliss-Moreau 2009). Experimentally, valence and arousal are also difficult to separate because most stimuli used to induce positivity or negativity also induce some change in arousal (e.g., the International Affective Picture System; Lang et al. 2005). Theoretically, the concept of arousal, itself, is vague and underspecified. As a psychological property, the term "arousal" is used to refer to enhanced attention, behavioral engagement, intensity of feeling, and/or physiological activation. In self-reports of affective experience, valence and arousal are separable properties for some, but not for all, individuals (Feldman 1995). Nonetheless, consistent with this arousal-based interpretation of our findings, increased activation within the affective workspace, such as the amygdala (Wilson-Mendenhall et al. 2013) and the anterior insula (Moriguchi et al. 2014), is associated with more intense subjective experiences of arousal. Stronger intrinsic connectivity within the affective workspace is also associated with more intense experiences of arousal (Touroutoglou et al. 2012). It thus certainly remains a possibility that activity in the brain regions we observed in our meta-analysis is in some way related to the arousing nature of affect.

A possibility that should be further investigated through more specifically controlled neuroimaging studies is that some aspects of the affective workspace encode general arousal and others encode valence more specifically. Our findings, in concordance with other recent neuroimaging studies (Wilson-Mendenhall et al. 2013; Chikazoe et al. 2014), suggest that the vmPFC and perhaps more specifically the mOFC is a candidate region for representing the subjective valence of otherwise ambiguous arousal.

Outstanding Questions and Future Directions

Questions about the neural representation of positivity and negativity remain that cannot be addressed by the present meta-analysis. First, as we have discussed throughout, it is possible that there are distributed functional networks that specifically support positive valence and other networks that specifically support negative valence and that these networks have an inhibitory relationship with one another. This hypothesis should be addressed in individual neuroimaging studies using functional connectivity analyses. It is sometimes assumed that neural "deactivations" can be harnessed to assess inhibition between brain regions in neuroimaging studies, but there are several reasons why "deactivations" would not help address this hypothesis in the context of a meta-analysis. As is standard in many meta-analyses of neuroimaging studies (e.g., Vytal and Hamann 2010; Lindquist, Wager, Kober et al. 2012), our database did not include contrasts reporting neural "deactivations" because they are reported infrequently across individual studies. Furthermore, "deactivations" are problematic because their interpretation depends on what other condition they are compared with (i.e., they might not represent "deactivation" per se; for a discussion, see Lindquist, Wager, Bliss-Moreau et al. 2012). More generally, since it is not clear whether BOLD activity generally reflects activations, inhibitions, or both, this hypothesis seems unanswerable by neuroimaging. Second, it remains a possibility that specificity will be observed at more fine-grained levels of analysis. As Chikazoe et al. (2014) demonstrate, positivity and negativity are represented by distributed populations of neurons spread throughout a given brain region. It might thus be more fruitful to regard function (i.e., the representation of positivity and negativity) as emerging out of populations of cells that interact with one another to produce a given mental representation in a given context. Of course, what we do not know from Chikazoe et al. (2014) is whether the same cells within the mOFC, lOFC, insula and ACC consistently represent positivity and negativity on every instance or whether they show degeneracy, with a single cell representing positive affect on 1 occasion and negative affect on another. Degeneracy occurs when different structures (i.e., cells and brain regions) perform the same function across instances (Edelman and Gally, 2001). The degenerate representation of valence would also ultimately be consistent with the valence-general affective workspace but appears unanswerable by neuroimaging. Non-human animal research that addresses this question seems to suggest that valence specificity is not the rule at the level of single cells, however.

These limitations notwithstanding, our findings make a key contribution toward understanding how valence is represented in regional brain activity. Twenty years after the advent of neuroimaging technology, the body of neuroimaging data suggests that the interpretation of behavioral data might be leading researchers to ask the wrong questions when it comes to the brain basis of valence. Although people can experience positivity and negativity as distinct states that contribute independently to attitudes and decisions (consistent with a bivalent view; Watson and Tellegen 1985; Cacioppo et al. 1999; Norris et al. 2010), they can also experience them as negatively correlated across instances (consistent with a bipolar view; Larsen and Diener 1992; Carroll et al. 1999). Yet, neither of these observations means that the brain is divisible into separate positive and negative systems or consists of a single system that encodes opposing positive and negative states. To assume so would be to conflate the content of experience (feelings of ambivalence or feelings of opposition between positivity and negativity) with the processes that produce those experiences.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/

Funding

This work was supported by National Institute of Health Director's Pioneer Award (DP1OD003312) and by the U.S. Army Research Institute for the Behavioral and Social Sciences (contract W5J9CQ-11-C-0046) awarded to L.F.B. *Conflict of Interest*

Notes

The views, opinion, and/or findings contained in this article are solely of the authors and should not be construed as an official

Department of the Army of DOD position, policy, or decision. *Conflict of Interest*: None declared.

References

Adolphs R, Tranel D, Damasio H, Damasio AR. 1995. Fear and the human amygdala. J Neurosci. 15:5879–5891.

Barrett LF, Bliss-Moreau E. 2009. Affect as a psychological primitive. Adv Exp Soc Psychol. 41:167–218.

Barrett LF, Russell JA. 1998. Independence and bipolarity in the structure of current affect. J Pers Soc Psychol. 74:967–984.

Barrett LF, Russell JA. 1999. Structure of current affect. Curr Direct Psychol Sci. 8:10–14.

Bartra O, McGuire JT, Kable JW. 2013. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage. 76:412–427.

Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD. 2001. Bad is stronger than good. Rev Gen Psychol. 54:323–370.

Belova MA, Paton JJ, Salzman CD. 2008. Moment-to-moment tracking of state value in the amygdala. J Neurosci. 28:10023–10030.

Berntson GG, Cacioppo JT, Quigley KS. 1991. Autonomic determinism: the modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. Psychol Rev. 98:459–487.

Biswal BB, Mennes M, Zuo XN, Gohel S, Kelly C, Smith SM, Beckmann CF, Adelstein JS, Buckner RL, Colcombe S, et al. 2010. Toward discovery science of human brain function. Proc Natl Acad Sci USA. 107:4734–4739.

Bliss-Moreau E, Bauman MD, Amaral DG. 2011. Neonatal amygdala lesions result in globally blunted affect in adult rhesus macaques. Behav Neurosci. 125:848–858.

Bradley MM, Codispoti M, Cuthbert BN, Lang PJ. 2001. Emotion and motivation I: defensive and appetitive reactions in picture processing. Emotion. 1:276–298.

Buckner RL, Carroll DC. 2007. Self-projection and the brain. Trends Cog Sci. 11:49–57.

Cacioppo JT, Gardner WL, Berntson GG. 1999. The affect system has parallel and integrative processing components: form follows function. J Pers Soc Psychol. 76:839–855.

Carroll JM, Yik M, Russell JA, Barrett LF. 1999. On the psychometric principles of affect. RevGen Psychol. 3:14–22.

Chang C, Lin C. 2011. LIBSVM: a library for support vector machines. ACM Transact Intell Syst Technol. 2:1–27.

Chikazoe J, Lee DH, Kriegeskorte N, Anderson AK. 2014. Population coding of affect across stimuli, modalities and individuals. Nat Neurosci. 17:1114–1122.

Clithero JA, Rangel A. 2013. Informatic parcellation of the network involved in the computation of subjective value. Soc Cogn Affect Neurosci. 9:1289–1302.

Corbetta M, Patel G, Shulman GL. 2008. The reorienting system of the human brain: from environment to theory of mind. Neuron. 58:306–324.

Craig AD. 2009. How do you feel–now? The anterior insula and human awareness. Nat Rev Neurosci. 10:59–70.

Edelman GM. 1989. The Remembered Present: a Biological Theory of Consciousness. New York: Basic Books.

Edelman GM, Gally JA. 2001. Degeneracy and complexity in biological systems. Proc Natl Acad Sci USA. 98:13763–13768.

Farroni T, Menon E, Rigato S, Johnson MH. 2007. The perception of facial expressions in newborns. Eur J Develop Psychol. 4:2–13.

Feldman LA. 1995. Variation in the circumplex structure of affect. Personal Soc Psychol Bull. 21:806–817.

Guillory SA, Bujarski KA. 2014. Exploring emotions using invasive methods: review of 60 years of human intracranial electrophysiology. Soc Cogn Affect Neurosci. 9:1880–1889.

Halgren E, Walter RD, Cherlow DG, Crandall PH. 1978. Mental phenomena evoked by electrical stimulation of the human hippocampal formation and amygdala. Brain. 101:83–115.

Haxby JV, Connolly AC, Guntupalli JS. 2014. Decoding neural representational spaces using multivariate pattern analysis. Annu Rev Neurosci. 37:435–456.

Hayes DJ, Northoff G. 2011. Identifying a network of brain regions involved in aversion-related processing: across species translational investigation. Front Integr Neurosci. 5:1–21.

Junghofer M, Bradley MM, Elbert TR, Lang PJ. 2001. Fleeting images: a new look at early emotion discrimination. Psychophysiology. 38:175–178.

Keil A, Bradley MM, Hauk O, Rockstroh B, Elbert T, Lang PJ. 2002. Large-scale neural correlates of affective picture processing. Psychophysiology. 39:641–649.

Kipps CM, Duggins AJ, McCusker EA, Calder AJ. 2007. Disgust and happiness recognition correlate with anteroventral insula and amygdala volume respectively in preclinical huntington's disease. J Cogn Neurosci. 19:1206–1217.

Kober H, Barrett LF, Joseph J, Bliss-Moreau E, Lindquist KA, Wager TD. 2008. Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. Neuroimage. 42:998–1031.

Kober H, Wager TD. 2010. Meta-analyses of neuroimaging data. Wiley Interdiscipl Rev. 1:293–300.

Kringelbach ML, Rolls ET. 2004. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. Prog Neurobiol. 72:341–372.

Kuhn S, Gallinat J. 2012. The neural correlates of subjective pleasantness. Neuroimage. 61:289–294.

Kuppens P, Tuerlinckx F, Russell JA, Barrett LF. 2013. The relationship between valence and arousal in subjective experience. Psychol Bull. 139:917–940.

Lagner R, Eickhoff SB. 2013. Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. Psychol Bull. 139:870–900.

Lang PJ, Bradley MM. 2010. Emotion and the motivational brain. Biol Psychol. 84:437–450.

Lang PJ, Bradley MM, Cuthbert BN, et al. 2005. International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. Gainesville, FL: NIMH, Center for the Study of Emotion & Attention.

Larsen JT, McGraw AP, Cacioppo J. 2001. Can people feel happy and sad at the same time? J Pers Soc Psychol. 81:684–696.

Larsen JT, McGraw AP, Mellers BA, Cacioppo JT. 2004. The agony of victory and the thrill of defeat: Mixed emotional reactions to dissapointing wins and releiving losses. Psychol Sci. 15:325–330.

Larsen RJ, Diener E. 1992. Promises and Problems with the Circumplex Model of Emotion. Newbury Park, CA: Sage. p. 25–59.

Leknes S, Tracey I. 2008. A common neurobiology for pain and pleasure. Nat Rev Neurosci. 9:314–320.

Lewis M. 1993. The emergence of human emotions. In: Lewis M, Haviland-Jones JM, editors. Handbook of Emotions. 2 ed. New York: Guilford. p. 253–322.

Lindquist KA, Barrett LF. 2012. A functional architecture of the human brain: Insights from emotion. Trends Cog Sci. 16: 533–554.

Lindquist KA, Wager TD, Bliss-Moreau E, Kober H, Barrett LF. 2012. What are emotions and where are they in the brain? Behav. Brain Sci. 35:175–184.

Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, Barrett LF. 2012. The brain basis of emotion: a meta-analytic review. Behav Brain Sci. 35:121–143.

Liu X, Hairston J, Schrier M, Fan J. 2011. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. Neurosci Biobehav Rev. 35:1219–1236.

Moratti S, Keil A, Stolarova M. 2004. Motivated attention in emotional picture processing is reflected by activity modulation in cortical attention networks. Neuroimage. 21:954–964.

Moratti S, Saugar C, Strange BA. 2011. Prefrontal-occipitoparietal coupling underlies late latency human neuronal responses to emotion. J Neurosci. 31:17278–17286.

Moriguchi Y, Touroutoglou A, Dickerson BC, Barrett LF. 2014. Sex differences in the neural correlates of affective experience. Soc Cogn Affect Neurosci. 9:591–600.

Morrison SE, Salzman CD. 2009. The convergence of information about rewarding and aversive stimuli in single neurons. J Neurosci. 29:11471–11483.

Murphy FC, Nimmo-Smith I, Lawrence AD. 2003. Functional neuroanatomy of emotions: a meta-analysis. Cogn Affect Behav Neurosci. 3:207–233.

Nichols T, Brett M, Andersson J, Wager T, Poline JB. 2005. Valid conjunction inference with the minimum statistic. Neuroimage. 25:653–660.

Norris CJ, Golan J, Berntson GG, Cacioppo JT. 2010. The current status of research on the structure of evaluative space. Biol Psychol. 84:422–436.

Olofsson JK, Nordin S, Sequeira H, Polich J. 2008. Affective picture processing: an integrative review of ERP findings. Biol Psychol. 77:247–265.

Osgood CE. 1952. The nature and measurement of meaning. Psychol Bull. 49:197–237.

Paton JJ, Belova MA, Morrison SE, Salzman CD. 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature. 439:865–870.

Poldrack RA. 2010. Mapping mental function to brain structure: How can cognitive neuroimaging succeed? Perspect Psychol Sci. 5:753–761.

Reynolds SM, Berridge KC. 2008. Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. Nat Neurosci. 11:423–425.

Robinson MJ, Berridge KC. 2013. Instant transformation of learned repulsion into motivational "wanting." Curr Biol. 23:282–289.

Roy M, Shohamy D, Wager TD. 2012. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. 16:147–156.

Russell JA, Barrett LF. 1999. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. J Pers Soc Psychol. 76:805–819.

Salimi-Khorshidi G, Smith SM, Keltner JR, Wager TD, Nichols TE. 2009. Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. Neuroimage. 45:810–823.

Salzman CD, Fusi S. 2010. Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. Annu Rev Neurosci. 33:173–202.

Schupp HT, Cuthbert BN, Bradley MM, Cacioppo JT, Ito T, Lang PJ. 2000. Affective picture processing: the late positive potential is modulated by motivational relevance. Psychophysiology. 37:257–261.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci. 27:2349–2356.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, et al. 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci. 106:13040–13045.

Stark CE, Squire LR. 2001. When zero is not zero: the problem of ambiguous baseline conditions in fMRI. Proc Natl Acad Sci. 98:12760–6.

Todd MT, Nystrom LE, Cohen JD. 2013. Confounds in multivariate pattern analysis: theory and rule representation case study. Neuroimage. 77:157–165.

Touroutoglou A, Hollenbeck M, Dickerson BC, Barrett LF. 2012. Dissociable large-scale networks anchored in the anterior insula subserve affective experience and attention/executive function. Neuroimage. 60:1947–1958.

Tsunozaki M, Chalasani SH, Bargmann CI. 2008. A behavioral switch: cGMP and PKC signaling in olfactory neurons reverses odor preference in *C. elegans*. Neuron. 59:959–971.

Vytal K, Hamann S. 2010. Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. J Cogn Neurosci. 22:2864–2885.

Wager TD, Barrett LF, Bliss-Moreau E, Lindquist KA, Duncan S, Kober H, Joseph J, Davidson M, Mize J. 2008. The neuroimaging of emotion. In: Lewis M, Haviland-Jones JM, Barrett LF, editors. Handbook of Emotions. 2nd ed. New York: Guilford. p. 513–530.

Wager TD, Lindquist M, Kaplan L. 2007. Meta-analysis of functional neuroimaging data: current and future directions. Soc Cogn Affect Neurosci. 2:150–158.

Wager TD, Phan KL, Liberzon I, Taylor SF. 2003. Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. Neuroimage. 19:513–531.

Watson D, Tellegen A. 1985. Toward a consensual structure of mood. Psychol Bull. 98:219–235.

Wierzbicka A. 1992. Semantics, culture, and cognition: universal human concepts in culture-specific configurations. USA: Oxford University Press.

Wilson-Mendenhall CD, Barrett LF, Barsalou LW. 2013. Neural evidence that human emotions share core affective properties. Psychol Sci. 24:947–956.

Wilson-Mendenhall CD, Feldman Barrett L, Barsalou LW. 2014. Variety in emotional life: within-category typicality of emotional experiences is associated with neural activity in large-scale brain networks. Soc Cogn Affect Neurosci. 10:62–71.

Wundt W. 1897/1998. Outlines of psychology. Bristol, UK: Thoemmes Press.

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zollei L, Polimeni JR, et al. 2011. The organization of the human cerebral cortex estimated by functional connectivity. J Neurophysiol. 106:1125–1165.