Published in final edited form as:
Perspect Psychol Sci. 2017 May ; 12(3): 508–526. doi:10.1177/1745691616685863.

A mathematical model captures the structure of subjective affect

Alison M. Mattek[1], **George Wolford**[1], and **Paul J. Whalen**[1]

[1]Dartmouth College

Abstract

While it is possible to observe when another person is having an emotional moment, we also derive information about the affective states of others from what they tell us they are feeling. In an effort to distill the complexity of affective experience, psychologists routinely focus on a simplified subset of subjective rating scales (i.e., dimensions) that capture considerable variability in reported affect: reported valence (i.e., how good or bad?) and reported arousal (e.g., how strong is the emotion you are feeling?). Still, existing theoretical approaches address the basic organization and measurement of these affective dimensions differently. Some approaches organize affect around the dimensions of bipolar valence and arousal (e.g., the circumplex model; Russell, 1980), whereas alternative approaches organize affect around the dimensions of unipolar positivity and unipolar negativity (e.g., the bivariate evaluative model; Cacioppo & Berntson, 1994). In this report, we (1) replicate the data structure observed when collected according to the two approaches described above, and re-interpret these data to suggest that the relationship between each pair of affective dimensions is conditional on valence ambiguity; then (2) formalize this structure with a mathematical model depicting a valence ambiguity dimension that decreases in range as arousal decreases (a triangle). This model captures variability in affective ratings better than alternative approaches, increasing variance explained from ~60% to over 90% without adding parameters.

Keywords

arousal; emotion; model of affect; subjective report; valence

The qualitative change in subjective experience that corresponds with an affective event (such as a painful injury or a pleasurable meal) has consistently been a topic of interest for psychologists (e.g. Wundt, 1912/1920/2014; James, 1884; Woodsworth & Schlosberg, 1938/1965). Subjective experience (by definition) can only be observed by the subject, so psychological research usefully relies on rating scales as a means of quantifying this aspect of affective responding. One dominant approach has been to focus on two rating scales (i.e., dimensions) that capture a large portion of variability in affective experience as it is subjectively described: (1) valence[1], measured with ratings of how positive or negative a particular event is and (2) arousal[2], measured with ratings of how emotionally provocative a

Correspondence to: Alison M. Mattek, Department of Psychological and Brain Sciences, Dartmouth College, Alison.M.Mattek@Dartmouth.edu.

Declaration of Conflicting Interests
The authors declare no conflicts of interest with respect to the authorship or publication of this article.

particular event is (e.g., Schlosberg, 1954; Russell, 1980; Cacioppo, Petty, Losch, & Kim, 1986; Lang, Bradley, & Cuthbert, 1998).

The theoretical conceptualization of these two dimensions has driven a substantial body of experimental work. For example, ratings of valence and/or arousal have been used to predict a number of other behaviors or physiological responses, including brain activity (e.g., Knutson, Katovich, & Suri, 2014; Chikazoe, Lee, Kriegeskorte, & Anderson, 2014; Lindquist, Satpute, Wager, Weber, & Barrett, 2015; Man, Nohlen, Melo, & Cunningham, 2016), skin conductance (e.g., Lykken & Venables, 1971; Greenwald, Cook, & Lang, 1989), and memory performance (e.g., Bradley, Greenwald, Petry, & Lang, 1992; Cahill, Haier, Fallon, Alkire, Tang, Keator, Wu, & McGaugh, 1996; Kensinger, 2004) to name but a few. Over time, this work has led to the establishment of two contrasting frameworks for describing these dimensions of affect—each adopting differing assumptions as to how the dimensions of valence and arousal are structured and should be measured. In practice, experimental designs within the domain of affective psychology often adopt the assumptions of one approach or the other, whether or not this is explicitly stated.

Two approaches to measuring subjective ratings of 2D affect

One approach (i.e., the bipolar framework; see Barrett & Russell, 1998; Russell, 2003; 2016 for reviews) is organized around the dimensions of bipolar valence and arousal, and is largely informed by the circumplex model of affect (Russell, 1980) and its precursors (e.g., Schlosberg, 1954; Russell & Mehrabian, 1977). The many experiments that adopt this basic framework implicitly make the following assumptions: (1) as positivity increases, negativity decreases (and vice versa); in other words, this framework employs a single bipolar scale ranging from positive to negative to measure subjective valence; and (2) changes in reported valence are subjectively different from changes in reported arousal, so this framework also employs an additional subjective arousal scale ranging from low (e.g., sleepy/inactive) to high (e.g., excited/active).

An alternative approach (i.e., the bivariate framework; see Cacioppo, Berntson, Norris, & Gollan, 2012 and Larsen, 2016 for reviews) has been developed in parallel, which is organized around the dimensions of unipolar positivity and unipolar negativity, and is largely informed by the bivariate evaluative model (Cacioppo & Berntson, 1994) and its precursors (e.g., Bradburn, 1969; Watson & Tellegen, 1985). Experiments that adopt this framework generally make a different set of assumptions: (1) as either positivity *or* negativity increases, the level of arousal also increases. Consequently, this framework generally does not treat

[1]The term “valence” was used by chemists before psychologists, and was not introduced to psychology until Kurt Lewin's works were translated from German to English: “A fairly precise translation of *Aufforderungschsrakter* is the term ‘demand value’ which [Edward] Tolman uses for the same concept. In order to avoid unnecessary misunderstandings, Professors Tolman and Lewin have agreed to use the same term and at Tolman's suggestion have chosen ‘valence’...the term is used in speaking of a stimulus-reaction process... in contrast to chemical valence, which is only positive, psychological valence...may be either positive (attracting) or negative (repelling)...an object or activity loses or acquires valence (of either kind) in accordance with the needs of the organism.” (from footnote and Translators’ note on p. 77 in Lewin, 1935).

[2]In this list of references, the subjective dimension we refer to as “arousal” is alternatively labeled “activation” (Schlosberg, 1954; Russell, 1980), “affective intensity” (Cacioppo et al., 1986), and “emotional arousal” (Lang et al., 1998). Although we treat these constructs as overlapping here, some work uses these terms to represent separable constructs (e.g., Reisenzein, 1994); see Table S2 for a list of definitions of the general “arousal” construct spanning the past several decades.

arousal as a dimension that is separate from valence intensity and does not measure subjective arousal directly; and (2) positivity can increase or decrease without any corresponding change in negativity, and negativity can increase or decrease without any corresponding change in positivity. Consequently, this approach measures valence (and by proxy arousal, given the first assumption) with two scales, unipolar positivity and unipolar negativity.

Both frameworks usefully distill the complexity of affective experience with two rating scales (i.e., they are both 2-dimensional frameworks). However, the underlying assumptions of these competing methodologies bring to light some conspicuous theoretical questions. First, is arousal equivalent to valence intensity, or does it represent a separable affective dimension? Second, is valence best represented by two unipolar dimensions—positivity and negativity—or rather by a single bipolar dimension that ranges from positive to negative? We will see that the answers to these questions dictate the geometric structure we use to represent these facets of reported affect.

The remainder of this report is organized into four sections. First, a Background section will provide the reader with a summary of existing findings that have been taken to support either of the two, 2-dimensional approaches outlined above. Second, a Data section will present both a replication of the findings described in the Background section as well as a re-interpretation of the observed data structure. Third, a Model section will propose a mathematical formalization of the data structure reviewed in the Background section and replicated in the Data section. We demonstrate that the model proposed here fits the observed rating data substantially better than existing alternatives. Finally, in the General Discussion section, we consider how this model can contribute to theoretical issues and methodological choices in the domain of affective psychology.

Background

In the previous section, we outlined the underlying methodological assumptions that have followed from the dominant theoretical descriptions of dimensional affect (bipolar versus bivariate). Logically, both sets of assumptions cannot be simultaneously correct. Indeed, we will see that the experimental effects taken to support one approach can be used as evidence for refuting the other. In both cases, we will be discussing effects that are large and have been consistently replicated. Reconciling the effects described in the next two subsections is a major aim of this report.

Findings supporting the bipolar approach

Evidence for the bipolar approach has been discussed at length (e.g., see Russell, 2016 for a review). In this section, we focus on two reliable effects found in subjective report data that clearly support this approach. First, dimensionality reduction techniques reveal bipolar valence and arousal to be two primary dimensions in various measures of reported affect[3], such as: adjective checklist responses following mood manipulations (Nowlis & Nowlis,

[3]Some variant of these dimensions have been shown to describe a large portion of variance in a host of different types of reported affect, including: Across the examples cited here, these two dimensions account for up to three-fourths of the variance in subjective

1956), similarity ratings of emotion adjectives (Bush, 1973), semantic groupings of emotion words (Russell, 1980), similarity ratings of emotional facial expressions (Abelson & Sermat, 1962; Shepard, 1962), patterns of categorization error for emotional facial expressions (Osgood, 1966), similarity ratings of non-verbal emotional vocalizations (Green & Cliff, 1975), a clustering analysis of low-level features that represent emotion in music and visual motion (Sievers, Polansky, Casey, & Wheatley, 2013), and many other comparable datasets (e.g., see Russell & Mehrabian, 1977; Russell, 1980; Feldman, 1995; Barrett & Russell, 1999). Arousal and bipolar valence appearing as separate dimensions supports the use of separate valence and arousal scales in the bipolar approach, and is a challenge to equating arousal with valence intensity in the bivariate approach (e.g., Kron, Goldstein, Lee, Gaurdhouse, & Anderson, 2013).

The second reliable effect in support of the bipolar approach (which comes from data collected according to the bivariate approach) is that ratings of unipolar positivity and ratings of unipolar negativity have a reliable inverse correlation. That is, the more something is ‘positive’, the more likely it is ‘not negative’, and vice versa (e.g., Green, Goldman, & Salovey, 1993; Russell & Carroll, 1999). This inverse relationship between positivity and negativity ratings suggests that it is generally appropriate to represent valence with a single bipolar dimension rather than with two separate unipolar dimensions (Barrett & Russell, 1998; Yik, Russell, & Barrett, 1999; Russell 2016), which is a notable challenge to the bivariate approach.

Findings supporting the bivariate approach

Evidence for the bivariate approach has also been discussed at length (see Cacioppo, Berntson, Norris, & Gollan, 2011 and Larsen, 2016 for reviews). In this section, we focus on two reliable effects found in subjective report data that clearly support this alternative approach. First, ratings within the Cartesian product created by the dimensions of bipolar valence and arousal (i.e., rating data collected according to the bipolar approach) tend to take on an inverted triangular form. Figure 1A-F is a re-presentation of figures from six previous reports showing that an inverted triangular space captures the majority of rating data derived from some variant of separate bipolar valence and arousal scales. The data depicted in Figure 1 demonstrate that in practice, bipolar valence ratings and arousal ratings are correlated on either side of the valence midpoint. This effect has been observed for ratings of visual scenes (Lang, Greenwald, Bradley, & Hamm, 1993; Ito, Cacioppo, & Lang, 1998; Anders, Lotze, Erb, Grodd, & Birbaumer, 2004; Britton, Taylor, Sudheimer, & Liberzon, 2006; Lang, Bradley, & Cuthbert, 2008; Kron, et al. 2013; Kurdi, Lozano, & Banaji, 2016), facial expressions (Schlosberg, 1952), sound clips (Bradley & Lang, 2007), olfactory cues (Bensafi, 2001; Bensafi, Rouby, Farget, Bertrand, Vigouroux, & Holley, 2002; Jin, Zelano, Gottfried, & Mohanty, 2015), gustatory cues (Small, Gregory, Mak, Gitelman, Mesulam, & Parrish, 2003), and affective words (Bradley & Lang, 1999; Lewis, Critchley,

rating measurements, with the bipolar valence dimension accounting for up to half of the variance. However, we note that the number of explanatory dimensions as well as their order of variance-accounted-for shifts somewhat across studies and individuals, and only a small number of examples are listed here. The question of whether valence and arousal are the most “important” dimensions is beyond the scope of this report and has been discussed elsewhere (e.g., Schlosberg, 1954; Russell & Mehrabian, 1977; Russell, 1980; Barrett & Russell, 1999).

Rothstein, & Dolan, 2007). The fact that valence intensity and arousal ratings are correlated on either side of the valence midpoint fits with the bivariate approach of not collecting a separate arousal rating, and is a challenge to the bipolar approach.

The second reliable effect in support of the bivariate approach is: although ratings of positivity and ratings of negativity generally have a strong inverse correlation (as discussed in the previous section), there are numerous studies showing that the relationship between positivity and negativity is more complicated. For example, these ratings can be uncorrelated (e.g., Diener & Emmons, 1984; Cacioppo, Gardner, & Berntson, 1997; Watson, Wiese, D., Vaidya, J., & Tellegen, 1999) or can show a curvilinear relationship (e.g., Diener & Iran-Nejad, 1986; Ito et al., 1998; Schimmack, 2001). That positivity and negativity are not necessarily inversely correlated fits with the bivariate approach of using two separate unipolar valence scales, and is a challenge to the use of a single bipolar valence scale in the bipolar approach.

Taken together, we are left with a bit of a conundrum. How is it that both approaches can provide such strong supporting evidence when the assumptions that each approach is based upon appear to be logically opposed? To get at this issue, the current approach will assume that the relationship between arousal and valence ratings (and also the relationship between positivity and negativity ratings) must not be constant – that is, the relationships must be changing or conditional upon the types of stimulus events under investigation.

The current approach

In the Data section to follow, we present subjective rating data that replicate a number of the findings discussed up to this point. Our aim is to (a) offer a novel interpretation of the rating data focused on predicting the degree to which valence and arousal ratings (or positivity and negativity ratings) are correlated and (b) formalize these observations with a model that can capture these relationships and, in turn, explain a greater amount of variance in the rating data.

To briefly anticipate our findings: we will show that the critical determiner of the relationship between these pairs of affective dimensions is the clarity of valence of the event being rated, something we refer to as *valence ambiguity* (this definition is described in more detail in the Data section and is formalized in the Model section)[4]. Specifically, we demonstrate that (1) the correlation between bipolar valence and arousal ratings approaches zero as valence ambiguity increases, and (2) the inverse correlation between unipolar positivity and unipolar negativity ratings approaches zero as valence ambiguity decreases. The interpretation and model offered here provides an explanation for (1) why bipolar valence and arousal ratings are often correlated yet these dimensions can be separable in some cases, and (2) why positivity and negativity ratings are often inversely correlated yet these dimensions can be separable in some cases.

[4]See Potential Caveats & Limitations section for a discussion on a related phenomenon in social psychology, termed “ambivalence”.

Data: subjective ratings in response to briefly presented affective events

We collected subjective rating data from participants (total N = 195 across seven datasets) responding to a wide variety of briefly presented affective events including: faces (Tottenham, Tanaka, Leon, McCarry, Nurse, Hare, et al. 2009), images (Lang et al., 2008; Bradley & Lang, 1994; Bouhuys, Bloem, & Groothuis, 1995), words (Russell, 1980; Barrett, 2004), sentences (Kim, Somerville, Johnstone, Polis, Alexander, Shin, & Whalen, 2004; Barrett & Russell, 1998), and auditory stimuli (see Supplemental Materials I for a detailed description of all experimental stimuli), with the aim of replicating the structure observed in previously reported comparable rating measurements (see Figure 1).

Ratings along the dimensions of bipolar valence and arousal were collected in response to the stimulus items listed in Table 1 (datasets 1 through 5). Similarly, ratings along the dimensions of unipolar positivity and unipolar negativity were collected for a subset of those items, which are listed in Table 2 (datasets 6 & 7). For details regarding all data collection procedures, see Supplemental Materials I.

Ratings of bipolar valence and arousal

Figure 2A-B presents a group-level analysis of bipolar valence and arousal ratings in response to facial expressions of emotion (Figure 2A) and other affective events across multiple modalities (Figure 2B). Note the consistency of these data when compared with previously reported datasets in Figure 1, in that the majority of data points fit within an inverted triangular structure, and bipolar valence ratings and arousal ratings are significantly correlated on either side of the valence midpoint.

Bipolar valence and arousal ratings are contained within a boundary—Gross visual inspection of the ratings in Figure 2 suggests that the relationship between arousal and valence ratings (on either side of the valence midpoint) reflect an inverted triangular boundary constraining these data, rather than a piecewise linear relationship (i.e., a V-shaped, or absolute value function). Here, we demonstrate that the parameters of this boundary (i.e., the slopes and intercepts) do not need to be estimated from the data, but rather can be determined by the range of the scales employed: both line segments of the boundary begin at the valence midpoint at the lowest point on the arousal scale, one extends to the upper right and the other to the upper left of the Cartesian product created by the range of the measurement scales (as drawn in Figure 2). Across all datasets in this report, 86% of the variance in arousal ratings fall within this boundary ($R^2 = 0.857$)[5].

The relationship between valence and arousal ratings is not uniform within the boundary—Previous work has interpreted the correlations between bipolar valence and arousal as evidence that valence intensity and arousal are equivalent (e.g., Lang 1995; Kron, Pilkiw, Banaei, Goldstein, & Anderson, 2015). However, visual inspection of the data suggests that ratings proximal to the boundary will show a strong correlation between

[5]To calculate the R^2 for this boundary, the squared errors with respect to the V-shaped boundary were *partitioned* into negative error (below the boundary) and positive error (above the boundary), and only negative errors were included in the error term (i.e., only data falling outside the boundary were considered error). See Supplementary Materials I to see this analysis performed on each dataset separately.

arousal and valence ratings, whereas data points that fall within the central aspects of the triangular space will not. Consequently, if the boundary is assumed to represent a functional relationship, it has no predictive value[6], despite visibly capturing the gross structure of the ratings. In other words, the substantial number of points that fall in the central aspects of the space (within the boundary) incur a very large error if one assumes a constant relationship between valence intensity and arousal.

Proximity to the boundary reflects valence ambiguity—Inspection of Figure 2A reveals an interesting pattern regarding where different facial expression categories are located within the inverted triangular space. Expressions that are more clearly valenced show a near-perfect correlation between arousal and valence ratings, and thus are located proximal to the boundary of the space. Compared to these, surprised facial expressions (which do not have a clear canonical valence in that both positive and negative valence interpretations are plausible) populate the central aspects of the space. Notably, for this ambiguously valenced facial expression, the correlation between valence and arousal ratings breaks down. Thus, these data suggest that proximity to the triangular boundary inversely reflects valence ambiguity (i.e., items with a clearer valence are located closer to the boundary), and in turn that valence ambiguity determines whether valence and arousal ratings will be highly correlated (i.e., items along the boundary are highly correlated while items in the central aspects of the space are not). (i.e., items along the boundary are highly correlated while items in the central aspects of the space are not; for example, see Supplemental Figure S12). Although this interpretation might be but one way to frame these data, in a later section, we will more rigorously test this interpretation with a formalized model.

Ambiguous ratings are not due to individual differences—Are ratings that fall in the central aspects of the space truly ambiguous, or are they simply an artifact of averaging across individuals, some of whom provide very positive ratings and some of whom provide very negative ratings? One can see from gross inspection of the standard errors in Figure 2A (and also in Supplementary Figures S1-S9) that the latter is not the case. The standard error of ratings for ambiguously valenced items (i.e., points relatively distant from the boundary, in the central aspects of the space) are not noticeably different[7] than the standard error of ratings for more unambiguously valenced items (which are near the boundary). That is, participants consistently rate these items with a valence rating somewhere *between* positive and negative. This implies that the participants *know* that these items can be interpreted both ways, and therefore they are not *uncertain* about their valence judgments regarding ambiguous items (using a Bayesian definition of uncertainty, which is a measure of variance). This across-participant consistency also fits with reaction time data: when participants are forced to choose either a positive or negative label, responses are slower for ambiguous compared to unambiguous items (e.g., surprised faces compared to happy or angry faces—even for participants with a strong interpretive bias; Neta, Norris, & Whalen,

[6]Technically, using this fixed boundary as a function results in a negative $R^2 = -0.241$, because the data points (on average) are further away from the fixed V-shaped function plotted in Figure 2 (sum of squared errors [*SSE*] = 1197.98) than they are from the grand mean of arousal ratings (sum of squares total [*SST*] = 965.97).

[7]Even if variance was somewhat smaller for items that are closer to the boundary, this would be consistent with the fact that the variance of any measure always approaches zero with increasing proximity to any real boundary, giving rise to a Poisson-like distribution (for example, when RT's get very close to the zero boundary).

2009). In sum, ratings within the central aspects of the inverted triangular space cannot be reduced to an artifact of averaging across multiple individuals.

Summary—Putting this all together, if we define *valence ambiguity* quantitatively, for any given rating, as the distance from the boundary illustrated in Figure 2, a parsimonious description of the data would be as follows: the (absolute) correlation between bipolar valence and arousal ratings approaches 1 as valence ambiguity decreases and approaches 0 as valence ambiguity increases. This effect is evident with gross inspection of Figure 2 and is also visualized in Supplemental Figure S12.

Nonetheless, because bipolar valence ratings and arousal ratings are frequently correlated in practice, the bivariate framework has shown the usefulness of an alternative measurement approach, which employs unipolar positivity and unipolar negativity ratings (e.g., Ito et al., 1998; Larsen, McGraw, & Cacioppo, 2001; Kron et al., 2013; Larsen, 2016). In the next section, we will examine the structure of ratings within this alternative (unipolar positivity × unipolar negativity) space.

Ratings of unipolar positivity and unipolar negativity

The unipolar positivity and unipolar negativity ratings collected for this report (Figure 3) are consistent with previously reported datasets (cited in the Introduction), in that unipolar positivity ratings and unipolar negativity ratings have a strong inverse correlation (dataset 6: $R^2 = 0.691$; dataset 7: $R^2 = 0.559$). Previous work has taken this inverse correlation between positivity and negativity to support a bipolar model of valence.

Positivity and negativity have a mathematically reciprocal relationship—It is clear from visual inspection of the rating data in Figure 3 that a reciprocal curve fits the data substantially better than a line (dataset 6: $R^2 = 0.882$, dataset 7: $R^2 = 0.845$; compare to the linear effect size in the previous paragraph). This implies that a bipolar model of valence is inadequate in the sense that it is a linear approximation of a curvilinear relationship (see also Diener & Iran-Nejad, 1986). This curvilinear relationship is consistent with a number of behavioral and physiological findings showing that a positive motivational system and a negative motivational system have inhibitory effects on each other (e.g., Miller, 1960; Cacioppo & Berntson, 1994; Lang, 1995; Kim, Somerville, Johnstone, Alexander, & Whalen, 2003; Correia, McGrath, Lee, Graybiel, & Goosens, 2016; Urry, Van Reekum, Johnstone, Kalin, Thurow, Schaefer, ... & Davidson, 2006; Leknes & Tracey, 2008; Nikolova, Knodt, Radtke, & Hariri, 2016). The very terms “positive” and “negative” imply that one system would functionally negate the activity of the other.

Summary—In this alternative bivariate framework (unipolar positivity × unipolar negativity), we can quantitatively define *valence ambiguity* the same way we did within the bipolar framework: as the distance from the positivity and negativity axes of the space (imagine the Cartesian axes in Figure 3 are a 45-degree rigid clockwise rotation of the inverted triangular boundary in Figure 2). Considering this definition, the reciprocal fit in the rating data can be characterized as follows: the correlation between positivity and negativity approaches −1 as valence ambiguity increases and approaches 0 as valence ambiguity decreases. This effect is also visualized in Supplementary Figure S13.

Interpreting the data as fitting a reciprocal curve also offers an explanation for a number of findings that other models are unable to explain. First, a reciprocal fit predicts that positivity and negativity will be uncorrelated in some cases, whereas a strictly linear (bipolar) model does not. Second, this reciprocal fit indicates that separate positivity and negativity ratings do *not* account for all of the variance in arousal ratings, so equating the arousal dimension with valence intensity (as often occurs in the bivariate approach) does not account for this nuance[8]. More specifically, variance in arousal ratings of ambiguous items (that are spatially more distant from the boundary) become collapsed onto a reciprocal curve when plotted with respect to unipolar positivity and unipolar negativity. In other words, valence intensity is not a proxy for arousal when valence is ambiguous.

Interim Discussion

The rating data observed here, which replicate the general structure observed in previous data collected with both the bipolar and bivariate approaches (see Introduction), can be summarized as follows:

Regarding the bipolar approach: (1) bipolar valence ratings and arousal ratings co-vary when valence is unambiguous (e.g., happy or angry facial expressions, etc.—Figure 2, Supplementary Figures S1 & S13), giving rise to the boundaries of the inverted triangular space containing the rating data. That is, as valence becomes more unambiguous, the absolute correlation between bipolar valence and arousal ratings approaches 1. However, (2) arousal ratings can vary independently from valence ratings when valence is more ambiguous. That is, arousal can be manipulated independently from valence, as valence approaches the midpoint (e.g., ambiguous surprised facial expressions, music, etc.—Figure 2, Supplementary Figures S1, S4, S6A, S7, S8A, & S15). Therefore, a V-shaped function alone is an incomplete description of the data, and an inverted triangular space bounded by this function is more accurate by including the central aspects of the space.

Regarding the bivariate approach: (3) unipolar positivity and unipolar negativity ratings can vary independently when valence is unambiguous. That is, when items are clearly positive, negativity is fixed at zero and the degree of positivity can vary. Similarly, when items are clearly negative, positivity is fixed at zero and the degree of negativity can vary. However, (4) unipolar positivity ratings and unipolar negativity ratings co-vary (inversely) when valence is ambiguous. That is, as valence becomes more ambiguous (i.e., when positivity and negativity are are both non-zero and equal to each other), the correlation between positivity and negativity will approach −1.

[8]Approximately half of the variance along the arousal dimension is *lost* when predicted using separate, unipolar valence scales (dataset 6: $R^2 = 0.390$; dataset 7: $R^2 = 0.714$; see Supplementary Materials III). These effect sizes are consistent with previously reported work (e.g., $R^2 = 0.52$ in Kron, et al. 2013). This effect can also be seen with gross visual inspection—imagine the data in Figure 2 rotated 45 degrees clockwise and compare to Figure 3. The warping of the data structure (i.e., variance along the arousal dimension is Figure 2 is collapsed onto a reciprocal curve in Figure 3) implies that the dimensions of unipolar positivity and unipolar negativity are *not* a rigid rotation of the dimensions of bipolar valence and arousal (as was originally suggested by Watson & Tellegen 1985; see Supplementary Materials IV.B for additional analyses that demonstrate this effect).

Overall, these observations in some ways conflict with the existing 2-dimensional frameworks of affect outlined in this report. First, the fact that arousal ratings often co-vary with valence intensity ratings contradicts the assumption that valence and arousal are largely independent constructs. On the other hand, the inverse correlation between unipolar positivity and unipolar negativity ratings contradicts the assumption that positivity and negativity are largely independent constructs. Neither pair of dimensions is reliably independent in practice, but rather, we find that their correlation systematically depends on the valence ambiguity of the events being rated. In the next section, we propose a mathematical model that formalizes this interpretation of the data.

Model: A mathematical formalization of the conditional relationships between affective ratings

The observed relationships between the dimensional ratings investigated here (arousal & bipolar valence; positivity & negativity)—which have been established in the literature (see Background section), quantitatively verified by the analyses presented in this report (see Data section), and verbally described as points 1 through 4 in the previous paragraph—can be mathematically stated with a set of three equations. First, for any given event, the relationship between its bipolar valence rating (v) and its arousal rating (a) will be proportional to its valence ambiguity ($P_+ - P_-$):

$$v = (P_+ - P_-)\, a \quad (1)$$

where P_+ and P_- are the probabilities that the event will be categorized as 'positive' or 'negative', respectively, in a 3-alternative forced choice task (3AFC) including the options: 'positive', or 'negative', or 'no emotion'. The outcome of this 3AFC task is a ternary variable [P_+ , P_-, P_0], where P_+ is the probability that a given event will be categorized as 'positive', P_- is the probability that a given event will be categorized as 'negative', and P_0 is the probability that a given event will be categorized as 'no emotion'.

Consistent with the rating data, equation 1 necessitates that valence ratings and arousal ratings will be equal when valence is unambiguous (i.e., if either $P_+ = 1$ or $P_- = 1$, this reduces equation 1 to $v = a$ or $v = -a$, respectively). However, as valence approaches the midpoint, arousal ratings become independent of valence (i.e., as ($P_+ - P_-$) approaches 0, then v approaches a constant [0] and is no longer a function of a).

Since P_+ and P_- are computed from a 3AFC task, they have properties consistent with the observed unipolar positivity and unipolar negativity ratings. First, P_+ can vary even when P_- is fixed at zero, and vice versa. Second, the probabilities for each response option must sum to 1 ($P_+ + P_- + P_0 = 1$), which necessitates that P_+ and P_- have an inverse correlation overall. Therefore, two additional equations are implied for predicting unipolar positivity ratings (*pos*) and unipolar negativity ratings (*neg*). First, for any given event, the positivity rating (*pos*) is equal to the arousal rating (*a*) times the probability that the event is positive (P_+):

$$pos = (P_+)\,a \quad (2)$$

Furthermore, for any given event, the negativity rating (*neg*) is equal to the arousal rating (*a*) times the probability that the event is negative (P_-):

$$neg = (P_-)\,a \quad (3)$$

In the remainder of this section, we test equations 1, 2, and 3, which together predict three variables of interest: bipolar valence ratings (*v*), unipolar positivity ratings (*pos*), and unipolar negativity ratings (*neg*), respectively, from two measured variables: arousal ratings (*a*) and a valence ambiguity measure ($P_+ - P_-$; derived from a 3AFC task with the choices positive, negative or no emotion).

Method

In dataset 5, we collected arousal and bipolar valence ratings (*a*, *v*) for a set of 83 items. In dataset 7, we collected unipolar positivity and unipolar negativity ratings (*pos*, *neg*) for these same items by an independent group of participants. Participants who provided the unipolar ratings for dataset 7 also categorized these items in a 3AFC task (positive, negative, no emotion). These categorizations were made either before or after providing the unipolar ratings (order of tasks was randomized). A detailed description of the 3AFC task is included in the description of data collection procedures for dataset 7 (Supplementary Materials I-E). These 3AFC data were then used to estimate ($P_+ - P_-$) for each item, by calculating the percentage of trials that each item was placed into a given category across participants (each participant categorized each stimulus item one time):

$$P_+ = (\#\ \text{of participants categorizing item as positive}) / (\text{TOTAL}\ \#\ \text{of participants}) \quad (5)$$

$$P_- = (\#\ \text{of participants categorizing item as negative}) / (\text{TOTAL}\ \#\ \text{participants}) \quad (6)$$

Results

To test if equation 1 held true for these 83 items, we examined whether the product of *a* (arousal ratings) and $P_+ - P_-$ (an estimate of valence ambiguity computed from the 3AFC categorizations made by an independent group of participants) was equivalent to *v* (the bipolar valence ratings). Indeed, the product of *a* and ($P_+ - P_-$) for each item provided a highly accurate prediction of the bipolar valence rating (v) for that item. One can see visibly that the bipolar valence ratings predicted by the model (Figure 4B) take on the same structure as the actual bipolar valence ratings (Figure 4A) when both are plotted against the arousal dimension. These predictions accounted for over 95% of the variance in the observed bipolar valence ratings ($R^2 = 0.97$; see Supplementary Figure S17A for a plot of this fit). Equation 1 was also highly accurate when predicting bipolar valence ratings for single participants (mean $R^2 = 0.88$ across all participants; see Supplementary Figure S17B for a plot of fits for each participant).

To test equations 2 and 3, we examined whether the product of a (again, the arousal ratings from dataset 5) and P_+ or P_-, respectively, could predict *pos* (unipolar positivity ratings from dataset 7) or *neg* (unipolar negativity ratings from dataset 7). The product of a and P_+ (equation 2) accounted for more than 90% of the variance in positivity ratings ($R^2 = 0.92$; Supplementary Figure 17C), and the product of a and P_- (equation 3) accounted for more than 90% of the variance in negativity ratings ($R^2 = 0.94$; Supplementary Figure 17D). Notably, the reciprocal relationship observed in the actual unipolar ratings (Figure 4C) is also observed in the model's predicted unipolar ratings (Figure 4D).

Summary

The proposed mathematical model predicts over 90% of the variance along three dimensions of interest (bipolar valence = 97%, unipolar negativity = 94%, and unipolar positivity = 92%) using two measurements (arousal, valence ambiguity). These R^2 values suggest that equations 1 through 3 provide a more accurate model of the observed rating data compared to other 2-dimensional frameworks (which in this report accounted for up to 70% of the variance in the rating data—an estimate that is consistent with previously reported data; e.g., Ito et al., 1998; Kron et al., 2013). Specifically, a framework that employs a bipolar valence rating cannot fully account for the variance in separate, unipolar positivity and unipolar negativity ratings (e.g., Supplemental Figure S16); on the other hand, a framework that employs unipolar positive and unipolar negative ratings cannot fully account for the variance in arousal ratings (e.g., Supplemental Figure S15). Therefore, the present model, which also employs only two measured variables (arousal and valence ambiguity), outperforms the other two alternatives.

General Discussion

The relationship between pairs of affective ratings (bipolar valence × arousal; positivity × negativity) is conditional on valence ambiguity. First, when valence is unambiguous (clearly positive or clearly negative), ratings of bipolar valence and arousal are highly correlated; however, as valence becomes more ambiguous, these measurements become more uncorrelated. Second, where, measurements of positivity and negativity are concerned, they are inversely correlated when valence is more ambiguous; however, as valence becomes more unambiguous, these measurements become more uncorrelated. We have demonstrated that these effects are readily observable in subjective rating data, and have presented a mathematical model that formally defines these primary affective dimensions in terms of these effects. The model capitalizes on these relationships to define the boundaries of a triangular 2-dimensional affective space and predicts where ratings will fall within this space. Therefore, this model captures the general structure of the observations at hand, predicting the joint distribution of the measured variables as opposed to predicting a particular outcome variable (that is, the model is generative rather than discriminative). The space captured by this model is depicted in Figure 5.

Theoretical considerations

The proposed model offers a theoretical description of the primary dimensions of subjectively reported affect. First, valence is depicted as a bipolar dimension that moves from negative to positive *through maximal ambiguity* (rather than neutrality). Second, in line with existing theoretical statements, the structure captured by the model suggests that the dimension of subjective arousal distinguishes affective events from those that lack an affective quality.

The midpoint of bipolar valence is maximal ambiguity—The model depicts maximal ambiguity as the midpoint of a bipolar valence dimension, in line with reliable effects observed in the rating data. First, the triangular boundary constraining bipolar valence and arousal ratings suggests that if neutrality is assumed to be the midpoint of a bipolar valence scale (as in the bipolar approach—see Introduction), valence intensity will be confounded with arousal. Second, the reciprocal fit between positivity and negativity reveals that when either positivity or negativity is maximal, the other is minimal, so positivity and negativity cannot be maximal at the same time (as allowed in the bivariate framework—see Introduction). Placing maximal ambiguity at the midpoint of valence, as in the proposed model, accounts for these effects in the data.

Defining valence ambiguity—In this report, valence ambiguity is a term used to describe any event that has both a positive and a negative meaning. This definition is consistent with general behavioral psychology theories where, for example, an extinguished cue that has acquired both a negative and a positive meaning (e.g., a tone that predicts shock sometimes and safety sometimes) is considered to be ambiguous (e.g., Bouton, 1994; Whalen, 1998). In the next two paragraphs, we elaborate on this specific example of ambiguity as a particularly potent illustration of how we modeled valence ambiguity in this report:

Imagine a rat learns that a tone consistently predicts a shock. After this acquisition phase, the valence of the tone is perfectly unambiguous (the tone is negative). If we then have the same tone consistently predict no shock in an extinction phase, the tone now has two meanings: one meaning is negative (shock) and one meaning is positive (no shock). The behavioral literature tells us that extinction is not unlearning—that is, during extinction, the animal does not unlearn the initial negative meaning acquired in the initial conditioning phase (Bouton, 2002). Instead, the animal now has two possible valence interpretations of the tone, so the valence of the tone is now ambiguous.

To bring this example back to the current data, we can think about what valence “rating” the rat will produce in response to the ambiguous, extinguished tone. When one measures how rats interpret the tone a day after their extinction training (by measuring their freezing behavior, for example) the group largely shows a range of intermediate freezing responses (i.e., not as high as during acquisition, and not as low as baseline freezing responses recorded before training began; see e.g., Fig. 2a in Milad & Quirk, 2002). Even though each rat's experience with the tone within each training phase was wholly unambiguous ($P_+ = 0$, $P_- = 1$ during acquisition; and $P_+ = 1$, $P_- = 0$ during extinction), the intermediate reactions to the tone after this inconsistent training looks like a version of intuitive averaging (see Epstein, 1983), whereupon the rats respond with some intermediate freezing magnitude (for example, $P_+ = 0.7$, $P_- = 0.3$, depending on the behavioral bias of the particular animal).

We think surprised facial expressions are similar to extinguished tones in that they have a similar inconsistent reinforcement history (see Oler, Quirk & Whalen, 2009) and when presented in an experimental context, participants provide valence ratings that also seem to intuitively average their previous inconsistent experiences with this expression category. The fact that valence ratings of surprised facial expressions fall within in the central aspects of affective space, but without an appreciable increase in response variance (i.e., uncertainty), is consistent with the model definition of valence ambiguity (i.e., $P_+ - P_-$).

The role of subjective arousal in 2-dimensional affect—In addition to valence ambiguity, the second dimension in this model is arousal. A notable feature of the proposed model is that the valence dimension gets more and more constrained as arousal decreases. This is analogous to the perceptual phenomenon of color, where the saturation of color becomes more and more constrained as brightness decreases. In other words, there can be no variation in color when brightness is minimal. Similarly, we observe little to no variation in valence ratings when arousal is minimal. Early dimensional models of emotion explicitly illustrate this phenomenon geometrically (Schlosberg, 1954; Osgood 1966).

This geometric characterization of arousal is consistent with many emotion theories proposing a necessary role for physiological arousal in emotion (e.g., James, 1884; Lange, 1885/1922; Schachter & Singer, 1962; Thayer, 1989; etc.). Still, the necessity of physiological arousal in emotional experience has been contested (e.g., Reisenzein, 1983). In the proposed model, arousal is defined by subjective report rather than physiology. The structure captured by the model suggests that the *subjective* dimension of arousal (see Table S2 for more specific definitions of this term, which has not always been consistently defined; e.g., Lindsley, 1988), separate from physiological arousal, plays a fundamental role in how affective experiences are distinguished from those that lack noticeable affective quality. This is not a new idea, as the subjective arousal dimension is sometimes described as a general "emotion" dimension[9] (e.g., Schlosberg 1954; Ortony, Clore, & Collins, 1988; Mehrabian, 1996).

Methodological considerations

Valence ambiguity & arousal

In this report we measure valence ambiguity, and show that when taken in combination with arousal ratings* (using equations 1-3), we can then derive bipolar valence ratings, unipolar

[9]From Mehrabian (1996): "The [arousal] scale can be viewed as measuring emotionality" (p. 266)

positivity ratings, and unipolar negativity ratings. In comparison, other measurement approaches (bipolar or bivariate) may not fully capture the variance in some particular affective rating: (1) using arousal and bipolar valence scales concedes the unique variance captured by separate positivity and negativity scales (specifically, variance in the non-dominant unipolar affect of ambiguously valenced events; Supplementary Materials IV-A) and (2) using separate positivity and negativity scales concedes a portion of the variance captured by an arousal scale (specifically, variance in the arousal of ambiguously valenced events; Supplementary Materials IV-B). By comparison, the use of a valence ambiguity measurement in combination with arousal ratings captures nearly all of the variance (over 90%) in subjectively reported bipolar valence, unipolar positivity, and unipolar negativity.

Measuring valence ambiguity

In this report, valence ambiguity was estimated for any given item using single-trial 3AFC categorizations from a group of participants. On the surface, this task seems a course measurement strategy, forcing participants to choose one of three extreme, possibly hypothetical labels for a given event. That is, there is likely no actual event that is purely negative, purely positive, or entirely lacking in affective quality. However, estimates of valence ambiguity gain precision with repeated measures. To elaborate, the present measure of valence ambiguity yields a continuous proportion variable that can take on any value between −1 and 1, with its precision limited by the number of repeated measures. In combination with arousal ratings, the current approach predicts bipolar valence ratings as well as separate unipolar positivity and unipolar negativity ratings—thus the apparent coarseness of this measure does not result in loss of information according to the predictions.

Measuring subjective arousal

In this report, arousal was estimated for any given item using single-trial Likert scale ratings from a group of participants, consistent with previous work (see Introduction). The precise way in which the arousal scale is worded or defined via instructions to participants has not been agreed upon (e.g., see Lindsley, 1988; Barrett & Russell, 1999). Traditionally, terms used to measure arousal are often positively valenced (Watson et al., 1999; e.g., “calm” as the low anchor or “excited” as the high anchor; e.g., Kron et al., 2013; see also Supplemental Figure S4). In this report, the arousal dimension was simply anchored with the terms “low” and “high”, and the wording of the question was designed to avoid any terms with a clear valence (e.g., “Rate the strength of your emotional response” or “Rate the emotional intensity of the face/image”; see Supplementary Materials V for more discussion on this methodological issue). Further investigation might be able to capture interesting differences that arise from changes in instructions, wording, or anchoring of this dimensional scale. Nonetheless, changes along these parameters do not seem to influence the gross structure of the data (e.g., Compare Figures 1 and 2).

---

[*]With respect to the question of whether an arousal rating is actually necessary (since arousal ratings correlate with P0 to some degree), or if the ternary valence ambiguity variable alone is sufficient for capturing this 2-dimensional space, see Supplemental Materials I-E (specifically Supplemental Figure S11 and the associated text).

How do I measure dimensional affect in my experiment?

The major conclusion of the model we have presented here, is that valence ambiguity dictates the degree to which particular affective ratings are confounded with each other (i.e., correlated). In general, there is no point in collecting any second measurement that will be perfectly confounded with some initial measurement. As one example, if an experimenter was interested in studying disgust responses, the entire design might be limited to unambiguously negative stimulus items. In this case, we can predict that unipolar negativity ratings, arousal ratings, and bipolar valence ratings will all be nearly perfectly correlated for these items, which will fall on a single line along the Negative Intensity axis illustrated in Figure 5. Because unambiguously negative items (such as disgusting images) only vary along a single axis in 2D affective space, such an experiment would need only a single scale to measure 2D affect (i.e., unipolar negativity). Alternatively, if an experimenter were investigating responses to both clearly negative items and clearly positive items, ratings of these items would fall on the V-shaped boundary, along both the Negative Intensity axis and the Positive Intensity axis illustrated in Figure 5. In this case, a single bipolar valence scale would be sufficient to capture variance in subjective ratings along these axes, since valence and arousal ratings will be correlated for these items. As a final example, if an experimenter wishes to dissociate subjective valence from subjective arousal, then measuring both valence ambiguity and arousal would be an optimal strategy, since the model predictions reveal these variables to more adequately capture variance in 2-dimensional affective ratings overall (we elaborate on this point in the next section).

Implications for experimental design: conditions of high valence ambiguity

Valence ambiguity influences the degree to which arousal and valence ratings co-vary, so experimental designs will want to take this into account when selecting stimulus items, if the goal is to dissociate effects of valence from effects of arousal (e.g., Small, et al. 2003; Anderson, Christoff, Stappen, Panitz, Ghahremani, Glover, ... & Sobel, 2003; Cunningham, Raye, & Johnson, 2004; Dolcos, LaBar, & Cabeza, 2004; Kensigner & Corkin, 2004; Lewis, et al. 2007). Ideally, investigating this dissociation would require an experiment that manipulates valence while holding arousal constant and/or manipulates arousal while holding valence constant. The latter requires items with valence ratings near the midpoint (ambiguous), because arousal and valence are confounded at the boundaries of the space. The former requires arousal to be substantially greater than zero, because valence can only be manipulated within a restricted range when arousal is close to zero. In order to hold arousal constant, however, valence must be manipulated from negative—through *ambiguous* —to positive.

Potential Caveats & Limitations

We have presented a model that proposes a mathematically defined space for two primary dimensions of affect (valence and arousal), offering a potential resolution for points of theoretical disagreement in existing models of these dimensions. However, there are recognizable limitations to this model that we highlight here.

Group-level effects versus participant-level effects

As noted throughout this manuscript, our analyses and model are focused on group-level analyses, in which ratings were averaged across participants. Previous work has usefully questioned whether the inverted triangular structure observed here is applicable at the level of the individual (Kuppens, Tuerlinckx, Russell, & Barrett, 2013). Here, we show that Equation 1 can predict bipolar valence ratings for individual participants, but we do not fully explore the application of this model to idiographic case studies. When items are treated as a random factor rather than participants, the rating data largely fall within triangular structure similar to the plots in the main text here (see Supplemental Figures S1B, S2B, S3B, S4B, S6C, and S10A), but the structure in these individual-participant data is visibly noisier. We expect that the valence ambiguity of any object/event/mood being rated will interact with individual differences in affective responding, although future work will be needed to fully test that idea.

The frequency of ambiguity

One may argue that the current report has over-emphasized the variance contributed by the arousal dimension because we highlight the role of items that may not occur all that frequently (i.e., ambiguously valenced items with high arousal). That is, items with very high arousal appear much more likely to fall at the extreme ends of the valence dimension. Does this mean that an item falling on the valence midpoint, but high in arousal, is purely theoretical? Possibly, but we note that the effects described in this report were achieved with a limited number of ambiguous items[10]. In other words, this limited subset of items is precisely what drives the observed dissociation between valence intensity and arousal in the rating data presented here. Therefore, although the distribution of ratings within the triangular space does not appear to be uniform or normal, this did not mitigate the capacity of the proposed model to describe the rating data better than the alternative measurement approaches overall. That said, perhaps in daily life it is actually the extremes of emotion that are experienced less frequently, as everyday life often consists of more subtle, ambiguous emotional moments that depend on contextual information to interpret valence and arousal value (Cannon, 1928; Ellsworth, 1994; Whalen, 1998).

Does the model generalize to ambivalent events?

It is possible that some of the items we have categorized as ambiguously valenced events include items that might be more appropriately included under a related phenomenon studied in social psychology, referred to as 'ambivalence' (e.g., Cacioppo et al., 1997; Ito et al., 1998; Larsen, 2016). In that line of research, the term 'ambivalence' is used to refer to the simultaneous occurrence of positive and negative feelings. Examples of such events include a moving day (Larsen et al., 2001), where you are excited to start over but 'still gotta carry this stuff outta here.'

[10]Examples of what we employed as ambiguously valenced stimulus items include: (1) surprised facial expressions, which can occur in response to both positive and negative events (e.g., Kim, et al., 2003); (2) standardized images that contain both positive and negative aspects (e.g., IAPS images used in Ito et al., 1998; for example, one image contains a picture of a boy with his dog—the boy has a broad smile on his face but the dog is snarling); (3) several music clips that evoke an emotional quality that could be seen as both positive and negative (e.g., a tense moment of Beethoven's 5th Symphony); (4) portrait shots of celebrities that the general population both praises and criticizes (i.e., Kim Kardashian, see Supplemental Figure S7).

Proponents of the bipolar approach have suggested that simultaneous positive and negative qualities of a given stimulus event will offset each other (see Russell, 2016). This view is consistent with our definition of valence ambiguity as the proportion of negativity versus positivity. On the other hand, proponents of the bivariate approach have suggested that positivity and negativity are separable under conditions of ambivalence (Larsen, Norris, McGraw, Hawkley, & Cacioppo, 2009; Cacioppo et al., 2012; Larsen, 2016). This view is consistent with our inclusion of the ‘no emotion’ option in the operationalization of valence ambiguity, which allows for positivity and negativity to be predicted separately.

Nonetheless, there is a possibility that the present study has under-sampled ambivalent events. Ambivalent qualities may inherently require more complexity and longer time courses (e.g., a moving day or a graduation; Larsen et al., 2001) compared to the events presented in this report (which lasted no more than 5 seconds). Furthermore, the items in this report were not selected to distinguish ambiguity from ambivalence, but rather, to sample a host of commonly used stimulus items in psychological research, with an emphasis on employing items that fall in the central aspects of 2-dimensional affective space (i.e., ambiguously valenced items). Our view, not directly tested here, is that ambiguous is a term that best describes events that could lead to positive or negative outcomes, while ambivalence is a term that best describes events that evoke positive and negative feelings simultaneously. Future studies might seek to determine the relevance of the present model for ambivalent events (if they are indeed distinct from what we are calling ambiguously valenced events), especially in comparison to the documented strengths of the bivariate approach.

Positivity offset

The proposed model does not perfectly fit the data, and we note an obvious error here for transparency. There is a pronounced non-linearity in predictions of unipolar positivity ratings (see the bottom left hand corner of Supplementary Figure S17C)[11]. Namely, predicting unipolar positivity ratings for neutral items consistently results in values that are lower than the actual unipolar positivity ratings for those items. Previously published work has established that unipolar positivity ratings are slightly higher for neutral items compared to negative items (an effect called “positivity offset”; e.g., Cacioppo, et al. 1997; Norris et al., 2010; Cacioppo et al., 2012), and the unipolar positivity ratings presented here also show this effect. In contrast, our measurement of valence ambiguity shows P_+ is approximately zero for both neutral and negative items (see Supplementary Figure S11C). Overall, this error accounts for ~5% of the variance in unipolar positivity ratings (which is most of the 8% of variance that is not accounted for by the proposed model). Adjustments to the proposed model's unipolar positivity predictions (equation 2) would be required to accurately capture the positivity offset effect.

[11]Presumably, this is why the R^2 for unipolar positivity ratings is lower than R^2 for both unipolar negativity ratings and bipolar valence ratings.

Perceived versus felt affect

The proposed model does not distinguish between affective ratings made with respect to (a) an external object or (b) an internal subjective feeling, which is often an important psychological distinction when interpreting affective ratings (e.g., Gabrielsson, 2002; Russell, 2003). Although evidence exists showing that both types of ratings can conform to a similar structure (e.g., Watson et al., 1999; Kurdi et al., 2016), other work suggests that a triangular structure may not constrain ratings of current feelings like ratings of experimental events (e.g., Kuppens et al., 2013). Additional work will be needed to investigate the generalizability of the present model in this respect.

Additional complexity in affective experience

Any 2-dimensional framework represents a deliberate simplification of affective experience, in the service of experimental expediency and theoretical traction. Much work has highlighted additional factors, beyond the two dimensions highlighted here, that are important for characterizing affective responses, such as: dominance (e.g., Russell & Mehrabian, 1977; Mehrabian, 1996), unpredictability (e.g., Fontaine, Scherer, Roesch, & Ellsworth, 2007), or more fine grained emotion categories (Tomkins, 1963; Ekman, 1992; Young, Rowland, Calder, Etcoff, Seth, & Perrett, 1997; Sievers et al., 2013; Jacks, Garrod, & Schyns, 2014; Watson & Stanton, 2016). In general, we do not see a 2-dimensional model as in conflict with more complex models, but hopefully complementary to them.

Conclusions

When encountering any environmental event of biological and social relevance, we implicitly ask ourselves two fundamental questions: Is this good or bad (valence), and how important is that answer to me (arousal)? In this report, we provided a reinterpretation of the relationship between valence and arousal ratings and the space that constrains them, as routinely measured by affective and social scientists. We showed that the relationship between these affective dimensions is conditional—dependent on the valence ambiguity of the event being rated. That is, arousal and valence are correlated when valence is clear, but not when valence is ambiguous. Positivity and negativity are inversely correlated when valence is ambiguous, but not when valence is clear. Critically, these conditional relationships constrain these ratings within a triangular space. We then presented a mathematical model that formalizes this structure, comprising two dimensions (valence ambiguity, arousal). This generative model captures a greater degree of variance in these affective ratings compared to alternative two-dimensional frameworks. Given the critical importance of our subjective reports in communicating our emotions to one another, our hope is that this model will inform new theory and methods to further our understanding of how to effectively think about and measure subjective affect, based upon the structure that organizes these experiences.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Our thanks to Bishma Chakrabarti, Alexander Shackman, Carien van Reekum, and Luke Chang for comments on earlier drafts of the model. Additional thanks to Julia Berkowitz, Holly Wakeman, & Kelly Gaudet for assisting with data collection, and Samantha Chavez, M. Justin Kim, James Taylor, Tracy d'Arbeloff, Daisy Burr, & Eshin Jolly, for additional support and feedback. We also thank the anonymous reviewers whose comments greatly improved the quality of this report. These data were presented at the annual Wisconsin Symposium on Emotion (Madison, WI, April 22, 2015), at the Social & Affective Neuroscience Society (SANS) annual meeting (Boston, MA, April 25, 2015), and at the Society for Social Neuroscience (S4SN) annual meeting (Chicago, IL, October 16, 2015).

Funding

This research was supported by the National Institutes of Mental Health of the National Institutes of Health under award number R01MH080716. The content of the reported research is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Abelson RP, Sermat V. Multidimensional scaling of facial expressions. Journal of Experimental Psychology. 1962; 63(6):546–554. [PubMed: 13858932]

Anders S, Lotze M, Erb M, Grodd W, Birbaumer N. Brain activity underlying emotional valence and arousal: A response related fMRI study. Human Brain Mapping. 2004; 23(4):200–209. [PubMed: 15449355]

Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, Sobel N. Dissociated neural representations of intensity and valence in human olfaction. Nature Neuroscience. 2003; 6(2):196–202. [PubMed: 12536208]

Barrett LF. Feelings or words? Understanding the content in self-report ratings of experienced emotion. Journal of Personality and Social Psychology. 2004; 87(2):266–281. [PubMed: 15301632]

Barrett LF, Russell JA. Independence and bipolarity in the structure of current affect. Journal of Personality and Social Psychology. 1998; 74(4):967–984.

Barrett LF, Russell JA. The structure of current affect controversies and emerging consensus. Current Directions in Psychological Science. 1999; 8(1):10–14.

Bensafi, M. Le Traitement Affectif des Odeurs: Aspects Implicites et Explicites. Université Lumière; Lyon, France: 2001. Unpublished Doctoral Thesis

Bensafi M, Rouby C, Farget V, Bertrand B, Vigouroux M, Holley A. Autonomic nervous system responses to odours: the role of pleasantness and arousal. Chemical Senses. 2002; 27(8):703–709. [PubMed: 12379594]

Bouhuys AL, Bloem GM, Groothuis TG. Induction of depressed and elated mood by music influences the perception of facial emotional expressions in healthy subjects. Journal of. 1995; 33(4):215–226.

Bouton ME. Context, ambiguity, and classical conditioning. Current directions in psychological science. 1994; 3(2):49–53.

Bouton ME. Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. Biological psychiatry. 2002; 52(10):976–986. [PubMed: 12437938]

Bradburn, NM. The structure of psychological well-being. Aldine; Chicago: 1969.

Bradley MM, Greenwald MK, Petry MC, Lang PJ. Remembering pictures: pleasure and arousal in memory. Journal of experimental psychology: Learning, Memory, and Cognition. 1992; 18(2):379.

Bradley, MM., Lang, PJ. The International Affective Digitized Sounds (IADS-2): Affective ratings of sounds and instruction manual. University of Florida; Gainesville, FL: 2007. Tech. Rep. B-3

Bradley, MM., Lang, PJ. Technical Report C-1, The Center for Research in Psychophysiology. University of Florida; 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings.; p. 1-45.

Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry. 1994; 25(1):49–59. [PubMed: 7962581]

Britton JC, Taylor SF, Sudheimer KD, Liberzon I. Facial expressions and complex IAPS pictures: common and differential networks. Neuroimage. 2006; 31(2):906–919. [PubMed: 16488159]

Bush LE. Individual differences multidimensional scaling of adjectives denoting feelings. Journal of Personality and Social Psychology. 1973; 25(1):50–57. [PubMed: 4688168]

Cacioppo JT, Berntson GG. Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. Psychological Bulletin. 1994; 115(3):401–423.

Cacioppo, JT., Berntson, GG., Norris, CJ., Gollan, JK. The evaluative space model.. In: Van Lange, P.Kruglanski, A., Higgins, ET., editors. Handbook of theories of social psychology. Vol. 1. Sage Press; Thousand Oaks, CA: 2012. p. 50-72.

Cacioppo JT, Gardner WL, Berntson GG. Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. Personality and Social Psychology Review. 1997; 1(1):3–25. [PubMed: 15647126]

Cacioppo JT, Petty RE, Losch ME, Kim HS. Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. Journal of Personality and Social Psychology. 1986; 50(2):260–268. [PubMed: 3701577]

Cahill L, Haier RJ, Fallon J, Alkire MT, Tang C, Keator D, McGaugh JL. Amygdala activity at encoding correlated with long-term, free recall of emotional information. Proceedings of the National Academy of Sciences. 1996; 93(15):8016–8021.

Cannon, WB. Neural organization for emotional expression.. In: Reymert, ML., editor. Feelings and emotions: the Wittenberg symposium. Clark University Press; Worcester, MA: 1928. p. 257-259.

Chikazoe J, Lee DH, Kriegeskorte N, Anderson AK. Population coding of affect across stimuli, modalities and individuals. Nature neuroscience. 2014; 17(8):1114–1122. [PubMed: 24952643]

Correia SS, McGrath AG, Lee A, Graybiel AM, Goosens KA. Amygdala-ventral striatum circuit activation decreases long-term fear. eLife. 2016; 5:e12669. [PubMed: 27671733]

Cunningham W, Raye C, Johnson M. Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. Journal of Cognitive Neuroscience. 2004; 16(10):1717–1729. [PubMed: 15701224]

Diener E, Emmons RA. The independence of positive and negative affect. Journal of Personality and Social Psychology. 1984; 47:1105–1117. [PubMed: 6520704]

Diener E, Iran-Nejad A. The relationship in experience between various types of affect. Journal of Personality and Social Psychology. 1986; 50(5):1031–1038.

Dolcos F, LaBar KS, Cabeza R. Dissociable effects of arousal and valence on prefrontal activity indexing emotional evaluation and subsequent memory: an event-related fMRI study. Neuroimage. 2004; 23(1):64–74. [PubMed: 15325353]

Ekman P. An argument for basic emotions. Cognition & emotion. 1992; 6(3/4):169–200.

Ellsworth PC. William James and emotion: is a century of fame worth a century of misunderstanding? Psychological Review. 1994; 101(5):222–229. [PubMed: 8022957]

Epstein S. Aggregation and beyond: Some basic issues on the prediction of behavior. Journal of Personality. 1983; 51(3):360–392. [PubMed: 28497598]

Feldman LA. Valence focus and arousal focus: Individual differences in the structure of affective experience. Journal of personality and social psychology. 1995; 69(1):153–166.

Fontaine JR, Scherer KR, Roesch EB, Ellsworth PC. The world of emotions is not two-dimensional. Psychological Science. 2007; 18(12):1050–1057. [PubMed: 18031411]

Gabrielsson A. Emotion perceived and emotion felt: Same or different?. Musicae Scientiae. 2002; 5(1): 123–147.

Green RS, Cliff N. Multidimensional comparisons of structures of vocally and facially expressed emotion. Perception & Psychophysics. 1975; 17(5):429–438.

Green DP, Goldman SL, Salovey P. Measurement error masks bipolarity in affect ratings. Journal of Personality and Social Psychology. 1993; 64(6):1029. [PubMed: 8326466]

Greenwald MK, Cook EW, Lang PJ. Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. Journal of Psychophysiology. 1989; 3(1):51–64.

Ito TA, Cacioppo JT, Lang PJ. Eliciting affect using the International Affective Picture System: Trajectories through evaluative space. Personality and Social Psychology Bulletin. 1998; 24(8): 855–879.

Jacks RE, Garrod OG, Schyns PG. Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. Current Biology. 2014; 24(2):187–192. [PubMed: 24388852]

James W. What is an Emotion?. Mind. 1884; 9(34):188–205.

Jin J, Zelano C, Gottfried JA, Mohanty A. Human Amygdala Represents the Complete Spectrum of Subjective Valence. The Journal of Neuroscience. 2015; 35(45):15145–15156. [PubMed: 26558785]

Kensinger EA. Remembering emotional experiences: The contribution of valence and arousal. Reviews in the Neurosciences. 2004; 15(4):241–252. [PubMed: 15526549]

Kensinger EA, Corkin S. Two routes to emotional memory: Distinct neural processes for valence and arousal. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(9):3310–3315. [PubMed: 14981255]

Kim H, Somerville LH, Johnstone T, Alexander AL, Whalen PJ. Inverse amygdala and medial prefrontal cortex responses to surprised faces. Neuroreport. 2003; 14(18):2317–2322. [PubMed: 14663183]

Kim H, Somerville L, Johnstone T, Polis S, Alexander A, Shin L, Whalen P. Contextual modulation of amygdala responsivity to surprised faces. The Journal of Cognitive Neuroscience. 2004; 16(10): 1730–1745. [PubMed: 15701225]

Knutson B, Katovich K, Suri G. Inferring affect from fMRI data. Trends in Cognitive Sciences. 2014; 18(8):422–428. [PubMed: 24835467]

Kron A, Goldstein A, Lee DHJ, Gardhouse K, Anderson AK. How are you feeling? Revisiting the quantification of emotional qualia. Psychological Science. 2013; 24(8):1503–1511. [PubMed: 23824581]

Kron A, Pilkiw M, Banaei J, Goldstein A, Anderson AK. Are valence and arousal separable in emotional experience? Emotion. 2015; 15(1):35–44. [PubMed: 25664950]

Kuppens P, Tuerlinckx F, Russell JA, Barrett LF. The relation between valence and arousal in subjective experience. Psychological Bulletin. 2013; 139(4):917–940. [PubMed: 23231533]

Kurdi B, Lozano S, Banaji MR. Introducing the Open Affective Standardized Image Set (OASIS). Behavior research methods. 2016:1–14.

Lang PJ. The emotion probe: studies of motivation and attention. American Psychologist. 1995; 50(5): 372–385. [PubMed: 7762889]

Lang PJ, Bradley MM, Cuthbert BN. Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. Biological psychiatry. 1998; 44(12):1248–1263. [PubMed: 9861468]

Lang, PJ., Bradley, MM., Cuthbert, BN. International affective picture system (IAPS): Affective ratings of pictures and instruction manual (Tech. Rep. A-8). University of Florida; Gainesville: 2008.

Lang PJ, Greenwald MK, Bradley MM, Hamm AO. Looking at pictures: Affective, facial, visceral, and behavioral reactions. Psychophysiology. 1993; 30(3):261–273. [PubMed: 8497555]

Lange, C. The emotions. (I. A. Haupt, Trans.). Williams & Wilkins; Baltimore: 1922. Original work published 1885

Larsen JT. Holes in the case for mixed emotions. Emotion Review. 2016

Larsen JT, McGraw AP, Cacioppo JT. Can people feel happy and sad at the same time?. Journal of Personality and Social Psychology. 2001; 81(4):684–696. [PubMed: 11642354]

Larsen JT, Norris CJ, Cacioppo JT. Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. Psychophysiology. 2003; 40(5):776–785. [PubMed: 14696731]

Larsen JT, Norris CJ, McGraw AP, Hawkley LC, Cacioppo JT. The evaluative space grid: A single-item measure of positivity and negativity. Cognition and Emotion. 2009; 23(3):453–480.

Lewin, K. A Dynamic Theory of Personality: Selected Papers. Adams, DK., Zener, KE., editors. McGraw-Hill; New York: 1935.

Lewis PA, Critchley HD, Rotshtein P, Dolan RJ. Neural correlates of processing valence and arousal in affective words. Cerebral Cortex. 2007; 17(3):742–748. [PubMed: 16699082]

Lindquist KA, Satpute AB, Wager TD, Weber J, Barrett LF. The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. Cerebral Cortex. 2015 Advance online publication. doi: 10.1093/cercor/bhv001.

Lindsley, DB. Emotions and the electroencephalogram.. In: Reymert, ML., editor. Feelings and Emotions. McGraw-Hill; New York: 1950.

Lindsley, DB. Activation, arousal, alertness, and attention.. In: Hobson, JA., editor. States of Brain and Mind. Springer; New York: 1988.

Lykken DT, Venables PH. Direct measurement of skin conductance: A proposal for standardization. Psychophysiology. 1971; 8(5):656–672. [PubMed: 5116830]

Leknes S, Tracey I. A common neurobiology for pain and pleasure. Nature Reviews Neuroscience. 2008; 9(4):314–310. [PubMed: 18354400]

Man V, Nohlen H, Melo H, Cunningham W. Hierarchical brain systems support multiple representations of valence and mixed affect. Emotion Review. 2016

Mehrabian A. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. Current Psychology. 1996; 14(4):261–292.

Milad MR, Quirk GJ. Neurons in medial prefrontal cortex signal memory for fear extinction. Nature. 2002; 420(6911):70–74. [PubMed: 12422216]

Miller NE. Learning resistance to pain and fear: effects of overlearning, exposure, and rewarded exposure in context. Journal of Experimental Psychology. 1960; 60(3):137–145. [PubMed: 14422565]

Neta M, Norris CJ, Whalen PJ. Corrugator muscle responses are associated with individual differences in positivity-negativity bias. Emotion. 2009; 9(5):640–648. [PubMed: 19803586]

Nikolova YS, Knodt AR, Radtke SR, Hariri AR. Divergent responses of the amygdala and ventral striatum predict stress-related problem drinking in young adults: possible differential markers of affective and impulsive pathways of risk for alcohol use disorder. Molecular psychiatry. 2016; 21(3):348–356. [PubMed: 26122584]

Nowlis V, Nowlis HH. The description and analysis of mood. Annals of the New York Academy of Sciences. 1956; 65(4):345–355. [PubMed: 13373215]

Oler, JA., Quirk, GJ., Whalen, PJ. Cingulo-amygdala interactions in surprise and extinction: Interpreting associative ambiguity. In: Vogt, BA., editor. Cingulate Neurobiology and Disease. Volume 1: Infrastructure, Diagnosis, and Treatment. Oxford University Press; Oxford: 2009. p. 207-218.

Ortony, A., Clore, GL., Collins, A. The cognitive structure of emotions. Cambridge University Press; Cambridge, MA: 1988.

Osgood CE. Dimensionality of the semantic space for communication via facial expressions. Scandinavian Journal of Psychology. 1966; 7(1):1–30. [PubMed: 5908205]

Reisenzein R. The Schachter theory of emotion: two decades later. Psychological Bulletin. 1983; 94(2):239–264. [PubMed: 6356197]

Russell JA. A circumplex model of affect. Journal of Personality and Social Psychology. 1980; 39(6): 1161–1178.

Russell JA. Mixed emotions viewed from the psychological constructionist perspective. Emotion Review. 2016

Russell JA, Carroll JM. On the bipolarity of positive and negative affect. Psychological Bulletin. 1999; 125(1):3–30. [PubMed: 9990843]

Russell JA, Mehrabian A. Evidence for a three-factor theory of emotions. Journal of research in Personality. 1977; 11(3):273–294.

Schachter S, Singer J. Cognitive, social, and physiological determinants of emotional state. Psychological Review. 1962; 69(5):379–399. [PubMed: 14497895]

Schlosberg H. Three dimensions of emotion. Psychological Review. 1954; 61(2):81–88. [PubMed: 13155714]

Schlosberg H. The description of facial expressions in terms of two dimensions. Journal of Experimental Psychology. 1952; 44(4):229–237. [PubMed: 13000062]

Schimmack U. Pleasure, displeasure, and mixed feelings: Are semantic opposites mutually exclusive?. Cognition & Emotion. 2001; 15(1):81–97.

Shepard RN. The analysis of proximities: Multidimensional scaling with an unknown distance function. II. Psychometrika. 1962; 27(3):219–246.

Sievers B, Polansky L, Casey M, Wheatley T. Music and movement share a dynamic structure that supports universal expressions of emotion. Proceedings of the National Academy of Sciences. 2013; 110(1):70–75.

Small DM, Gregory MD, Mak YE, Gitelman D, Mesulam MM, Parrish T. Dissociation of neural representation of intensity and affective valuation in human gustation. Neuron. 2003; 39(4):701–711. [PubMed: 12925283]

Thayer, RE. The biopsychology of mood and arousal. Oxford University Press; 1989.

Tomkins, SS. Affect, imagery, consciousness: Vol. 2. The negative affects. Springer Publishing Company; New York: 1963.

Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Nelson C. The NimStim set of facial expressions: judgments from untrained research participants. Psychiatry research. 2009; 168(3):242–249. [PubMed: 19564050]

Urry HL, Van Reekum CM, Johnstone T, Kalin NH, Thurow ME, Schaefer HS, Davidson RJ. Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. The Journal of neuroscience. 2006; 26(16):4415–4425. [PubMed: 16624961]

Watson D, Tellegen A. Toward a consensual structure of mood. Psychological Bulletin. 1985; 98(2): 219–235. [PubMed: 3901060]

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

Watson D, Stanton K. Emotion blends and mixed emotions in the hierarchical structure of affect. Emotion Review. 2016

Whalen PJ. Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. Current directions in psychological science. 1998; 7(6):177–188.

Woodsworth, RS., Schlosberg, H. Experimental Psychology (Rev. ed.). Holt, Rinehart, and Winston, Inc; New York: 1954. First edition published in 1938

Wundt, WM. An Introduction to Psychology.. In: Pintner, R., D'Hooghe, M., editors. George Allen & Unwin, Ltd; London: 2014. Available from www.gutenberg.org/ebooks/46677. (Original work published in 1912, reprinted in 1920)

Yik MS, Russell JA, Barrett LF. Structure of self-reported current affect: Integration and beyond. Journal of Personality and Social Psychology. 1999; 77(3):600–619.

Young AW, Rowland D, Calder AJ, Etcoff NL, Seth A, Perrett DI. Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. Cognition. 1997; 63(3):271–313. [PubMed: 9265872]

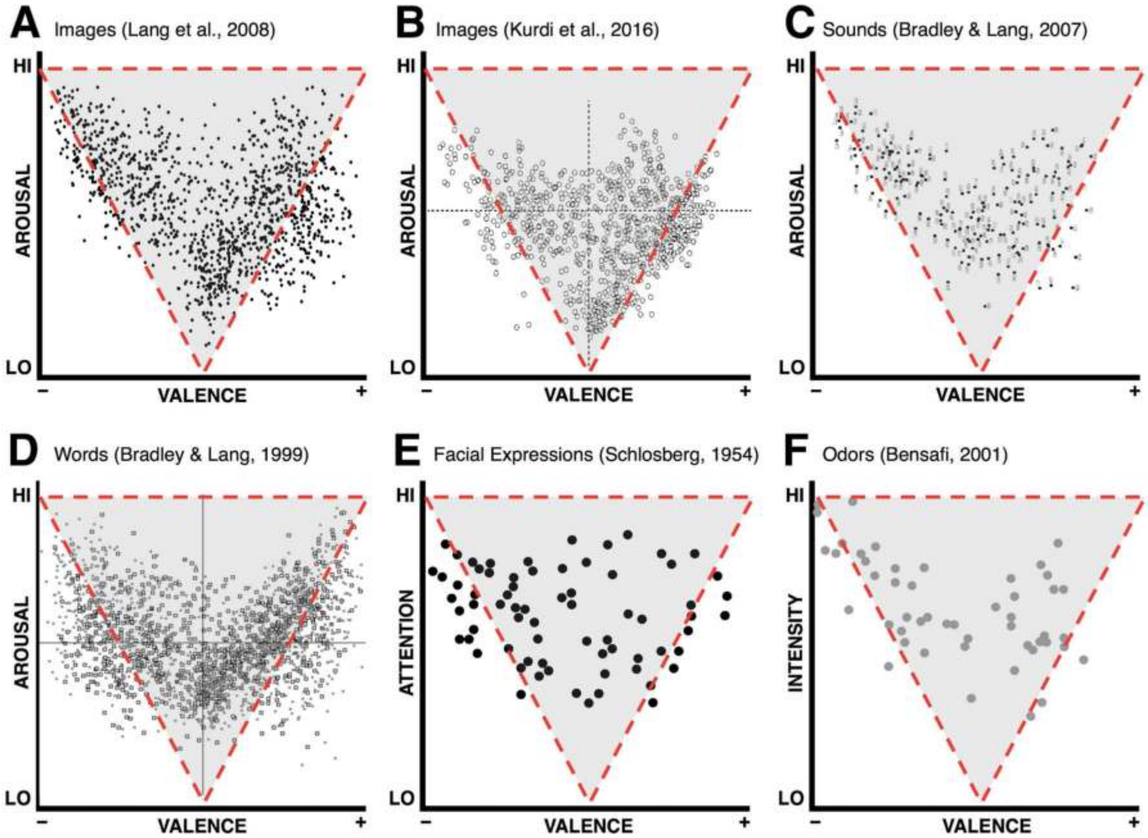

Figure 1.

Ratings of bipolar valence and arousal are often correlated on either side of the valence midpoint, such that the majority of data points fall within an inverted triangular envelope. Although some points fall outside this general constraint, even these outlying points tend to follow the overall structure. Note that, while the rating data in this figure were derived from Likert scale ratings, this inverted triangular pattern is also observed when these dimensions are derived indirectly with dimensionality reduction techniques (e.g., Shepard, 1962; Abelson & Sermat, 1962; Green & Cliff, 1975; Watson et al., 1999).

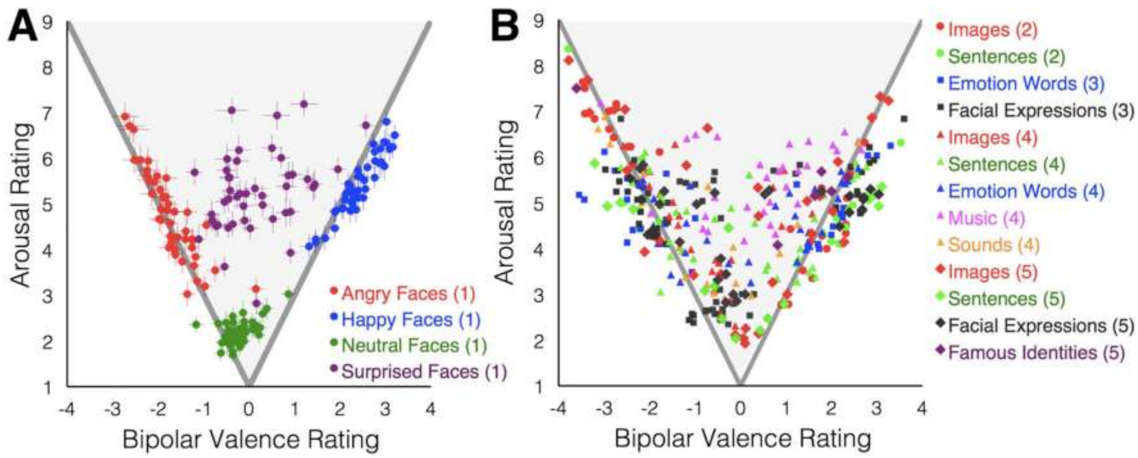

Figure 2. Bipolar valence ratings and arousal ratings fall within a triangular boundary

This structure occurs for ratings of **(A) facial expressions (dataset 1)**. Note that clearly valenced facial expressions (happy, angry) fall near the boundary, while ambiguously valenced facial expressions (surprise) fall more distant from the boundary. In turn, valence and arousal ratings are highly correlated for clearly valenced expressions (forming the V-shaped boundary seen in many datasets), but for ambiguously valenced expressions, valence and arousal ratings are notably less correlated. This structure also occurs for ratings of (B) other affective events of variable stimulus modalities (datasets 2-5), including images, music, sounds, words, and sentences. (For plots that include the standard errors for these items—which are comparable to the standard errors for dataset 1 items—see Supplementary Materials I). **Note:** Each dataset described in Table 1 is depicted here by plotting each item according to its mean bipolar valence rating (x-axis) and mean arousal rating (y-axis), averaged across participants. Plotting each participant (averaging across items) reveals a structure that is consistent with the data here (see Supplementary Figures S1B, S2B, S3B, S4B, S6C, and S10A).

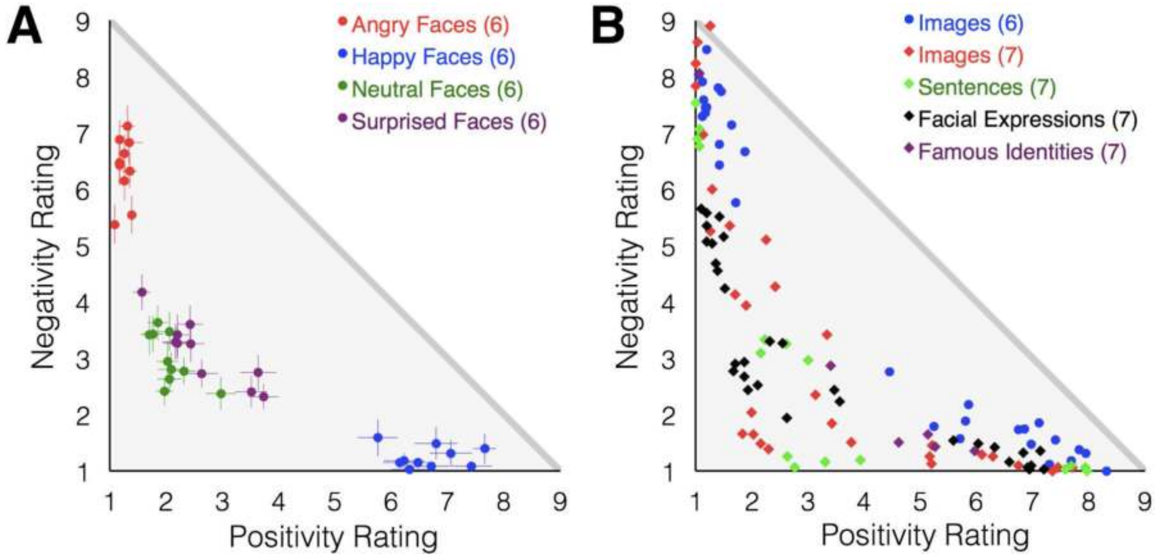

Figure 3. Unipolar positivity ratings and unipolar negativity ratings have a mathematically reciprocal relationship (y = 1 / x), and so do not fall in the upper diagonal of the space

(A) Facial expressions (dataset 6) show this structure **and (B) other modalities (datasets 6 & 7)** also show this structure: IAPS images (dataset 6); additional facial expression categories, famous identities, images, and sentences (dataset 7). Across both of these panels, each item is plotted according to its mean unipolar positivity rating (x-axis) and mean unipolar negativity rating (y-axis). The ratings shown here have been averaged across participants for each item, but this same structure can be observed when the ratings are plotted for each participant (see Supplementary Figures S5B & S10B).

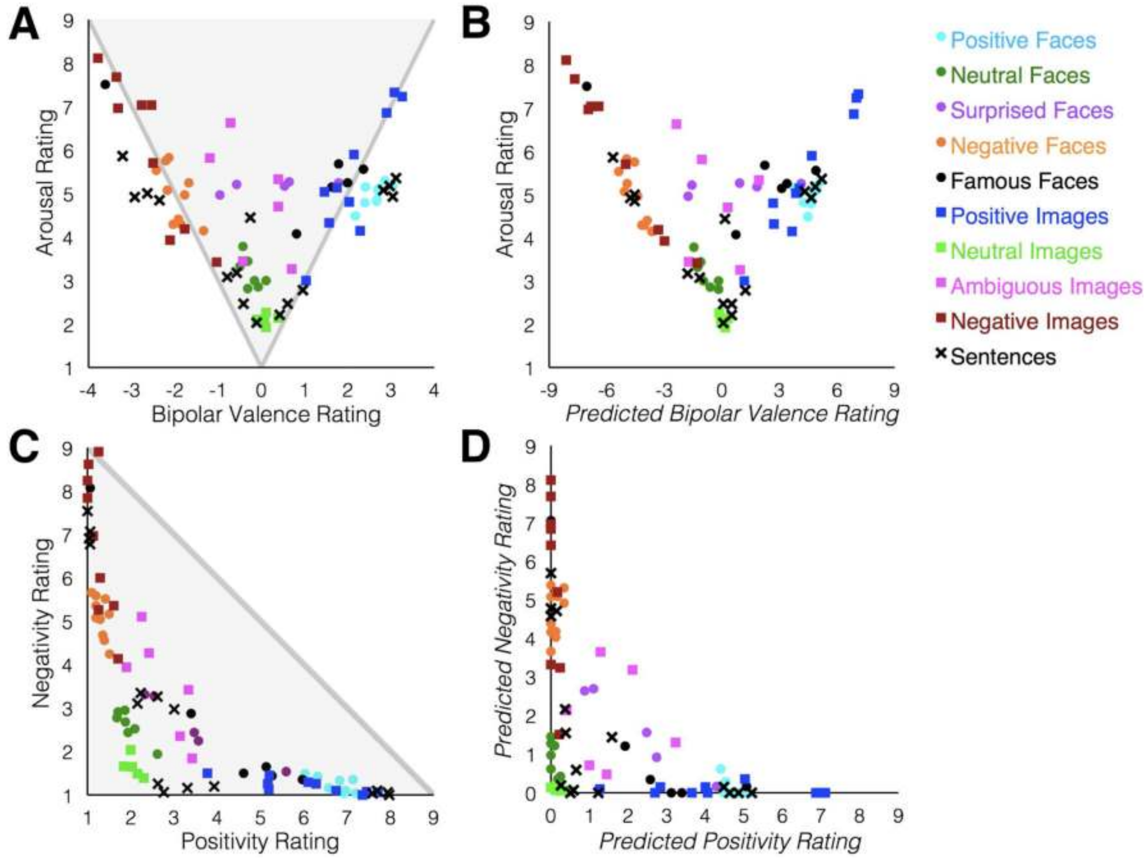

Figure 4. The structure of (A) observed bipolar valence × arousal ratings and (C) observed unipolar positivity × negativity ratings can be reproduced by 3AFC valence categorizations in combination with arousal ratings (B and D), respectively, using the model equations

(A) *Observed* bipolar valence ratings (x-axis) and *observed* arousal ratings (y-axis); these data are also plotted in Figure 2B and Supplementary Figures S7A, S8A, S9A, & S10A. **(B)** Equation 1's *predicted* bipolar valence ratings (x-axis) plotted against *observed* arousal ratings (y-axis). **(C)** *Observed* positivity rating (x-axis) and *observed* negativity rating (y-axis) for a set of 83 items; these data are also plotted in Figure 3B and Supplementary Figures S7B, S8B, S9B, & S10B. **(D)** Equations 2's *predicted* positivity ratings (x-axis) and Equation 3's *predicted* negativity ratings (y-axis).

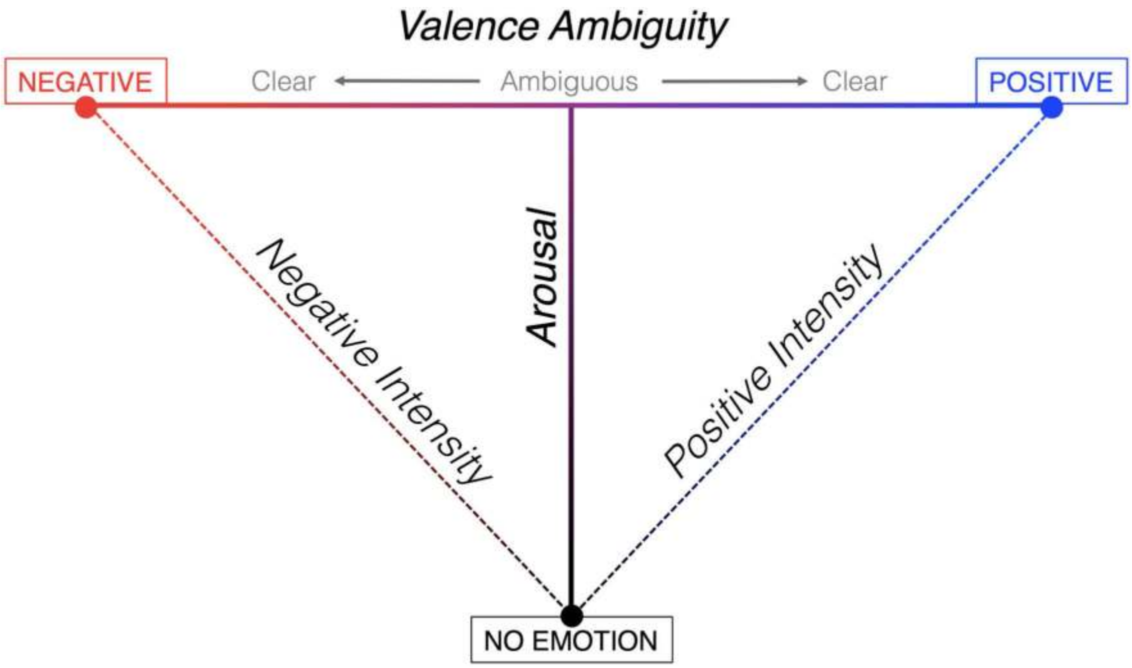

Figure 5.
A graphical depiction of the space defined by equations 1 through 3. The range of a bipolar valence dimension becomes smaller as arousal approaches zero. Stimulus items that are ambiguous with respect to valence will fall in the more central aspects of the space, while more clearly valenced items will fall along the boundaries of the space.

Table 1

Overview of datasets where affect was measured with bipolar valence and arousal scales.

Dataset	$N_{PARTICIPANTS}$ (N_{FEMALE})	Description of Stimulus Items	N_{ITEMS}
1	45 (29)	**Faces** (45 identities expressing: happy, angry, surprised, neutral; from in-house photos)	180
2	19 (12)	**Images** (30 images from the IAPS; Lang et al., 2008) **Sentences** (12 affective scenarios; adapted from Kim et al., 2004)	42
3	35* (26)	**Words** (28 emotion words; from Russell, 1980) **Faces** (10 identities expressing: happy, calm, afraid, sad, surprised, neutral; from Tottenham et al. 2009)	88
4	34 (22)	**Images** (15 Self Assessment Manikin icons, from Bradley & Lang, 1994; 12 schematic facial expression, from Bouhuys et al., 1995) **Sentences** (40 emotion statements; from Barrett & Russell, 1998) **Words** (31 emotion adjectives; from Barrett, 2004) **Music** (32 clips of 5 seconds; symphonic, rock, or pop; no lyrics) **Sounds** (16 clips of ≲5 seconds; human vocalizations, etc.)	146
5	29 (22)	**Images** (30 images; subset of images in dataset 2 + Google image search) **Sentences** (16 sentences adapted from sentences in datasets 2 & 4) **Faces** (37 faces; subset of faces in datasets 1 & 3 + additional identities)	83

*These participants are also listed in Table 2.

Table 2

Overview of datasets where affect was measured with unipolar positivity and unipolar negativity scales.

Dataset	$N_{PARTICIPANTS}$ (N_{FEMALE})	Description of Stimulus Items	N_{ITEMS}
6	35* (26)	**Faces** (40 faces; subset of faces in dataset 1) **Images** (30 images; same as dataset 2)	70
7	34 (26)	**Faces** (37 faces; same as dataset 5) **Images** (30 images; same as dataset 5) **Sentences** (16 sentences; same as dataset 5)	83

*These participants are also listed in Table 1.