2012

Subjective emotions vs. verbalizable emotions in web texts

International Journal of Psychology and Behavioral Sciences,Rosemead : Scientific and Academic,v. 2, n. 5, p. 173-184, 2012

http://www.producao.usp.br/handle/BDPI/49954

International Journal of Psychology and Behavioral Sciences 2012, 2(5): 173-184
DOI: 10.5923/j.ijpbs.20120205.08

Subjective Emotions vs. Verbalizable Emotions in Web Texts

Sergey Petrov[1], José F. Fontanari[2], Leonid I. Perlovsky[3,*]

[1]Independent consultant, Cambridge, 02138, USA
[2]Instituto de Física de São Carlos, Universidade de São Paulo, São Paulo, Caixa Postal 369, Brazil
[3]Athinoula A. Martinos Center for Biomedical Imaging, Harvard University, Charlestown, 02129, USA

Abstract Cognition and emotions are inseparable. Still, it is not clear to which extent emotions can be characterized by words and how much of emotional feelings are non-verbalizable. Here we approach this topic by comparing the structure of the emotional space as revealed by word contexts to that in subjective judgments, as studied in the past. The number of independent emotions and categories of emotions is a key characteristic of the emotional space. Past research were based exclusively on perceived subjective similarities by participants of experiments. Here we propose and examine a new approach, the similarities between emotion names are obtained by comparing the contexts in which they appear in texts retrieved from the World Wide Web. The developed procedure measures a similarity matrix among emotional names as dot products in a linear vector space of contexts. This matrix was then explored using Multidimensional Scaling and Hierarchical Clustering. Our main findings, namely, the underlying dimension of the emotion space and the categories of emotion names, were consistent with those based on subjective judgments. We conclude that a significant part of emotional experiences is verbalizable. Future directions are discussed.

Keywords Subjective Emotions, Verbalizable Emotions, Emotion Contexts, Basic Emotions, Multidimensional Scaling, Hierarchical Clustering, WWW Texts

1. Introduction

The concept of emotion has been used in various ways by different authors, sometimes denoting different physiological or mental mechanisms, and often without any attempt to clarify its intended meaning[1]. There has been a variety of attempts to define emotion in the many disciplines where emotions are relevant. For instance, according to Grossberg & Levine[2] emotions are neural signals indicating satisfaction or dissatisfaction of instinctual needs (drives) – not unlike Simon's view on emotions[3]. Emotions have been related to survival[4] and to facial expressions[5]. Emotions have been associated with references to internal, mental conditions[6].

Emotions have been tied to social interactions, arising from the dissonance people feel between competing goals and conflicting interpretations of the world[7]; this view has also been influential in the belief–desire theory of emotion[8,9]. Cabanac[10] defined emotions as any mental experience with high intensity and high hedonic content(pleasure-displeasure). Emotions underlie human creativity[11,12]. Most authors seem to agree on emotions performing appraisals of situations or events[13,14,15]. But even if appraisal is the main function of emotions, it is still not clear if emotional appraisal is equivalent to conceptual appraisal, and to which extent it could be expressed in words.

Another controversial issue related to emotions is the attempt to define the basic emotions. Although it is in general agreed that basic emotions evolved with fundamental life tasks[16,17,18,19,20], there are about 14 different proposals of emotion candidates for this category, whose size vary from 2 to 11 members[21]. Similarly to the role of primary colors in vision, all other (non-basic) emotions could be thought of as a composition of a few basic emotions[4,20]. The idea of basic emotions having specialized neurophysiological and anatomical substrates or being the primitive building blocks of other, non-basic, emotions has been criticized[21]; Izard[22] argued in support of this idea based on the interactions between emotions and cognition. We refer the reader to Gratch et al.[8] for a recent overview of the interplay of cognition and emotion. Some cognitive scientists use the name 'discrete' instead of basic[1,23]. We would also mention that language acquisition is separate to some extent from acquisition of cognitive representations[24,25,26,27,28,29,30]; one only requires experience with surrounding language, whereas the other requires life experience in the real world[31,32]. Emotions in language

* Corresponding author:
leonid@seas.harvard.edu(Leonid I. Perlovsky)
Published online at http://journal.sapub.org/ijpbs

connect language to real life experience.

Grossberg and Levine[2] theory relating emotions to instincts could have been used for relating basic emotions to basic (or bodily) instincts; this direction of research was not pursued to our knowledge. Perlovsky[33,34,35,36,37,38] argued that human 'higher' emotions, such as aesthetic, musical, and cognitive dissonance emotions are related to the instinct for knowledge, a 'higher' instinct[23,39,40,41, 42]; that these emotions are principally different from basic emotions, and their number is much larger being better described by a continuum of emotions rather than by discrete labels. Steps toward experimental test of this hypothesis were made in[39,43,44,45].

Regardless of the experts' theories and disputes on emotions, people have an informal and implicit naïve theory of emotion which they use in their daily routines to interact and influence other individuals. Emotion words are labels for the categories of the folk taxonomy of emotional states, and have an immense importance in clinical and personality psychological evaluations which use mood adjective (emotion name) checklists to assess the patients' emotional states[46]. A relationship of emotion words to 'true' psychological emotions is a separate scientific problem, that we touch in this contribution. As the manner humans perceive color similarities can tell much about the physical distance (in terms of the wavelengths) between the colors[47,48] it could be expected that the way people think and talk about emotions may bear some relationship to psychological emotional states[49,45].

Most, if not all, quantitative approaches to understanding the underlying organization of the emotion categories have focused on perceived similarities among (English) emotion names. A remarkable outcome of this research avenue was the finding that emotion names are not independent of each other[46]. Attempts to produce a representation or structural model to capture the relationships among the emotion word categories have led to the proposal of the Circumplex model in which emotion names are arranged in a circular form with two bipolar dimensions interpreted as the degree of pleasure and the degree of arousal[50,51]. In that sense, emotion names mix together in a continuous manner like hues around the color circle[46,52]. This suggestion of the limited variety of basic emotions corresponds to defining emotions in terms of prototypical scripts or scenarios[53].

A complementary approach to the structural models of emotion names categories is the exploration of the hierarchical structure of those categories[54]. This more intuitive approach allows the immediate identification of the basic emotions categories as those that are closer to the root of the hierarchical tree. Both approaches use mathematical techniques to extract relevant information from a similarity matrix produced by asking individuals to rate the similarity between a given set of distinct emotion words.

The procedure to obtain the similarities among the emotion words is the main feature that distinguishes our contribution from the landmark papers mentioned in the previous paragraphs. Rather than asking individuals to rate the similarities using a fixed discrete scale, we search for texts in the Web that contain emotion names and define the similarity among a specified set of target emotion names - essentially the same set used in the study of Shaver et al[54] – as the number of common words in the close neighborhoods (contexts) of the target emotion names. This definition allows us to express the similarities as dot-products between vectors in the space of contexts and then use the full power of linear algebra for its analysis. In particular, we follow the original multidimensional scaling framework[55] and re-express these similarities as dot-products of orthogonal vectors in the space spanned by the target emotion names. These vectors, known as principal coordinates, are rescaled eigenvectors of the similarity matrix. The estimated dimension of the emotion space is consistent with the estimates based on the individuals' judgment of the similarities. Regarding the hierarchical clustering analysis, our clusters exhibit a good correlation to those produced by Shaver et al.[54].

The rest of this paper is organized as follows. In Section 2 we describe the procedure used to extract texts from the Web. This section also contains our definition of the similarity between pairs of emotion names and its mathematical interpretation as a dot product in the linear vector space of contexts. The resulting similarity matrix S is analyzed in Section 3. We begin with elementary statistical measures and then proceed to the Multidimensional Scaling Analysis, S. Torgerson[55]. The section ends with the presentation of the categories into which the emotion names are grouped according to Ward's minimum variance hierarchical clustering algorithm[56]. Our findings are discussed in Section 4. Finally, Section 5 presents concluding remarks and outline future research directions.

2. Methods

Practically all methods employed in the literature to investigate the closeness of common emotion names were based on querying participants about the similarity and differences between a given set of emotion names[46,54,57,58]. Our approach departs from the traditional psychology methods in that we gauge the similarity between two emotion names by comparing the contexts in which they are used in documents extracted from the Web. At the present stage, we do not explore the semantic information contained in those texts; rather our comparison is based solely on the shared vocabulary between documents.

2.1. Target Emotion Names

Although contemporary English contains hundreds of terms with emotional connotations[59], apparently there is no consensus on which of these terms can be considered emotion names or emotion prototypes. An ingenious approach to this issue was offered by Shaver et al[54], who presented a list of 213 candidate emotion names to 112 students and asked them to rate those terms on a 4-point scale ranging from ‘I definitely would not call this an emotion’ to ‘I definitely would call this an emotion’. This procedure resulted in a much shorter list containing 135 emotion names that the participants rated highest on the 4-point ‘emotionness’ scale. In addition to these 135 emotion names we have included 7 more names, namely, anticipation, acceptance, wonder, interest, aversion, pain, and courage in order to take into account a few widely recognized ‘basic’ emotions[21], which were not in the original list of that study. Table 1 shows the 135 emotion names from the list of Shaver et al[54] together with the 7 names mentioned above, totaling 142 emotion names which we use as target words in our Web queries, as described next.

2.2. Context Retrieval

For every target emotion name listed in Table 1, we retrieve 99 documents containing the target word from the Web using the Yahoo! search engine. Thus, the documents were ordered by Yahoo! relevance criteria. Since, as expected, almost every target emotion word is used in a variety of semantic contexts which are unrelated to emotions (e.g., ‘ecstasy’) and many of them appear in advertising (e.g., names of restaurants), our search focused on documents in which the target emotion word is combined with the word ‘emotion’. This combination more or less restricted the retrieved documents to ones where a particular emotion – or at least an emotion word – was the subject of the text. This combination – target emotion name plus the word ‘emotion’ - increased considerably the average length of the retrieved documents.

These retrieved texts were then cleaned up for the purpose of forming the so called bags of words. A bag of words is a list of words in which the grammatical rules are ignored. During the cleanup, all words of length 2 or shorter were eliminated. In addition, we have also filtered out conjunctions, prepositions, pronouns, numbers, punctuations marks and all formatting signs. In what remained of each document, we then selected a sequence of 41 consecutive words with the target emotion name in the middle, i.e., 20 words before and 20 words after the target word. Only the 50 more relevant (according to Yahoo!) contexts were retained for every emotion name. For some of the emotion names used by Shaver et al[54], namely, tenderness, thrill, caring, sentimentality, longing, cheerfulness, enjoyment, contentment, enthrallment, amazement, astonishment and nervousness, we were unable to retrieve 50 contexts of the prescribed length out of the 99 retrieved ones, and so we excluded those words (numbered 131 to 142 in Table 1) from our list of target emotion names. We indexed these 130 emotion words by $i = 1,\dots 130$.

In summary, for each of the first 130 target emotion names exhibited in Table 1, we produced 50 distinct sequences of words, each containing 20 valid words before and 20 valid words after the target word in question. A valid word is a word that escaped the cleanup procedure applied to the Web documents retrieved by the Yahoo! search engine. The final step is to lump all the 50 sequences corresponding to a given target emotion name, say word i, into a single bag of words which we denote by W_i. (Note that W_i is not a set since an element can be present there more than once). Hence, the number of elements in a bag of words is $50 \times (20+1+20) = 2050$, regardless of the emotion index-name $i = 1,\dots 130$. We note that there are only $K = 34244$ distinct words among the 266500 words that make up the 130 word bags.

2.3. Similarity Measure

The basic similarity $\hat{S}_{ij}$ between the two target emotion words i and j is calculated using their corresponding bags of words, W_i and W_j, as follows. Let us denote by $w_{ij}(k)$ the number of times word k from W_i appears in the bag W_j. Note that $w_{ij}(k)$ and $w_{ji}(k)$ have different domains since there might be words that belong to W_i but not to W_j, and vice-versa. The unprocessed similarity $\hat{S}_{ij}$ is defined as

$$\hat{S}_{ij} = \sum_k w_{ij}(k) \tag{1}$$

where k runs over all words (repetitions included) in W_i. This procedure takes into account multiple appearances of words in bags W_i and W_j. For example, if word k from bag W_i appears m times in bag W_j and n times in bag W_j then it contributes with the factor mn to the unprocessed similarity $\hat{S}_{ij}$. From this example we can easily realize that the similarity measure defined by eq. (1) is symmetric, i.e., $\hat{S}_{ij} = \hat{S}_{ji}$ for all pairs of target emotion words i and j. In the case the bags W_i and W_j consist of the same word repeated n times we have $\hat{S}_{ij} = n^2$, whereas if W_i and W_j do not have any element in common we have $\hat{S}_{ij} = 0$.

We note that our definition of the unprocessed similarity, eq. (1), is equivalent to a dot-product of two vectors. In fact, let us order all the $K = 34244$ distinct words alphabetically (the specific order is not essential for the argument). Then we can represent each bag W_i uniquely by a K-dimensional vector $\mathbf{Y}_j = (Y^1{}_j,\dots Y^K{}_j)$ in which the component $Y^l{}_j$ is the number of times that word l appears in W_j. Hence our unprocessed similarity measure can be written as the dot-product

$$\hat{S}_{ij} = \sum_k Y^l{}_i \; Y^l{}_j \tag{2}$$

Of course, the vectors $\mathbf{Y}_j$ are very sparse, i.e., most of their components are zeros. We will offer a more economic representation of the similarity entries in terms of a much lower dimension vector space in Section 3. Equation (2) is important because it shows that $\hat{S}_{ij}$ is a dot-product which allows us then to use the full power of linear algebra for its analysis. Even more important, however, is the observation that $\hat{S}_{ij}$ provides little information about the proximity or closeness of the vectors $\mathbf{Y}_i$ and $\mathbf{Y}_j$, unless these vectors are normalized. In fact, on the one hand almost orthogonal vectors can have a very high unprocessed similarity value if their norms are large and, on the other hand, two parallel vectors may have a low similarity if their norms are small. This problem can be easily corrected by defining the normalized similarity as

$$S_{ij} = \hat{S}_{ij} / \sqrt{(\hat{S}_{ii}\ \hat{S}_{jj})} \qquad (3)$$

where $\hat{S}_{ii}$ is the squared norm of the vector $\boldsymbol{Y}_j$, $\| \boldsymbol{Y}_j \|^2$. Note that $S_{ij} \in [0, 1]$ and $S_{ii} = 1$. Henceforth we will refer to the normalized similarity, eq. (3), simply as the similarity between emotion index-names *i* and *j*.

2.4. Null Random Model

Since our approach is based on the statistics of word contexts, we should also define a 'null' model using random contexts, so that we could identify which results depend on contexts specific to emotion words, and which ones characterize random contexts. A null model to compare our results can be obtained as follows. First, we lump together the 130 word bags into a single meta-bag comprising 266500 elements. Next we pick 2050 (41 x 50) words at random and without replacement to form the random bag F_1. This drawing step is repeated to form the remaining word bags F_2,… F_{130}, which ends when the meta-bag is emptied. Given these randomly assembled word bags, we then follow the procedure described before to calculate the normalized entries R_{ij} of the random similarity matrix $\boldsymbol{R}$ between emotion names *i* and *j*.

Table 1. List of the 135 emotion names obtained by Shaver et al[54] (1987) plus the 7 basic emotion names (anticipation, acceptance, wonder, interest, aversion, pain, and courage). Only the first 130 emotion names in this list were used to generate the 130 x 130 similarity matrix **S**

1 : acceptance	30 : disappointment	59: grouchiness	88 : melancholy	117 : sympathy
2 : adoration	31 : disgust	60 : grumpiness	89 : misery	118 : tenseness
3 : affection	32 : dislike	61 : guilt	90 : mortification	119 : terror
4 : aggravation	33 : dismay	62 : happiness	91 : neglect	120 : torment
5 : agitation	34 : displeasure	63 : hate	92 : optimism	121 : triumph
6 : agony	35 : distress	64: homesickness	93 : outrage	122 : uneasiness
7 : alarm	36 : dread	65 : hope	94 : pain	123 : unhappiness
8 : alienation	37 : eagerness	66: hopelessness	95 : panic	124 : vengefulness
9 : amusement	38 : ecstasy	67 : horror	96 : passion	125 : woe
10 : anger	39 : elation	68 : hostility	97 : pity	126 : wonder
11 : anguish	40 : embarrassment	69 : humiliation	98 : pleasure	127 : worry
12 : annoyance	41 : enthusiasm	70 : hurt	99 : pride	128 : wrath
13 : anticipation	42 : envy	71 : hysteria	100 : rage	129 : zeal
14 : anxiety	43 : euphoria	72 : infatuation	101 : rapture	130 : zest
15 : apprehension	44 : exasperation	73 : insecurity	102 : regret	131 : tenderness
16 : arousal	45 : excitement	74 : insult	103 : rejection	132 : thrill
17 : attraction	46 : exhilaration	75 : interest	104 : relief	133: caring
18 : aversion	47 : fear	76 : irritation	105 : remorse	134: sentimentality
19 : bitterness	48 : ferocity	77 : isolation	106 : resentment	135: longing
20 : bliss	49 : fondness	78 : jealousy	107 : revulsion	136: cheerfulness
21 : compassion	50 : fright	79 : jolliness	108 : sadness	137: enjoyment
22 : contempt	51 : frustration	80 : joviality	109 : satisfaction	138: contentment
23 : courage	52 : fury	81 : joy	110 : scorn	139: enthrallment
24 : defeat	53 : gaiety	82 : jubilation	111 : shame	140: amazement
25 : dejection	54 : gladness	83 : liking	112 : shock	141: astonishment
26 : delight	55 : glee	84 : loathing	113 : sorrow	142: nervousness
27 : depression	56 : gloom	85 : loneliness	114 : spite	
28 : desire	57 : glumness	86 : love	115 : suffering	
29 : despair	58 : grief	87 : lust	116 : surprise	

3. Results

In this section we use two techniques often employed in the investigation of the underlying psychological structure of the use of emotion words by English speakers, namely, multidimensional scaling and hierarchical clustering analysis (Shepard, 1980; Shaver et al., 1987). However, before we introduce these more involved exploratory tools, we present a simple statistical description of the 130 x 130 similarity matrix S, as well as of its random counterpart R, generated according to the procedure described in the previous section. In the following analysis we disregard the diagonal elements of both matrices.

3.1. Simple Statistical Measures

We begin with the most elementary measures, namely, the mean **Sm** and the standard deviation σ_S of the entries of the similarity matrix **S**. We find **Sm** = 0.376 and σ_S = 0.069. For the random null model, these two quantities are **Rm** = 0.686 and σ_R = 0.020. The considerable differences between even these simple measures indicate a rich underlying structure of the matrix S. A better visualization of the en-tries of these matrices is obtained by ordering the entries according to their rank, from the largest to the smallest. Disregarding the diagonal elements, there are 130 x 120 /2 = 8385 distinct entries in each of these matrices and in Fig. 1 we present their values as function of their ranks r = 1,… 8385. These distributions are remarkably symmetric around their mean values, shown by the horizontal lines in the figure. For the most part of the rank order range, say 2000 < r < 7000, the similarity values decrease linearly with the rank r. In particular, in this range we find the following equations for trends $S_r \approx 0.463-2.1\ 10^{-5}\ r$ and $R_r \approx 0.463-2.1\ 10^{-6}\ r$. These results indicate that the random similarity matrix ***R*** is much more homogenous than ***S***, which is expected since in the random model the contexts do not provide information to distinguish between the target words.

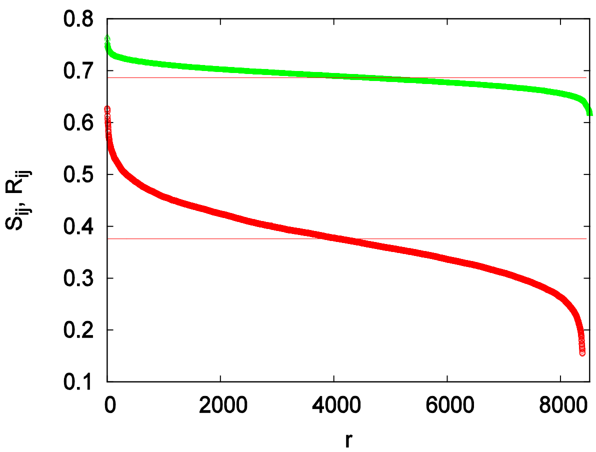

Figure 1. Plot of the off-diagonal entries of the similarity matrix *S* (circles, lower curve) and its random version *R* (triangles, upper curve) as function of their ranks r = 1,… 8385. The horizontal solid lines are the mean values *Sm* = 0.376 and *Rm* = 0.686. These results corroborate the expectation that *R* is more homogeneous than *S*

To conclude this section it is instructive to write out the pair of words corresponding to the extremes of the rank order distribution. For instance, the pair of distinct words with the largest similarity, S_1 = 0.627, is *aggravation* and *irritation*; the pair with the second largest, S_2 = 0.626, is *anguish* and *gloom*; the pair with the third largest, S_3 = 0.623, is *mortification* and *shock*; and the pair with the smallest similarity, S_{8385} = 0.155, is *hostility* and *glee*. The high similarity of emotion words of close emotional content lends credibility to our experimental procedure.

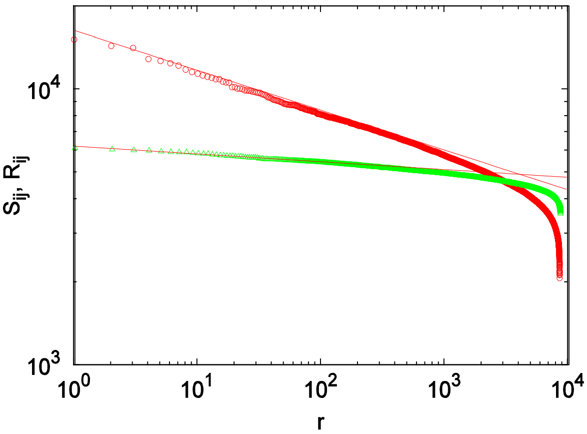

Figure 2. Log-log plot of the off-diagonal entries of the unprocessed similarity matrix $\hat{S}$ (circles) and its random version $\hat{R}$ (triangles) as function of their ranks r = 1,…8385 . The solid lines are the fitting of equation (4) with b = 16280 and a = 0.15 for $\hat{S}$ and b = 6200 and a = 0.03 for $\hat{R}$. These results corroborate the expectation that $\hat{R}$ is more homogeneous than $\hat{S}$

3.2. Multidimensional Scaling

Similarly to most psychological experimental settings aiming at exploring the relationships between *n* emotion words[46,54], the end product of our data-mining methodology is a *n x n* similarity matrix ***S***. It is thus tempting to assume the existence of a subjacent 'emotion' vector space of dimension *m* that contains vectors whose Euclidian scalar product generates ***S***. The mathematical procedure to derive a base of this vector space is known as Multidimensional Scaling Analysis[55,60,61,62]. More explicitly, we want to find the set of *m* orthogonal vectors of length *n*, $(x_1^a, x_2^a, \ldots x_n^a)$ with $a = 1,\ldots m$, such that

$$S_{ij} = \sum_{a=1\ldots m} x_i^a \; x_j^a \qquad (5)$$

for all pairs i,j = 1,…n . This problem has a simple and neat solution in the case *m*=*n*. In fact, denoting the eigenvectors of S by $(u_1^a, u_2^a, \ldots u_n^a)$ we write the well-known formula for the decomposition of the entries of a matrix

$$S_{ij} = \sum_{a=1\ldots n} \lambda^a u_i^a \; u_j^a \qquad (6)$$

so that the prescription

$$x_i^a = \sqrt{\lambda^a}\, u_i^a \qquad (7)$$

provides the desired solution to our problem. Of course, in the case of interest *m* < *n* and for a general matrix ***S***, eq. (5)

has no solution. A popular approach is to use (7) considering only the *m* largest eigenvalues in the expansion (6). The quality of the approximation can then be measured by the stress function[62,63]

$$Q = \sqrt{\Sigma_{ij}(S_{ij} - S^*_{ij})^2 / \Sigma_{ij} S_{ij}^2} \qquad (8)$$

where S_{ij} is taken as (6) $a = 1,\ldots m$, with the eigenvalues ordered such that $\lambda^1 \leq \lambda^2 \leq,\ldots \lambda^n$. Nowadays the designation Multidimensional Scaling Analysis is applied to the numerical minimization of Q using gradient descent techniques, in which only the rank order of the entries of $\boldsymbol{S}$ are used[48,63]. Our procedure follows the original formulation of the Multidimensional Scaling proposed by Torgerson[55].

The issue now is to pick a 'representative' value for the dimension *m*. A large value of *m* yields a very low stress value (we recall that $Q = 0$ for $m = n$ by construction) but then most of the dimensions may not be relevant to describe the underlying structure of the similarity matrix as they are likely to be determined more by noise than by the essential structure of S. Alternatively, a too small value of m may not reproduce the similarity matrix with sufficient accuracy. A popular heuristic method to determine the 'optimal' dimensionality is the so-called elbow test in which the stress Q is plotted against *m*, as done in Fig. 3. Ideally, such a graph should exhibit an 'elbow' indicating that, after a certain value of *m*, the rate of decrease of the stress function becomes negligible. The results of Fig. 3 are not so discrepant from this idealistic expectation, for $m \geq m^* \approx 30$, the dimension at which the concavity of the stress function nearly vanishes, the rate of decrease of Q is approximately 0.001, the slope of the solid straight line shown in the figure. However, there is a lot of subjectivity in the estimate of the critical dimension m^* based on the elbow test, as illustrated by the two fitting straight lines in Fig. 3. In fact, the fitting of the dashed line, in which we have eliminated the first point ($m=1$) because of its interpretation as random noise, yields a much lower estimate for this critical dimension, $m^* \approx 5$.

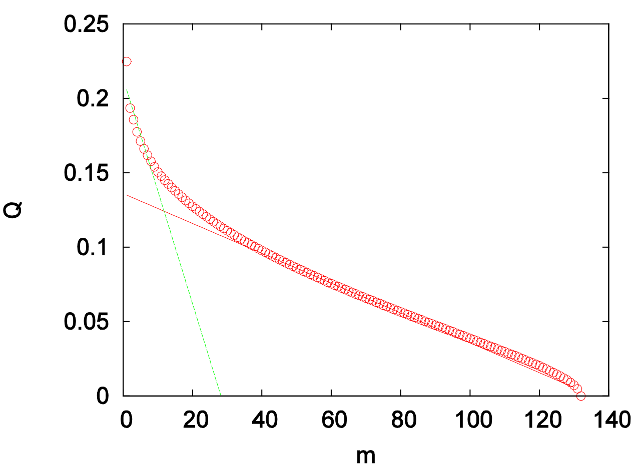

Figure 3. Elbow test showing the stress function Q as defined in Eq. (8) against the number of dimensions *m* of the underlying word emotion space. The solid straight line has the slope -0.001 whereas the dashed line has slope -0.007

It is instructive to compare the stress function Q with the numerical values of the eigenvalues in Fig. 4. As expected, there is a qualitative resemblance between these quantities, since the spectral decomposition of the matrix $\boldsymbol{S}$, eq. (5), tells us that the largest eigenvalues are the most important to the reconstruction of $\boldsymbol{S}$. A small slope (-0.004) of the fitting straight line indicates that eigenvalues with a > 30 represent a small percentage of the data for all pairs i, j. A similar argument holds for a > 5. Thus the issue boils down to quantifying how small the percentage representation of the neglected eigenvectors must be; at present, this is a subjective decision of the researcher.

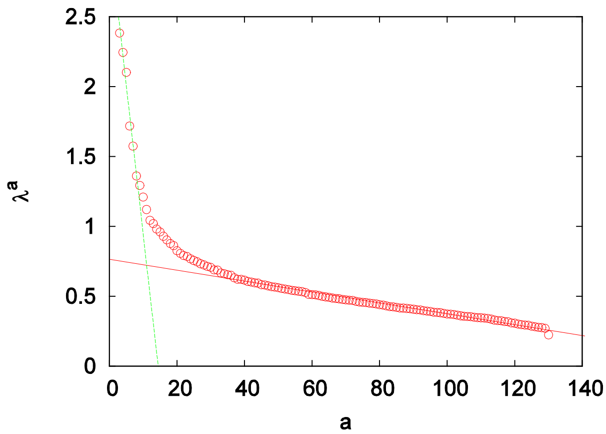

Figure 4. Eigenvalues of the similarity matrix *S* ordered according to their rank from largest to smallest. The first ($\lambda^1 = 49.99$) and the second ($\lambda^2 = 4.44$) largest eigenvalues are omitted from the figure. The solid straight line has the slope 10.004, whereas the dashed line has slope -0.21

As already pointed out, a similar mathematical analysis of the dimension of the emotion word space in which participants rated the similarity between the emotion words yielded a much smaller value for m^*, typically $m^* = 2$[46] and $m^* = 3$[54]. Of utmost interest for the Multidimensional Scaling Analysis (as well as for the Factor Analysis) is the interpretation of the eigenvectors associated to the largest eigenvalues. This type of analysis resulted in the claim that the emotion space can be described by essentially two axes (eigenvectors), namely, the degree of pleasure and the degree of arousal, and provided the main evidence in support of the Circumplex model of emotion[50]. Although this is clearly not the case here, since our critical dimension m^* is definitely greater than two, it is instructive to look more closely to the first three eigenvectors of our similarity matrix S, shown in Fig. 5, and seek an emotional interpretation for them. We note that whereas for the main coordinate (rescaled eigenvector) x^1 all components are positive (upper panel of Fig. 5), all other 129 coordinates fluctuate between positive and negative values. The same behavior pattern was found in the spectral analysis of the null model similarity matrix $\boldsymbol{R}$.

To interpret the principal coordinate x^1, which is the first eigenvector rescaled by the square root of its corresponding eigenvalue[see eq. (7)], first we note that its eigenvalue λ^1 is about ten times greater than λ^2 (see caption of Fig. 4). We attribute such a large contribution, as well as the fact that all the entries of x^1 are positive, to the noisy portion of the similarity matrix ***S***. In fact, the average of the entries of x^1, 0.617, (shown as the horizontal line in the upper panel of Fig. 5), is close to the mean of the entries of the random matrix R. This is expected since a great part of the similarity between any two target emotion names will be due to the coincidence between emotion unrelated words such as many classes of verbs, for instance.

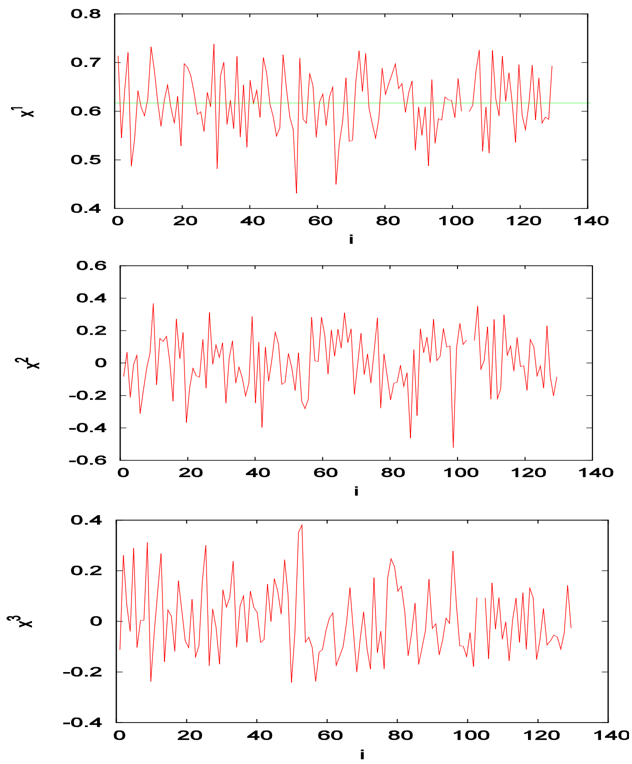

Figure 5. The coordinate vectors x^a associated to the three largest eigenvalues of the similarity matrix S. The labels $i = 1,\ldots 130$ stand for the emotion words listed in Table 1. The horizontal line in the upper panel indicates the mean value 0.617 of the entries of x^1

Next, to interpret the other principal coordinates, x^2 and x^3, shown in the middle and lower panels of Fig. 5, respectively, we consider only the components (words) with the largest positive and negative weights and then imagine these words aligned in a line (axis). For example, if a coordinate assigns a large negative entry to *frustration* and a large positive entry to *gladness*, then this coordinate could be said to measure the degree of pleasure. On the other hand, if another coordinate assigns a large negative entry to *dismay* but a positive entry to *arousal*, then we might say that coordinate measures the degree of arousal. Of course, the roles of positive and negative entries can be interchanged without affecting the interpretation of the axes. In Table 2 we present the 8 largest entries (in absolute value) of the four principal coordinates. It is difficult to find a clear emotional interpretation of these axes (actually, such an interpretation is not always possible) but overall we can see that, except for x^1, which is interpreted as random noise, there is a psychologically meaningful contraposition between negative and positive emotion words. This is clear for the principal coordinates x^2, x^3, and x^4, which differ by the intensity of the positive and negative emotions, whose possible interpretations are, correspondingly, ‘anger’, ‘pleasure’, and ‘attractiveness’.

3.3. Hierarchical Clustering

Given the similarity matrix ***S***, it is natural to attempt to group or categorize the emotion names in clusters or families. In fact, this seems to be the only unbiased manner of defining (and characterizing) a few ‘basic’ emotions amidst the hundreds of emotions described by people. The outcome of such an analysis, carried out by Shaver et al[54] when the similarities between emotion words were rated by humans, is that there are six basic-emotion categories, namely, *love*, *joy*, *surprise*, *anger*, *sadness*, and *fear*.

The variance or spread of a set of points (i.e., the sum of the squared distances from the centre) is the key element of many clustering algorithms[61]. In Ward’s minimum variance method[56] we agglomerate two distinct clusters into a single cluster such that the within-class variance of the partition thereby obtained is minimal. Hence the method proceeds from an initial partition where all objects (130 emotion names, in our case) are isolated clusters and then begin merging the clusters so as to minimize the variance criterion. Tables 3 summarizes the results of the hierarchical clustering algorithm when the objects (i.e., the target emotion names) are partitioned into 25 categories, which is the highest level of hierarchy described in Shaver et al[54].

Table 2. The four smallest and the four largest entries, shown in the parentheses, of the first four principal coordinates

x^1	x^2	x^3	x^4
glee (0.431)	rapture (-0.521)	frustration (-0.242)	infatuation (-0.438)
horror (0.449)	melancholy (-0.462)	anger (-0.239)	fondness (-0.295)
disgust (0.481)	euphoria (-0.397)	grief (-0.238)	attraction (-0.267)
agitation (0.487)	bliss (-0.366)	hurt (-0.201)	love (-0.242)
shock (0.725)	hostility (0.310)	delight (0.300)	apprehension (0.287)
sadness (0.726)	depression (0.316)	amusement (0.311)	alarm (0.290)
anguish (0.732)	resentment (0.352)	gaiety (0.351)	dread (0.335)
disappointment (0.737)	anger (0.366)	gladness (0.379)	fright (0.385)

Table 3. The partition of the 130 target emotion names into 25 clusters according to Ward's minimum variance hierarchical clustering algorithm for the case the contexts surrounding the target emotion words are formed by 40 valid words

1	acceptance, courage, elation, joy, sadness, sorrow, wonder
2	agitation
3	adoration, affection, fondness, infatuation, liking, love, lust
4	alarm, apprehension, dismay, dread, fear, fright, tenseness, uneasiness
5	aggravation, annoyance, exasperation, fury, grouchiness, grumpiness, irritation, rage, wrath
6	amusement, delight, gaiety, gladness, jolliness, joviality
7	arousal, bliss, distress, ecstasy, euphoria, hysteria, melancholy, mortification, rapture, shock, triumph
8	anger, attraction, grief, hurt, resentment, worry
9	bitterness, guilt, pity, pride, regret, remorse
10	aversion, dislike, hate
11	desire, eagerness, enthusiasm, excitement, exhilaration, jubilation, passion, zeal, zest
12	anxiety, panic
13	alienation, compassion, defeat, disappointment, frustration, insult, loneliness, pain, relief, sympathy
14	horror, terror
15	agony, anguish, gloom, glumness, misery, suffering, torment, unhappiness, woe
16	glee
17	dejection, depression, despair, hopelessness
18	anticipation, happiness, interest, optimism, satisfaction, surprise
19	embarrassment, shame
20	ferocity, outrage, spite
21	envy, insecurity, jealousy
22	contempt, disgust, loathing, revulsion, scorn, vengefulness
23	hope, humiliation
24	displeasure, hostility, neglect, pleasure
25	homesickness, isolation, rejection

We note that although these classifications are overall reasonable there are a few pairs of antonymous words that are lumped together in the same cluster, e.g., *joy/sadness* (cluster 1) and *pleasure/displeasure* (cluster 24). At this stage, it is not clear whether this finding is the result of an imperfect context filtering scheme or whether it reflects some intrinsic property of the retrieved texts (see Section 4).

Some words about the clusters produced by Ward's algorithm are in order. First, as already pointed out, the initial partition contains 130 singleton clusters. The first agglomeration, which reduced the number of clusters to 129, grouped the words *aggravation* and *irritation*; the second agglomeration grouped the words *anguish* and *gloom*; the third, the words *mortification* and *shock*; the fourth, the words *eagerness* and *enthusiasm*; and the fifth, the words *euphoria* and *rapture*. Not surprisingly, these pairs of words happen to be those with the highest similarity values (see Section 3.1). At these stages of the hierarchy, the good performance of our context comparison method in clustering words of similar meanings, without employing any explicit semantic information, is truly remarkable. The consequences of this finding – a self-organized dictionary - certainly deserve further research.

3.4. The Effect of the Length of the Contexts

Table 4. The partition of the 130 target emotion names into 25 clusters according to Ward's minimum variance hierarchical clustering algorithm for the case the contexts surrounding the target emotion words are formed by 10 valid words

1	acceptance, amusement, anticipation, gladness, happiness, joy, sadness, surprise
2	arousal, courage, distress, melancholy, mortification, rapture, shock, sorrow, wonder
3	adoration, affection, fondness, liking
4	alarm, dread, fear, fright, horror, terror
5	aggravation, agitation, annoyance, grouchiness, grumpiness, irritation
6	delight, gaiety, glee, jolliness, joviality
7	bliss, ecstasy, elation, euphoria, jubilation, optimism, triumph
8	anger, exasperation, frustration, fury, outrage, rage, wrath
9	attraction, aversion, desire, hysteria, interest, passion, satisfaction
10	apprehension, tenseness, uneasiness, worry
11	eagerness, enthusiasm, excitement, exhilaration, zeal, zest
12	anxiety, panic
13	compassion, pity, sympathy
14	agony, pain, relief, suffering, torment
15	disappointment, dismay, misery, unhappiness, woe
16	contempt, disgust
17	anguish, dejection, depression, despair, grief, hopelessness, hurt
18	infatuation, love, lust
19	embarrassment, guilt, pride, regret, remorse, shame
20	gloom, glumness
21	envy, jealousy
22	bitterness, dislike, ferocity, hate, hostility, loathing, resentment, revulsion, scorn, spite, vengefulness
23	defeat, hope, humiliation, insecurity, insult
24	displeasure, pleasure
25	alienation, homesickness, isolation, loneliness, neglect, rejection

Up to now we have considered only the case where the contexts are comprised of 20 valid words before and 20 valid words after the target emotion word, yielding thus contexts of length 40, as described in Section 2. However, we can use the same Web retrieved documents to produce contexts of any length less than 40. For the purpose of comparison, we offer here a summary of the results obtained for a similarity matrix produced by contexts of length 10, i.e. 5 valid words before and 5 valid words after the target emotion name. Regarding the basic statistic measures we find **Sm** = 0.133 and $\mathbf{s}_s$ = 0.047 which shows that the matrix **S** becomes more heterogeneous as we reduce the context size: the ratio $\mathbf{s}_s$/**Sm** is about two times higher than in the case of contexts of length 40. As for the rank order statistics we find that the rank-order distribution is qualitatively similar to those exhibited in Fig. 1. For this shorter context size, the pair of distinct words with the largest similarity, S_1 = 0.493, is *jolliness* and *joviality*, and the pair with the smallest similarity, S_{8385} = 0.030, is *displeasure* and *glee*. In addition, the elbow test (see Fig. 3) applied to this case yields the same prediction for the emotion space dimension, i.e., m* » 30 or m* » 5, depending on the region we choose to locate the elbow.

The hierarchical clustering analysis, however, exhibited some remarkable sensitivity to the length of the context. In fact, in Table 4 we show the partition of the 130 emotion names into 25 categories using the similarity matrix resulting from the comparison of contexts of length 10 (see Table 3 for the same partition using contexts of size 40). There are some subtle differences between the categories resulting from these different context lengths and we feel Table 4 provides a more intuitive clustering. The shortening of the context length resulted in the decreasing of the noise contribution to the similarity measure as indicated by the two-fold increase of the heterogeneity of the similarity entries mentioned above. (We recall that the random null model is characterized by a very homogeneous similarity matrix as shown in Fig. 1.) In any event, these two classifications illustrate the sensitivity of the resulting clustering to changes in the contexts of the emotion names.

4. Discussion

The classification of the 130 emotion words into 25 groups, as summarized in Tables 3 and 4, is the main result of the preceding analysis, and permits a direct comparison of our approach with the traditional querying of participants used by Shaver et al[54]. The procedure employed by those authors to produce their clusters of emotion names was to ask a group of 100 students to sort n cards, each one containing the name of an emotion (we recall that $n = 135$ in that study), into categories representing their best judgment about which emotions are similar and which are different from each other[54]. In addition it was explicitly pointed out to the students that *there was no correct way to sort the cards*. As a result, category size ranged from 1 to 90 elements; one participant classified all names into two categories according to the positive or negative connotation of the emotion name. Then for each pair of words, say i and j, an integer number $b_{ij} = 0,\dots 100$, is defined corresponding to the number of students that placed words i and j in the same category. This $n \times n$ co-occurrence matrix was then analyzed using a standard clustering algorithm[54].

Given the two very different procedures applied to produce the classification of the emotion words, it is most reassuring that the classifications results are similar, provided one overlooks some of the obvious 'misclassifications' of our procedure based on web retrieved texts, such as the *displeasure/pleasure, attraction/aversion* and the *joy/sadness* associations. In fact, considering Table 4, there are four clusters (namely, clusters 1, 2, 9 and 24) that exhibit this type of misclassification; the other 21 clusters offer a surprisingly sensible classification which bears a strong correlation with the classification presented by Shaver et al[54]. We warn, however, that there is no 'correct' classification of the emotion words.

Inspection of the bags of words associated to *pleasure* and *displeasure* for contexts of length 10 (Table 4) reveals the reason for their placement in the same isolated category: the word *displeasure* appears 3 times in the bag of the word *pleasure*, but the word *pleasure* appears 22 times in the bag of the word *displeasure*. As a result, the unprocessed similarity between these two target emotion names acquires a large numerical value ($3 \times 22 = 66$). It is interesting that people frequently use the word pleasure to characterize and talk about displeasure, but the reverse is not true. Although we could easily eliminate this type of misclassification we choose not do so at this stage, because it may be a genuine characteristic of the written language. In addition, we must bear in mind that some words (e.g., colorless, infinity, insanity, freedom, etc.) have meanings definable only with reference to their opposites; this effect may underlie the explanation for the observed high similarity among some antagonistic words.

5. Conclusions

In this contribution we compare the structure of the emotional space estimated from word contexts and from subjective judgments. We find that both are similar.

We present the first step towards the ambitious goal of exploring the vast amount of texts readily available in the Web to obtain information about human emotionality and psychology. Our paper addresses the categories of English emotion names – an extensively investigated research topic in social psychology[46,54,57,58]. Future research should extend our results to other cultures[64,65]. As noted by Plutchik[16], the appearance of words like angry, afraid, and happy in all languages suggests that they represent universal experiences. Emotional words categorize a part of personal and social reality[64] and these categories are important constituents of people's psychology. This paper demonstrated that a significant part of emotional experiences can be expressed in words.

This paper opens more questions than gives answers. Emotions named words could be possibly a minor part of human emotion abilities. Aesthetic emotions related to knowledge[66,67,68,69,70], musical emotions[35,42,71,72, 73] and emotions of cognitive dissonances[39,42], for instance, are not described by emotion words. Could they be studied by using the context comparison method similar to this paper? For example, could cognitive dissonance emotions be measured by substituting emotional words in this study with choices? Provided we can measure the perceived similarities of the emotions evoked by the musical stimulus or by the tension of holding conflicting thoughts, the mathematical methods used here can be applied to characterize those types of emotion as well.

Our method resulted in all reasonable categories (see Table 4), highly correlated with the categories obtained from the subjective judgment[54]. Our estimate of the dimension of the subjacent emotion space m^* is consistent with those obtained from people's subjective judgment.

An advantage of our method includes a possibility of investigating cultural evolution of emotions and their perceptions. Studying contexts of emotion words is possibly the only way to understand emotions existing centuries and millennia ago. For example, by studying usage contexts, Konstan[74] suggested that even such a basic idea as ‘forgiveness’ in its contemporary meaning appeared only two or three centuries ago, and did not exist in antiquity, or in the Church Fathers, or in the Bible. Homer's characters in the Iliad and the Odyssey had no concept of ‘guilt’ either[75]. Another advantage of our approach is the easiness of investigating how languages and cultures differ in emotionality. Experimental studies demonstrated different emotional content in different languages[76,77], and[34] suggested that the grammar affects the emotionality of a language. We refer the reader to Russell[64] for a lucid review of ethnographic and cross-cultural studies of emotion lexicons. The method developed here can be easily applied to different languages. Additional topics of investigation are comparisons between categories of emotion words in prose and poetry, as well as among different writers. Finally, we would suggest that other aspects of cognition could be explored using similar methods, it could be a fascinating subject for future research.

ACKNOWLEDGEMENTS

The research of JFF at São Carlos was supported by The Southern Office of Aerospace Research and Development (SOARD), Grant No. FA9550-10-1-0006, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). The research of LIP was supported by the AFOSR.

REFERENCES

[1] Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. Behavioral and Brain Sciences, 31, 559-575.

[2] Grossberg, S., & Levine, D.S. (1987). Neural dynamics of attentionally modulated Pavlovian conditioning: Blocking, inter-stimulus interval, and secondary reinforcement. Applied Optics, 26, 5015-5030.

[3] Simon, H. A. (1955). On a class of skew distribution functions. Biometrika, 42, 425–440.

[4] Plutchik, R. (1962). The Emotions: Facts, Theories and a New Model. New York: Random House.

[5] Ekman, P. (1957). A methodological discussion of nonverbal behavior. Journal of Psychology, 43, 141-149.

[6] Ortony, A., Clore, G. L. & Foss, M. A. (1987). The referential structure of the affective lexicon. Cognitive Science, 11, 341-364.

[7] Perlovsky, L.I. (2012). Brain: conscious and unconscious mechanisms of cognition, emotions, and language. Brain Sciences, Special Issue "The Brain Knows More than It Admits" - in press

[8] Gratch, J., Marsella, S., & Petta, P. (2009). Modeling the cognitive antecedents and consequences of emotion. Cognitive Systems Research, 10, 1-5.

[9] Reisenzein, R. (2009). Emotions as metarepresentational states of mind: Naturalizing the belief-desire theory of emotion. Cognitive Systems Research, 10, 6–20.

[10] Cabanac, M. (2002). What is emotion? Behavioural Processes, 60, 69-83.

[11] Thagard, P. & Stewart, T. C. (2011). The AHA! Experience: Creativity through emergent binding in neural networks. Cognitive Science, 35(1), 1–33

[12] Perlovsky, L.I. & Levine, D. (2012). The Drive for Creativity and the Escape from Creativity: Neurocognitive Mechanisms. Cognitive Computation, DOI 10.1007/s12559-012-9154-3; in press.

[13] Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition, 50, 7-15.

[14] Damasio, H., Grabowski, T., Frank, R., Galaburda, A.M., & Damasio, A. R. (1994). The Return of Phineas Gage: Clues about the Brain from the Skull of a Famous Patient. Science, 264, 1102-1105.

[15] Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.) (2001). Appraisal processes in emotion: Theory, methods, research. New York: Oxford University Press

[16] Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.), Emotion: Theory, research, and experience (Vol. 1: Theories of emotion, pp. 3-33). New York: Academic Press.

[17] Tooby, J., & Cosmides, L. (1990). The past explains the present: emotional adaptations and the structure of ancestral environment. Ethology and Sociobiology, 11, 375-424.

[18] Lazarus, R. S. (1991). Emotion and Adaptation. New York: Oxford University Press.

[19] Johnson-Laird, P. N., & Oatley, K. (1992). Basic emotions: a cognitive science approach to function, folk theory and empirical study. Cognition and Emotion, 6, 201-223.

[20] Ekman, P. (1999). Basic Emotions. In T. Dalgleish and M. Power (Eds.), Handbook of Cognition and Emotion. Sussex, U.K.: John Wiley & Sons.

[21] Ortony, A., & Turner, T. J. (1990). What's Basic About Basic Emotions? Psychological Review, 97, 315-331.

[22] Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. Psychological Review, 99, 561-565.

[23] Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., & Barrett, L.F. (2011). The brain basis of emotion: A meta-analytic review. Brain and Behavior Sciences, in print.

[24] Tikhanoff, V., Fontanari, J. F., Cangelosi, A. & Perlovsky, L. I. (2006). Language and cognition integration through modeling field theory: category formation for symbol grounding. In Book Series in Computer Science, v. 4131, Heidelberg: Springer.

[25] Fontanari, J.F. and Perlovsky, L.I. (2008b). A game theoretical approach to the evolution of structured communication codes, Theory in Biosciences, 127, pp.205-214; e-version http://dx.doi.org/10.1007/s12064-008-0024-1.

[26] Chater, N. & Christiansen, M. H. (2010). Language acquisition meets language evolution. Cognitive Science, 34(7), 1131–1157.

[27] Perlovsky, L.I., Deming R.W., & Ilin, R. (2011). Emotional Cognitive Neural Algorithms with Engineering Applications. Dynamic Logic: from vague to crisp. Springer, Heidelberg, Germany.

[28] Perlovsky, L.I. (2012). Mirror Neurons, Language, and Embodied Cognition. Neural Networks, in press.

[29] Perlovsky, L.I. (2011). Language, Emotions, and Cultures: Emotional Sapir-Whorf Hypothesis. WebmedCentral PSYCHOLOGY 2011;2(2):WMC001580.

[30] Perlovsky L.I. (2011). Language and Cognition Interaction Neural Mechanisms, Computational Intelligence and Neuroscience, 2011, Article ID 454587. Open Journal, doi:10.1155/ 2011/454587,http://www.hindawi.com/journals/cin/contents/

[31] Perlovsky, L.I. (2009). Language and Cognition. Neural Networks, 22(3), 247-257.

[32] Perlovsky, L.I. (2012). Emotionality of Languages Affects Evolution of Cultures. Review of Psychology Frontier, in press.

[33] Perlovsky, L.I. (2006). Music – The First Priciple. Musical Theatre, http://www.ceo.spb.ru/libretto/kon_lan/ogl.shtml

[34] Perlovsky, L.I. (2009). Language and Emotions: Emotional Sapir-Whorf Hypothesis. Neural Networks, 22, 518-526.

[35] Perlovsky, L.I. (2010a). Musical emotions: Functions, origin, evolution. Physics of Life Reviews, 7, 2-27.

[36] Perlovsky, L.I. (2011). Emotions of “higher” cognition, Comment to Lindquist at al ‘The brain basis of emotion: A meta-analytic review.’ Brain and Behavior Sciences, in print.

[37] Perlovsky, L.I. (2012). Cognitive Function of Music, Part I. Interdisciplinary Science Reviews, 37(2), 129–42.

[38] Perlovsky, L.I. (2012). The Cognitive Function of Emotions of Spiritually Sublime. Review of Psychology Frontier, 1(1), 1-10; www.j-psy.org.

[39] Fontanari, J. F., Bonniot-Cabanac, M.-C., Cabanac, M., & Perlovsky, L.I. (2012). A structural model of emotions of cognitive dissonances, preprint arXiv1202.6388, Neural Networks, in press.

[40] Perlovsky L.I. (2010). Physics of The Mind: Concepts, Emotions, Language, Cognition, Consciousness, Beauty, Music, and Symbolic Culture. WebmedCentral PSYCHOLOGY 2010;1(12):WMC001374

[41] Perlovsky, L.I. (2011). Music. Cognitive Function, Origin, and Evolution of Musical Emotions. WebmedCentral PSYCHOLOGY 2011;2(2):WMC001494.

[42] Perlovsky, L.I., Cabanac, A., Bonniot-Cabanac, M.-C., & Cabanac, M. (2012). ‘Mozart effect’, cognitive dissonance, and origin of music, in press.

[43] Perlovsky, L.I., Bonniot-Cabanac, M.-C., & Cabanac, M. (2010). Curiosity and Pleasure. WebmedCentral PSYCHOLOGY 2010;1(12):WMC001275

[44] Masataka, N. & Perlovsky, L.I. (2012). Music can reduce cognitive dissonance. Nature Precedings: hdl:10101/npre.2012.7080.1.

[45] Perlovsky, L.I. & Ilin, R. (2012). Mathematical Model of Grounded Symbols: Perceptual Symbol System. Journal of Behavioral and Brain Science, 2, 195-220; doi:10.4236/ jbbs.2012.22024; http://www.SciRP.org/journal/jbbs.

[46] Russell, J. A. (1989). Measures of emotion. In R. Plutchik & H. Kellerman (Eds.), Emotion: Theory, research, and experience (Vol. 4: The measurement of emotions, pp. 83-111). New York: Academic Press.

[47] Ekman, G. (1954). Dimensions of color vision. Journal of Psychology, 38, 467-474.

[48] Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. Part I. Psychometrika, 27, 125-140.

[49] Perlovsky, L.I. & Ilin, R. (2010). Neurally and Mathematically Motivated Architecture for Language and Thought. Special Issue "Brain and Language Architectures: Where We are Now?" The Open Neuroimaging Journal, 4, 70-80. http://www.bentham.org/open/tonij/openaccess2.htm

[50] Russell, J. A. (1980). A Circumplex Model of Affect. Journal of Personality and Social Psychology, 39, 1161-1178.

[51] Russell, J. A., & Feldman Barrett, L. (1999). Core Affect, Prototypical Emotional Episodes, and Other Things Called Emotion: Dissecting the Elephant. Journal of Personality and Social Psychology, 76, 805-819.

[52] Watson, D. & Tellegen, A. (1985). Toward a consensual structure of mood. Psychological Bulletin, 98, 219-235.

[53] Wierzbicka, A. (1992). Defining emotion concepts. Cognitive Science, 16(4), 539–581.

[54] Shaver, P., Schwartz, J., Kirson, D., & O’Connor, C. (1987). Emotion Knowledge: Further Exploration of a Prototype Approach. Journal of Personality and Social Psychology, 52, 1061-1086.

[55] Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. Psychometrika, 17, 401-419.

[56] Ward, J.H., Jr. (1963). Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association, 48, 236-244.

[57] Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. Journal of Personality and Social Psychology, 48, 813-838.

[58] Schimmack, U., & Reisenzein, R. (1997). Cognitive processes involved in similarity judgments of emotions. Journal of Personality and Social Psychology, 73, 645-661.

[59] Averill, J. R. (1975). A semantic atlas of emotional concepts. JSAS: Catalog of Selected Documents in Psychology, 5, 330.

[60] Shepard, R. N. (1980). Multidimensional Scaling, Tree-Fitting, and Clustering. Science, 210, 390-398.

[61] Murtagh, F., & Heck, A. (1997). Multivariate Data Analysis. Dordrecht: Kluwer.

[62] Borg, I., & Groenen, P. J. F. (2005). Modern Multidimensional Scaling: Theory and Applications. New York: Springer.

[63] Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. Psychometrika, 29, 28-42.

[64] Russell, J. A. (1991). Culture and the Categorization of Emotions. Psychological Bulletin, 110, 426-450.

[65] Perlovsky, L.I. (2007). Evolution of Languages, Consciousness, and Cultures. IEEE Computational Intelligence Magazine, 2(3), 25-39.

[66] Perlovsky, L.I. (2001). Neural Networks and Intellect: using model-based concepts. Oxford University Press, New York, NY.

[67] Perlovsky, L.I. (2007). Neural Dynamic Logic of Consciousness: the Knowledge Instinct. In Neurodynamics of Higher-Level Cognition and Consciousness, Eds. Perlovsky, L.I., Kozma, R. Springer Verlag, Heidelberg, Germany.

[68] Perlovsky, L.I. (2010). Neural Mechanisms of the Mind, Aristotle, Zadeh, & fMRI, IEEE Trans. Neural Networks, 21(5), 718-33.

[69] Perlovsky, L.I., Kozma, R., Eds. (2007). Neurodynamics of Higher-Level Cognition and Consciousness. Springer-Verlag, Heidelberg, Germany.

[70] Mayorga, R., Perlovsky, L.I., Eds. (2008). Sapient Systems. Springer, London, UK.

[71] Perlovsky, L.I. (2008). Music and Consciousness, Leonardo, Journal of Arts, Sciences and Technology, 41(4), 420-421.

[72] Perlovsky, L.I. (2012). Cognitive function of musical emotions. Psychomusicology, in press.

[73] Perlovsky, L.I. (2012). Cognitive function, origin, and evolution of musical emotions. Musicae Scientiae, e-version: doi: 10.1177/1029864912448327, in press.

[74] Konstan, D. (2010). Before Forgiveness: The Origin of a Moral Idea. New York, NY: Cambridge University Press.

[75] Dodds, E. R. (1951). The Greeks and the irrational. Berkeley: University of California Press.

[76] Guttfreund, D. G. (1990). Effects of language usage on the emotional experience of Spanish-English and English-Spanish bilinguals. Journal of Consulting and Clinical Psychology, 58, 604-607.

[77] Harris, C. L., Ayçiçegi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. Applied Psycholinguistics, 24, 561-579.

[78] Perlovsky, L.I. & Ilin, R. (2012). Mathematical Model of Grounded Symbols: Perceptual Symbol System. Journal of Behavioral and Brain Science, 2, 195-220; http://www.SciRP.org/journal/jbbs.

[79] Levitin, L., Schapiro, B., & Perlovsky, L.I. (1996). Zipf's Law Revisited: Evolutionary Model of Emergent Multi-resolution Classification. Proceedings of the Conference on Intelligent Systems and Semiotics (pp. 65-70). Gaithersburg, MD: National Institute of Standards and Technology.

[80] Ortony, A., Clore, G. L. & Foss, M. A. (1987). The referential structure of the affective lexicon. Cognitive Science, 11, 341-364.

[81] Perlovsky, L.I. (2010b). Intersections of Mathematical, Cognitive, and Aesthetic Theories of Mind. Psychology of Aesthetics, Creativity, and the Arts, 4, 11-17.

[82] Simon, H. A. (1967). Motivational and emotional controls of cognition. Psychological Review, 74, 29-39.