Measuring Cultural Relativity of Emotional Valence and Arousal using Semantic Clustering and Twitter

Eugene Y. Bann (eugene@aeir.co.uk)
Advanced Emotion Intelligence Research (AEIR)

Joanna J. Bryson (J.J.Bryson@bath.ac.uk)
Department of Computer Science, University of Bath, United Kingdom

Abstract

Researchers since at least Darwin have debated whether and to what extent emotions are universal or culture-dependent. However, previous studies have primarily focused on facial expressions and on a limited set of emotions. Given that emotions have a substantial impact on human lives, evidence for cultural emotional relativity might be derived by applying distributional semantics techniques to a text corpus of self-reported behaviour. Here, we explore this idea by measuring the valence and arousal of the twelve most popular emotion keywords expressed on the micro-blogging site Twitter. We do this in three geographical regions: Europe, Asia and North America. We demonstrate that in our sample, the valence and arousal levels of the same emotion keywords differ significantly with respect to these geographical regions — Europeans are, or at least present themselves as more positive and aroused, North Americans are more negative and Asians appear to be more positive but less aroused when compared to global valence and arousal levels of the same emotion keywords. Our work is the first in kind to programatically map large text corpora to a dimensional model of affect.

Keywords: Semantic Clustering, Emotion Analysis; Twitter; Core Affect Model.

Introduction

The question as to whether the experience and expression of emotions is universal or relative to specific cultures has resulted in a wide variety of studies, with theories ranging from the universality hypothesis to culture-specific facial expressions. Here we present evidence that culture is a necessary framework for researchers studying variation in emotions. Independent of the question of biological differences in the experience of emotions, it would be unsurprising if culture shapes our conscious perception, expression and experience of emotions, as has been hypothesised for other cognitive phenomena (Hunt & Agnoli, 1991; Fuhrman et al., 2011). Here, we use Latent Semantic Clustering on an emotional text corpus mined from Twitter to discern how the primary properties normally attributed to emotional keywords — valence and arousal — differ as the keywords are used in the same language (English) as exploited across different global regions.

The Conceptualisation of Emotion Qualia

Emotion qualia refers to the raw feel of an emotion. The actual phenomenon of a particular emotion experienced may differ according to each person's perception or understanding of that emotion, with perception being the result of the individual's past and hypothesised responses, unique to each human being. Barrett (2006) describes the act of conceptualising core affect, or in other words, why people attach emotion labels to the experience of emotion qualia. Since emotion keywords are constructed from conceptual knowledge about the world, emotions themselves may be concepts that humans begin learning in infancy and continuously extend and revise throughout life (Lindquist & Barrett, 2008). This repeated experience of labelling a combination of core affect and the context in which it occurs as an emotion provides training in how to recognise and respond to that emotion. In this sense, Barrett describes emotions as simulations. This skill of conceptualising core affect as an emotion could be a core aspect of emotional intelligence, in much the same way as conceptual thinking is core to cognitive intelligence. Each person learns the label in association with their unique experience, thus each person's conceptualisation of their emotional spectrum is unique. Cultures, formed of communicating individuals, may therefore also be unique if individual experiences vary somehow systematically. We base our analysis on this hypothesis. The reader should bear in mind that we are not analysing emotion keywords in particular, rather, we are analysing emotion *conceptualisations*, or what cultures understand specific emotion keywords to mean, using Latent Semantic Clustering to infer these meanings.

Core Affect

Core affect is an emerging paradigm in affective neuroscience, and postulates a continuous approach to defining emotions (Posner et al., 2005). Several core-affect, or *circumplex* models have been proposed (e.g. Watson & Tellegen, 1985; Russell, 1980; Cacioppo & Berntson, 1994), yet all have one thing in common: they represent emotions as a single point in a continuous space defined by two (or rarely three) dimensions. Different labels have been assigned to these two dominant dimensions by various theorists, such as *pleasure* and *engagement*, however most commonly, *valence* and *arousal* are chosen. Thus far, there has been no attempt to computationally pinpoint emotions or documents within a core affect model using 'online' and 'big' data; to date, research regarding the core affect model has either been theoretical (e.g. Watson & Tellegen, 1985), or conducted via a limited survey (e.g. Russell, 1980).

Core affect is one of two main theories regarding the representation of emotions, the other being the Basic Emotion model, however, neither has thus far received unequivocal support. Basic emotions could turn out to map to multiple subtypes of coherent emotion networks, but this implies we need to split basic emotion categories into further subtypes to better reflect these emotion networks (Hamann, 2012; Bann,

2012). Here we extend this view and suggest that the core affect model enables us to quantify the properties of the basic emotions themselves.

Previous Work

There is growing evidence that aspects of a person's psychology can be predicted from their language usage. In the 1990s, human semantics was shown to be recoverable from linguistic corpra independent of any further grounding (Lowe, 1997; Bryson, 2008). Recent applications to individual psychology include discovering individual differences in personality (Pennebaker & King, 1999), discovering cultural change in moral beliefs (Bilovich & Bryson, 2008), as well as for emotion categorization (Fitzpatrick & Logan, 2011). French discovered that co-occurrence techniques such as LSA does not detect personality from short text samples (French & Gill, 2007), but do reveal that texts expressing particular emotions have a greater semantic similarity to corresponding exemplar words (Gill et al., 2008).

A recent study by Jack et al. (2012) found significant evidence that facial expressions are indeed culture-dependent; that is, different cultures represent the same emotions differently. However, whether or not this is because they *experience* different emotion qualia is another question. Using language, rather than facial expressions, as an accessor to emotion will enable a much more detailed and less ambiguous analysis, increasing significance by "throwing more data at the problem" (Recchia & Jones, 2009, p.3).

Currently, there have been few attempts to analyse cultural differences using language semantics. Language plays a key role in how emotions are conceptualised (and thus perceived); Lindquist states "language can be no more removed from emotion, than flour can be removed from an already baked cake" (Lindquist, 2009, p.1). Recently, Bann & Bryson (2012) demonstrated how conceptualisations of emotions can be inferred by performing Latent Semantic Analysis on a corpus of self-reported emotional tweets. Their DELSAR algorithm analysed 21,000 tweets each labelled with an emotion, and clustered each document in the corpus to its most similar corresponding emotion label using Latent Semantic Clustering. Here we use the same algorithm as the basis for our analysis.

Corpus

Typing emotion keywords into the Internet is increasingly becoming a significant technique for individual expression. There now exists a rich available source of information about emotions on the Internet, because so many people spend time expressing how they feel in blogs, forums, social networking websites and the like. We use data from the microblogging website Twitter to perform large-scale analysis of the language used in thousands of expressions of emotions within tweets. Acquiring a significantly larger corpus than Bann & Bryson (2012), we use the Gardenhose level of Twitter's streaming API[1] to create a corpus of 5,625,844 tweets[2] collected between 19th October 2012 and 18th January 2013. Each emotion keyword (see selection criteria below) is given a five-minute streaming window in turn for the duration of the period, ensuring an even temporal distribution of Tweets is collected. Table 1 describes our corpus, split by 'cultural' region. We use the tweet's timezone as an indication of the corresponding user's geographical location; seeing as it is very unlikely that a Twitter user would select a timezone other than that which they reside in, it is somewhat safe to assume that this reflects the cultural origin of each user.

Table 1: Distribution of tweets within our corpus.

Emotion	Asia	Europe	NA	All
Angry	12194	27070	61293	200024
Ashamed	1008	5097	17107	46486
Calm	5975	10181	36681	102827
Depressed	3078	11615	43129	120473
Excited	30923	100792	292822	847679
Happy	149129	186709	730839	2201874
Interested	3527	9728	31891	86763
Sad	46351	83075	341912	966165
Scared	15435	42500	194130	517715
Sleepy	26031	10787	120473	290666
Stressed	2587	8774	41716	109295
Surprised	3032	12454	56332	135877
Total	**299270**	**508782**	**1968325**	**5625844**

Region definitions. We only include those timezones that have over 5000 tweets within our corpus. The Asia region consists of the timezones *Kuala Lumpur, Beijing, Singapore, Jakarta, Bangkok, Hong Kong, Tokyo*; the Europe region consists of the timezones *London, Amsterdam, Athens, Edinburgh, Dublin, Berlin, Paris*; the North American (NA) region consists of the timezones *Eastern Time (US & Canada), Central Time (US & Canada), Mountain Time (US & Canada), Pacific Time (US & Canada)*.

Selection of emotions. As opposed to strictly using the basic emotions as identified by Bann & Bryson (2012), we use the most popular emotions that are used on Twitter, that is, those emotions that have the highest stream rate. Twelve emotions were selected that had a high rate and that equally divided into positive/negative and engaged/disengaged theoretical categories (see Table 2).

Subcorpus creation. Each subcorpus is created using a limit of 1000 documents per emotion for all subcorpora to ensure consistency within our results; we chose 1000 as it is the lowest value in Table 1. To mitigate micro-temporal effects, if the number of documents for a particular emotion is significantly greater than 1000, we use a modulus function to extract 1000

---

[1] https://dev.twitter.com/docs/streaming-apis.

[2] Having first removed 34,725 duplicate tweets. Corpus and code is available to download at www.aeir.co.uk/code.

documents equally spaced across the subcorpus — for example, if a particular emotion in a particular subcorpus has 6000 documents, we take one document every six documents. We also create six control subcorpora so to compare our region-specific results with a baseline. We use the same modulus function to extract 1000 equally spaced tweets, but without any timezone clause, selecting six random starting points.

Proposed Method

We use DELSAR (Bann & Bryson, 2012, see Algorithm 1) to generate the clustering matrix for each subcorpus — the three regions Asia, Europe and NA, and six random controls.

Algorithm 1 DELSAR

Require: Corpus **C** and Keyword Set **K**, where each document in **C** is mapped to one emotion keyword, *emotion*, in **K** (through corpus generation)

Generate cosine document similarity matrix of LSC(**C**, **K**) (*document* × *document* similarity matrix)

for each *emotion* ∈ **K** **do**

    for each *document* that has emotion *emotion* **do**

        delete *emotion* within the *document*

        Find the closest document *nearest* where *nearest* ≠ *document*

        Increment the count for the emotion that *nearest* is labelled as in *emotion_vector*

    end for each

    return *emotion_vector*

end for each

For each subcorpus, DELSAR uses LSA (Landauer & Dumais, 1997) to create a document-document matrix of cosine similarities (*Similarity Matrix*), in which similar documents are closer to one (i.e. the cosine of the angle between their vectors). It creates a clustering matrix that represents the corpus as an emotion-emotion matrix, describing how each emotion is similar to each other emotion.

All analysis was performed on a 64-bit Intel Core i5 CPU 2x2.67GHz with 4GB RAM using the GENSIM framework for Python (Řehůřek & Sojka, 2010) to create LSA spaces. For all tasks, we use a dimension of 36 and use Log-Entropy normalisation as our Association Function, found to generate optimal results (Nakov et al., 2001) and recommended for LSA (Landauer & Dumais, 1997).

Valence and Arousal

Here we take valance to mean the theoretical positive or negative attribution of an emotion keyword, and similarly arousal to mean the implied level of engagement. We should use the keywords *theoretical* valence and *theoretical* arousal as we are measuring emotion keywords relative to their generally accepted categorisation, although there does seem to be consistency in these categorisations between theorists. Table 2 shows the theoretical definitions of our keywords, accumulated using several circumplex models of affect (Watson & Tellegen, 1985; Russell, 1980; Cacioppo & Berntson, 1994).

Table 2: Valence and arousal categorisation of the twelve emotion keywords analysed.

Emotion	Valence	Arousal
Angry	Negative	Engaged
Ashamed	Negative	Disengaged
Calm	Positive	Disengaged
Depressed	Negative	Disengaged
Excited	Positive	Engaged
Happy	Positive	Disengaged
Interested	Positive	Engaged
Sad	Negative	Disengaged
Scared	Negative	Engaged
Sleepy	Positive	Disengaged
Stressed	Negative	Engaged
Surprised	Positive	Engaged

We calculate the valence and arousal levels of each emotion for each subcorpus as follows. First, we run DELSAR on the subcorpus to generate clustering vectors for each emotion. Each emotion's valence is then calculated as the number of *positive* elements within its vector, as defined in Table 2, divided by the total number of documents across all elements (which will always be 1000), or in other words, the percentage of *positive* elements within its vector. Similarly, each emotion's arousal is calculated as the percentage of *engaged* elements within its vector, again as defined in Table 2. We then normalise each valence and arousal value by taking away the average valence and arousal value, respectively, for *all* subcorpora analysed — Asia, Europe and NA regions and the six control subcorpora. This ensures relativity of the resulting circumplex model between these analysed groups; these groups can now be compared to one another to establish similarities and differences between them.

Results

Figure 1 shows a plot of our circumplex of selected subcorpora. We can see that some emotions are more tightly packed than others, and interestingly, that low-valence-high-arousal and high-valence-low-arousal emotions are much more universally similar when compared to the other two quadrants of the circumplex. In order to visualise each separate region more clearly we illustrate the aggregate theoretical positivity and engagement for each subcorpus, shown in Figure 2. This clearly illustrates that our three regions do indeed have different conceptualisations of the same emotion keyword; we see that the region *Europe* is a much more *positive* and *engaged* culture; in other words, Europeans find the same emotion keywords to be more positive and engaging when compared to other cultures and indeed our control samples. Also, we discover that Asians find the same emotion keywords to be somewhat more positive, and North Americans somewhat more negative, with negligible arousal differences.

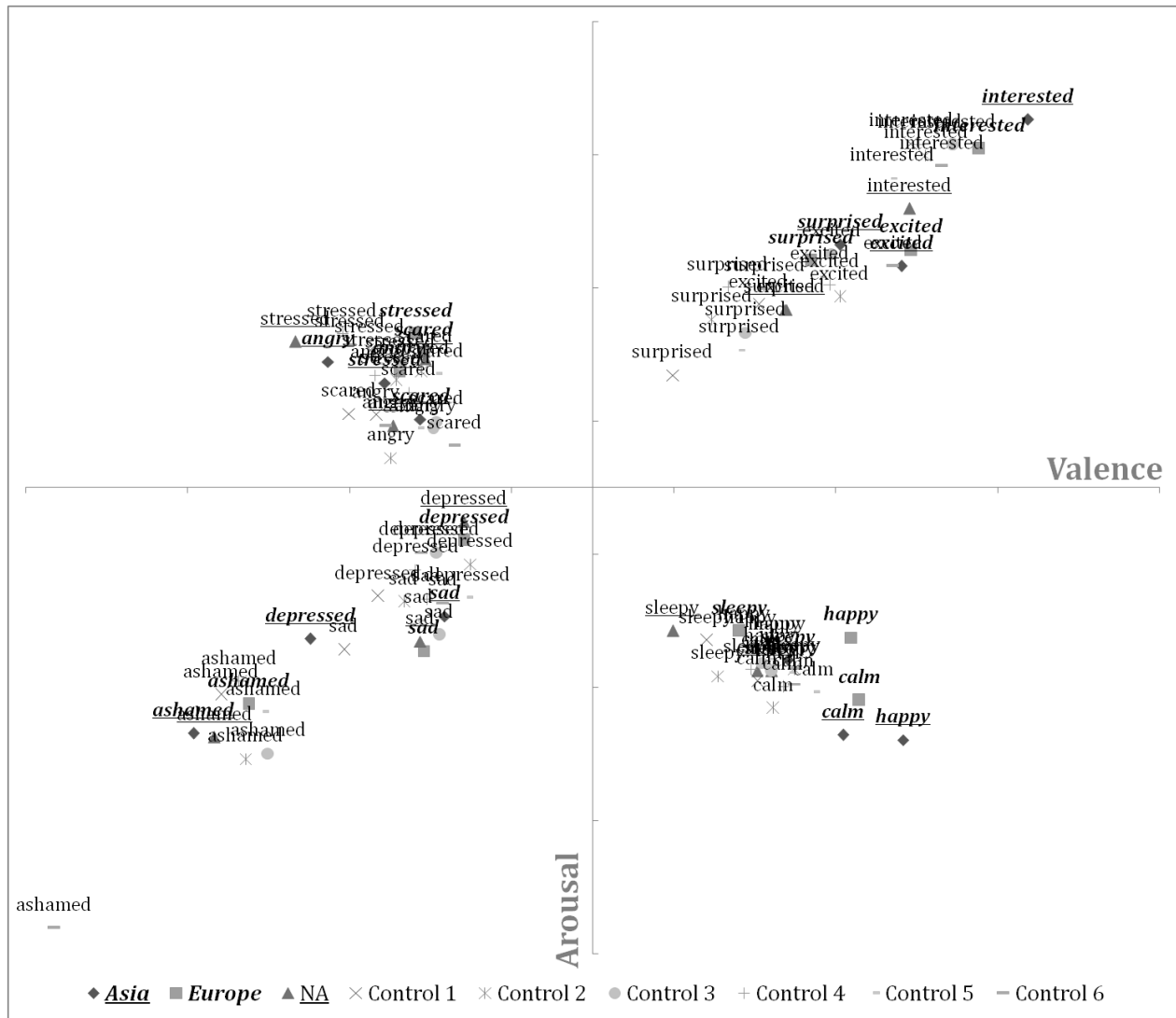In order to analyse how tightly packed our emotion clus-

Figure 1: Circumplex of three regions and six controls.

ters are in Figure 1, we conducted K-Means cluster analysis to determine the centroids for each emotion, calculating the distances of each emotion to its centroid. We plot our centroids, shown in Figure 3, resulting in a circumplex that could be thought of as a universal emotion circumplex, illustrating what people think emotions to be, relative to each other emotion. We can see that the emotions *scared*, *depressed* and *sad* have a very similar valence, yet varying arousal levels; so too do the emotions *sleepy* and *sad*. We can also see, albeit less definitively, that the emotions *stressed* and *surprised* have a similar arousal level, but opposite valence; so too do the emotions *sad* and *sleepy*.

In order to identify which emotions have the most and least similar conceptualisations across cultures, we calculate the distance of each emotion to its respective centroid for each region, and calculate the sum of these distances for each emotion across all subcorpora, shown in Table 3. We discover that

the emotions *sad* and *stressed* have the most similar conceptualisations across all cultures; in other words, people understand these two emotions to mean the same thing independent of culture. Similarly, we find that the emotions *surprised* and *depressed* have the most widely varying conceptualisations across cultures; in other words, different cultures have very different valence and arousal attributions towards these two emotions. Note that we do not include the emotion *ashamed* in the top two due to a strange anomaly in control group 6 which skews an otherwise relatively tight cluster.

Discussion

We would expect that the control groups would be tightly clustered around the centre of the circumplex in Figure 2, and for the most part, they are. The exceptions are control groups one and four, possibly due to the fact the the corpus is skewed in favour of tweets originating from NA (see Table 1);

Figure 2: Aggregate theoretical positivity and engagement for each subcorpus.

Figure 3: Centroid emotion circumplex.

Table 3: Sum of subcorpus distances to respective centroids.

Emotion	Distance	Emotion	Distance
Sad	**17.94**	Scared	23.27
Stressed	**19.66**	Happy	26.10
Calm	20.86	Excited	27.00
Interested	22.72	Depressed	**29.56**
Angry	23.18	Surprised	**32.89**
Sleepy	23.19	Ashamed	*40.70*

this is somewhat verified by their closeness to the NA subcorpus. Other than these anomalous subcorpora, the circumplex does illustrate how different cultures significantly conceptualise emotions differently, in keywords of valence and arousal. Interestingly, there are certain emotions in certain regions that stick out of our analysis. One example is the emotion *depression*; Asians find this emotion much more negative than all other cultures and control groups. This could be due to cultural differences such as coping strategies (Aldwin & Greenberger, 1987). Another example concerns the emotions *happy* and *calm*; Europeans and Asians find these emotions much more positive than North Americans and all control groups. Another suggests that Asians find *interest* a very positive and aroused emotion, compared to North Americans who conceptualise the same emotions, relatively, as negative and disengaged.

Limitations

We document several limitations of our approach. Firstly, our database may still contain duplicate tweets, as some users duplicate tweets by appending, for example, a number at the end, making th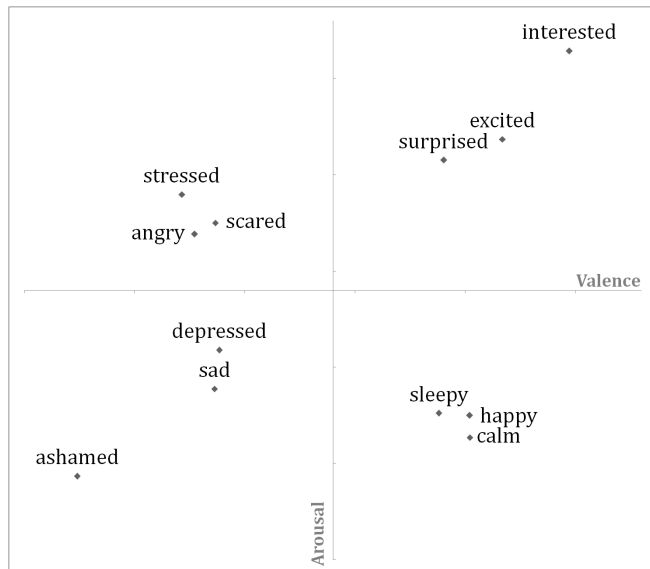em unique from one another. Second, our modulus function does not take an even sample for our control groups at the country level, so they may be skewed in favour of countries with a higher frequency of documents within the database (our corpus on the whole is in fact skewed in favour of NA). Thirdly, we assume that the emotion keywords we have selected are in fact emotion *qualia* as opposed to adjectives. Fourth, our corpus is essentially a snapshot in time and may reflect, for example, the political or economic climate at the time, or skew due to global events such as the US election. Finally, our corpus consists entirely of English tweets, which skews our results in favour of Western cultures; our Asia, and to some extent, Europe subcorpora may not be entirely representative of their respective cultures as we disregard all native languages other than English. In addition, the subpopulations of those regions who choose to use Twitter, and do so in English, may be a biased sample.

Conclusions

Emotions are being increasingly expressed online, and being able to understand these emotions is rapidly becoming a concern of AI and Cognitive Science. By mapping culture-specific emotion circumplexes, we hope to be better able to understand culture-specific perceptions or even experience of emotions. From the work presented here we can conclude the following:

Emotional semantics depends on culture. The same emotion keyword in one culture may describe different valence and arousal properties in another. This seems to be more true of some keywords than others, and could be critical where, for example, a significantly differing conceptualisation of the emotion *depression* would require a different understanding and response.

Emotions vary by geographic region. Europeans are more likely to express positiveness and engagement. Asians are

also more positive than North Americans, both relative to each other and to the control subcorpora. Note that this may reflect cultural differences in the public expression of emotion rather than its actual qualia — our method cannot disambiguate these.

Some emotions do seem to be conceptualised universally. The emotion keywords *sad* and *stressed* have the same conceptualisation across cultures, whereas cultures have the most disagreement regarding the conceptualisation of *surprised*.

We hope that our research paves the way for a better understanding of how language can be used to identify specific properties of emotions, and we encourage the reader to verify our results by downloading our code and corpus at http://www.aeir.co.uk/code.

References

Aldwin, C., & Greenberger, E. (1987). Cultural differences in the predictors of depression. *Am J Community Psychol*, *15*(6), 789-813.

Bann, E. Y. (2012). *Discovering basic emotion sets via semantic clustering on a Twitter corpus* (Vol. CSBU-2013-01; Tech. Rep.). University of Bath, Department of Computer Science.

Bann, E. Y., & Bryson, J. J. (2012, July). The conceptualisation of emotion qualia: Semantic clustering of emotional Tweets. In *Proceedings of the 13th Neural Computation and Psychology Workshop (NCPW13)*.

Barrett, L. F. (2006, February). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*(1), 20-46.

Bilovich, A., & Bryson, J. J. (2008, November). Detecting the evolution of semantics and individual beliefs through statistical analysis of language use. In *Proceedings of the AAAI Fall Symposium on Naturally-Inspired Artificial Intelligence* (pp. 21–26). Arlington, VA: AAAI Press.

Bryson, J. J. (2008, June). Embodiment versus memetics. *Mind & Society*, *7*(1), 77–94.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*(3), 401-423.

Fitzpatrick, J. K., & Logan, J. (2011). Categorizing emotion in spoken language: An analysis of semantic and prosodic contributions to emotional communication. In *Proceedings of the 33rd Annual Cognitive Science Conference.* (pp. 762–767).

Fuhrman, O., McCormick, K., Chen, E., Jiang, H., Shu, D., Mao, S., et al. (2011). How linguistic and cultural forces shape conceptions of time: English and Mandarin time in 3D. *Cognitive Science*, *35*(7), 1305–1328.

Gill, A. J., Gergle, D., French, R. M., & Oberlander, J. (2008). Emotion rating from short blog texts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1121–1124). ACM.

Gill, A. J., & French, R. M. (2007). Level of representation and semantic distance: Rating author personality from texts. In *Proceedings of the 2nd European Cognitive Science Conference.*

Hamann, S. (2012). Mapping discrete and dimensional emotions onto the brain: controversies and consensus. *Trends in Cognitive Sciences*, *16*(9), 458–466.

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, *98*(3), 377.

Jack, R., Garrod, O., Yu, H., Caldara, R., & Schyns, P. (2012, May). Facial expressions of emotion are not culturally universal. In *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 109, pp. 7241–7244).

Landauer, T. K., & Dumais, S. T. (1997). Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction and Representation of Knowledge. *Psychological Review*(104), 211–240.

Lindquist, K. A. (2009, January). Language is powerful. *Emotion Review*(1), 16-18.

Lindquist, K. A., & Barrett, L. F. (2008). Constructing emotion: The experience of fear as a conceptual act. *Psychological Science*, *19*(9), 898-903.

Lowe, W. (1997, August). Meaning and the mental lexicon. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence* (pp. 1092–1097). Nagoya: Morgan Kaufmann.

Nakov, P., Popova, A., & Mateev, P. (2001). Weight functions impact on LSA performance. In *Recent Advances in Natural Language Processing (RANLP 2001)* (p. 187-193).

Pennebaker, J. W., & King, L. A. (1999, December). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*(6), 1296-1312.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734.

Recchia, G., & Jones, M. N. (2009). More data trumps smarter algorithms: Comparing pointwise mutual information with latent semantic analysis. *Behavior Research Methods*, *41*(3), 647.

Řehůřek, R., & Sojka, P. (2010, May 22). Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (pp. 45–50). Valletta, Malta: ELRA.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161-1178.

Watson, D., & Tellegen, A. (1985, September). Toward a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219-235.