Supplementary Materials: Lexica Used

As described in the text, we use eight different sentiment analysis lexica. Each is widely used by scholars interested in the sentiment analysis of various types of texts. Although a number of them are constructed from the same sources, the actual overlap between them is surprisingly small: although the smallest of the lexica contains 3,731 words, only 331 words are captured by all eight lexica with the same polarity (positive or negative).

In addition to containing different sets of words, the lexica also vary in how they assign valence: Four identify words simply as positive or negative, while the others assign words a range of values indicating how strongly positive or negative they are. In addition, two of the lexica specify word stems, indicating they will accept any endings to a word (such 'wildcard' specifications increase the effective size of these lexica considerably). Finally, the ratio of negative to positive words included varies considerably, from 0.40 (labMT, the only lexicon with more positive than negative terms) to 2.39 (HuLiu), with an average ratio of close to 1.5. Table S1 offers a brief overview of the different lexica.

Table S1. Sentiment Analysis Lexica.

Name & citation	**Construction of original lexicon**	**Additional processing here / comments**	**Positive terms**	**Negative terms**
HuLiu (Hu & Liu, 2004)	Manually constructed by scholars at the University of Illinois in Chicago, based on WordNet (Miller, 1995).	Developed for social media; contains terms such as "f*ck".	2003 (+1)	4783 (-1)
labMT (Dodds, Harris, Kloumann, Bliss, & Danforth, 2011)	Used Mechanical Turk coders to code the 'happiness level' of the most frequent 5,000 words from four separate sources: Twitter, Google Books (English), music lyrics (1960 to 2007), and the *New York Times* (1987 to 2007). Full lexicon has 10,222 entries	Filtered out words with low valence scores (absolute value < 1), as recommended by the creators of the lexicon.	2668 (range from 1 to 3.5)	1063 (range from -1 to -3.5)
LexicoderSD (Young & Soroka, 2012)	Manually constructed; starting point was all words from the General Inquirer (GI) (Stone & Hunt, 1963), the Regressive Imagery Dictionary (RID) (Martindale, 1975), and Roget's Thesaurus with the same valence in all 3 dictionaries or with the same valence in 2 and omitted from the third. Targeted at political and	Includes wildcards to specify any ending acceptable for a given stem.	1615 (+1), of which 1043 stems	2768 (-1), of which 1971 stems

	economic news.			
MPQA (Wilson, Wiebe, & Hoffmann, 2005)	Used words from GI, from Hatzivassiloglou and McKeown (1997), and from their own prior work (Riloff & Wiebe, 2003)	Used only single words (no multi-word phrases) Averaged valence for words with multiple entries. 'strong' polarity is given a value of 1, 'weak' polarity gets ½.	2299 (range from 0.175 to 1.00)	4150 (range from -0.175 to -1)
NRC (Mohammad & Turney, 2011; Mohammad & Yang, 2011)	Coded all words from Roget's thesaurus that occur at least 120,000 times in Google's n-gram corpus, using 5 different MT coders for each word.		2312 (+1)	3243 (-1)
SentiWordNet (Baccianella, Esuli, & Sebastiani, 2010)	Assigns valences to the synonym sets (synsets) in the online semantic dictionary WordNet. Starting from 'paradigmatically' positive or negative words, propagated valence across the entire dictionary using the network structure implied by synsets sharing words. Full lexicon has 29,436 entries.	For words with multiple valences (e.g. in multiple synsets), averaged the values. Filtered out words with low aggregate valence scores (absolute value < 0.1)	11116 (range from 0.1 to 1)	13106 (range from -0.1 to -1)
SOCAL (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011)	"Sentiment Orientation CALculator", manually constructed from all words in a 400-text corpus of Epinions reviews, movie reviews from the Polarity Dataset (Pang, Lee, & Vaithyanathan, 2002), and GI.		3716 (range from 0.5 to 5.0)	6341 (range from -0.5 to -5.0)
WordStat	Constructed by Provalis (makers of WordStat), by combining word lists from GI, RID, and Pennebaker's Linguistic and Word Count dictionary (LIWC) (Tausczik & Pennebaker, 2010) and	Includes wildcards to specify any ending acceptable for a given stem.	5539 (+1), of which 337 stems	9539 (-1), of which 578 stems

	searching WordStat's internal dictionary for potential synonyms.			

References

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). *SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining.* Paper presented at the LREC, url: http://nmis.isti.cnr.it/sebastiani/Publications/LREC10.pdf.

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS one, 6*(12), 1-25.

Hatzivassiloglou, V., & McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. *Proceedings of the 35th annual conference of the Association for Computational Linguistics*, 174-181, url: http://www.aclweb.org/anthology/P97-1023.

Hu, M., & Liu, B. (2004). *Mining and summarizing customer reviews.* Paper presented at the The ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD-2004), Seattle, WA, url: https://www.cs.uic.edu/~liub/publications/kdd04-revSummary.pdf.

Martindale, C. (1975). *Romantic progression: The psychology of literary history*. Washington, DC: Hemisphere.

Miller, G. A. (1995). WordNet: A Lexical Database for English. *Communications of the ACM, 38*(11), 39-41.

Mohammad, S. M., & Turney, P. D. (2011). Crowdsourcing a word-emotion association lexicon *Computational Intelligence*: Wiley-Blackwell.

Mohammad, S. M., & Yang, T. (2011). *Tracking sentiment in mail: How genders differ on emotional axes.* Paper presented at the Workshop on computational approaches to subjectivity and sentiment analysis, Portland, OR, url: https://arxiv.org/abs/1309.6347.

Pang, B., Lee, L., & Vaithyanathan, S. (2002). *Thumbs up? Sentiment classification using machine learning techniques.* Paper presented at the Conference on Empirical Methods in NLP, Philadelphia, PA, url: https://www.cs.cornell.edu/home/llee/papers/sentiment.pdf.

Riloff, E., & Wiebe, J. (2003). Learning extraction patterns for subjective expressions. *Proceedings of the 2003 conference on Empirical Methods in Natural Language Processing*, url: https://www.cs.utah.edu/~riloff/pdfs/emnlp03.pdf.

Stone, P. J., & Hunt, E. B. (1963). A computer approach to content analysis: Studies using the General Inquirer. *Proceedings of the Spring Joint Computer Conference*, 241-256.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics, 37*(2), 267-307.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24-54.

Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. *Proceedings of the Human Language Technologies Conference / Conference on Empirical Methods in Natural Language Processing*, url: https://people.cs.pitt.edu/~wiebe/pubs/papers/emnlp05polarity.pdf.

Young, L., & Soroka, S. (2012). Affective news: The automated coding of sentiment in political texts. *Political Communication, 29*(2), 205-231.