Political Analysis of Social Media Data

Sentiment Analysis

Instructor: Gregory Eady
Office: 18.2.10
Office hours: Fridays 13-15

Today

- Sentiment analysis
- Video lectures & exercises

How can we measure the sentiment of social media (or other text) data?

- The *tone* of content may matter as much as the substance
- Affect is linked to decision-making and judgment
- Tone changes how we process news media
- Various cases where sentiment is useful to measure on social media (e.g. positive or negative attitudes toward immigrants; tone of news coverage; tone of political campaigns)

The key and obvious problem for large-scale sentiment analysis

- Time
  - Manually coding millions of tweets is effectively impossible
- Money
  - Finding funding to have annotators code thousands of tweets is challenging

Automated sentiment analysis

- Allows us to analyze massive amounts of data easily
- Are two broad possible approaches:
    1. Statistical
    2. Non-statistical

1. Statistical (data-driven) approaches

- **Supervised learning**
  - Some texts in our data (e.g. social media posts) have labels indicating whether (or the degree to which) they express positive or negative sentiments
  - We then use machine-learning to develop a model that predicts whether any arbitrary text expresses positive or negative sentiments
  - We apply that model to predict the sentiment of the remaining unlabeled texts
- **Unsupervised learning**
  - Use unlabeled data to learn from word co-occurrences in the data to infer unknown categories (e.g. topic models)

2. Non-statistical (dictionary-based) approaches

- Simple interpretation and implementation
- Basic idea:
  - Categorize words or phrases that one believes express positive or negative sentiment
  - In a given document (e.g. social media post), count the number of words that are positive and negative
  - Take the difference between the number of positive and negative words to calculate an overall sentiment score for that document
- The challenge is to develop a good dictionary, but many already existing in multiple languages

Benefits to dictionary-based approaches

- **Efficient**: Once a dictionary is constructed, dictionary approaches are extremely simple and fast
- **Scope**: Are applicable across wide ranges of texts
- **Reliable**: Will produce the same results if you apply them to documents produced now or if you apply them to the same documents in ten years from now

Drawbacks to dictionary-based approaches

- **Not domain-specific**: Word meaning may differ strongly across document types. (e.g. Twitter versus face-to-face conversation)
- **Changes in vocabulary over time**: Words have different meaning today than they used to ("LOL" used to be used sincerely)

Challenges in creating a dictionary itself

- Most are based on a bag of words with unigrams only
- Homographs (words that are spelled the same, but have different meanings, e.g. lie)
- Context specificity, especially for negative modifiers (e.g. "not happy", "no good", "never suitable")
- Words likely to carry more weight than others ("evil" versus "bad")

"Manual coding may be likened to the perspective of a beat cop in a specific neighborhood, rich in context and detail-oriented, while computer automation offers a bird's eye view, like a helicopter pilot circling the city to monitor overall crime patterns" (Young et al. 2012)

On the one hand, using manual coding in conjunction with supervised learning will almost always dominate a dictionary-based approach

- You can define the concept as you want as a researcher
- You can tailor the coding scheme to your particular research objective
- Annotators can take into account the context

On the other hand, manual coding can be very costly, and if a dictionary approach will be sufficient, then use it. No need to use fancier methods if you don't need to.

Are many sentiment dictionaries

- LIWC (Linguistic Inquiry and Word Count)
- ANEW (Affective Norms for English Words)
- LSD (Lexicoder Sentiment Dictionary)
- Bing (Bing Liu et al. dictionary)
- NRC (National Research Council)
- AFINN (?)

Basic idea in application...

- Simply count up the words that are negative and those that are positive (or a similar approach)

Example of a negative sentiment Tweet:

@UK_Together: Brian Wilson says **breaking** away from the UK is a **risk** rural communities seem **unwilling** to take. (4 August 2014)

Example of a positive sentiment Tweet:

@YesScotland: With a Yes, we can make Scotland's **wealth** work better – creating better jobs, providing **decent** pensions and investing in childcare. (15 July 2014)

An early application to social media data

Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures

Scott A. Golder* and Michael W. Macy

We identified individual-level diurnal and seasonal mood rhythms in cultures across the globe, using data from millions of public Twitter messages. We found that individuals awaken in a good mood that deteriorates as the day progresses—which is consistent with the effects of sleep and circadian rhythm—and that seasonal change in baseline positive affect varies with change in daylength. People are happier on weekends, but the morning peak in positive affect is delayed by 2 hours, which suggests that people awaken later on weekends.

Individual mood is an affective state that is important for physical and emotional well-being, working memory, creativity, decision-making (*1*), and immune response (*2*). Mood is influenced by levels of dopamine, serotonin, and other neurochemicals (*1*), as well as by levels of hormones (e.g., cortisol) (*3*). Mood is also externally modified by social activity, such as daily routines of work, commuting, and eating (*4*, *5*). Because of this complexity, accurate measurement of affective rhythms at the individual level has proven elusive.

1878 30 SEPTEMBER 2011 VOL 333 **SCIENCE** www.scie

Research motivation

- Little data to examine diurnal and seasonal mood
- Most studies use US undergraduates
- Most data are self-reported
- Social media data provide massive fine-grained behavioral measures to capture individual mood over time

Research setup

- Collect Twitter posts from 2.4 million users and 509 million messages in 84 countries with English speakers (February 2008-January 2010)
- Use the LIWC sentiment dictionary to measure positive and negative sentiment for each tweet

Diurnal mood based on LIWC dictionary across days

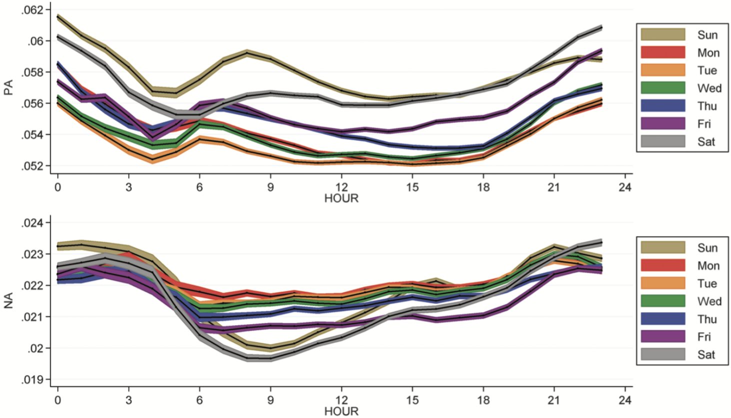

Fig. 1. Hourly changes in individual affect broken down by day of the week (top, PA; bottom, NA). Each series shows mean affect (black lines) and 95% confidence interval (colored regions).

Diurnal mood based on LIWC dictionary across cultures

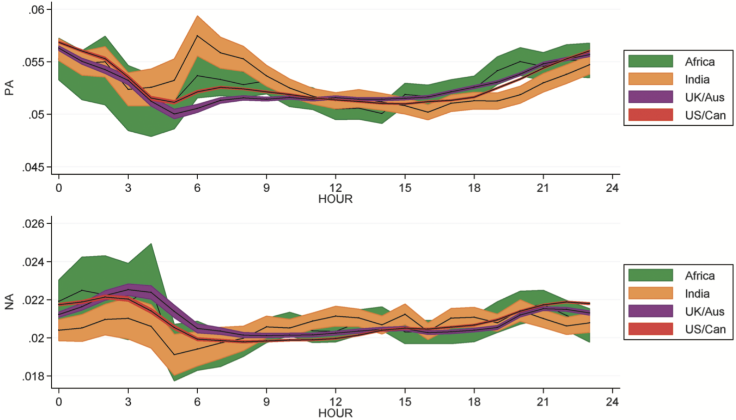

Fig. 2. Hourly changes in individual affect in four English-speaking regions. Each series shows mean affect (black lines) and 95% confidence interval (colored regions).

TECHNOLOGY

Everything We Know About Facebook's Secret Mood Manipulation Experiment

It was probably legal. But was it ethical?

ROBINSON MEYER JUNE 28, 2014

A controversial application in political science

Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer[a,1], Jamie E. Guillory[b], and Jeffrey T. Hancock[c,d]

[a]Core Data Science Team, Facebook, Inc., Menlo Park, CA 94025; [b]Center for Tobacco Control Research and Education, University of California, San Francisco, CA 94143; and Departments of [c]Communication and [d]Information Science, Cornell University, Ithaca, NY 14853

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

Emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness. Emotional contagion is well established in laboratory experiments, with people transferring positive and negative emotions to others. Data from a large real-world social network, collected over a 20-y period suggests that longer-lasting moods (e.g., depression, happiness) can be transferred through networks [Fowler JH, Christakis NA (2008) *BMJ* 337:a2338], although the results are controversial. In an experiment with people who use Facebook, we test whether emotional contagion occurs outside of in-person interaction between individuals by reducing the amount of emotional content in the News Feed. When positive expressions were reduced, people produced fewer positive posts and more negative posts; when negative expressions were reduced, the opposite pattern occurred. These results indicate that emotions expressed by others on Facebook influence our own emotions, constituting experimental evidence for massive-scale contagion via social networks. This work also suggests that, in contrast to prevailing assumptions, in-person interaction and nonverbal cues are not strictly necessary for emotional contagion, and that the observation of others' positive experiences constitutes a positive experience for people.

computer-mediated communication | social media | big data

demonstrated that (*i*) emotional contagion occurs via text-based computer-mediated communication (7); (*ii*) contagion of psychological and physiological qualities has been suggested based on correlational data for social networks generally (7, 8); and (*iii*) people's emotional expressions on Facebook predict friends' emotional expressions, even days later (7) (although some shared experiences may in fact last several days). To date, however, there is no experimental evidence that emotions or moods are contagious in the absence of direct interaction between experiencer and target.

On Facebook, people frequently express emotions, which are later seen by their friends via Facebook's "News Feed" product (8). Because people's friends frequently produce much more content than one person can view, the News Feed filters posts, stories, and activities undertaken by friends. News Feed is the primary manner by which people see content that friends share. Which content is shown or omitted in the News Feed is determined via a ranking algorithm that Facebook continually develops and tests in the interest of showing viewers the content they will find most relevant and engaging. One such test is reported in this study: A test of whether posts with emotional content are more engaging.

The experiment manipulated the extent to which people ($N =$ 689,003) were exposed to emotional expressions in their News Feed. This tested whether exposure to emotions led people to

Is emotion contagious?

- Offline, people are shown to take on the same emotions as those around them
- Well-established in laboratory research
- But much past research is observational
- Some think that "contagion" is due to *interacting* with a happy or sad person, not due to the emotion itself
- Others think that happiness might depress emotion through social comparison

Experimental setup:

- Manipulate the news feed of Facebook users
- Two different experimental treatments:
  - Reduce exposure to negative emotional content from friends
  - Reduce exposure to positive emotional content from friends
- Determine positive or negative posts based on a dictionary-based method (researchers saw none of the actual content)
- Roughly 155,000 users per condition

Measurement of positive and negative sentiment

Posts were determined to be positive or negative if they contained at least one positive or negative word, as defined by Linguistic Inquiry and Word Count software (LIWC2007) (9) word counting system, which correlates with self-reported and physiological measures of well-being, and has been used in prior research on emotional expression (7, 8, 10). LIWC was adapted to run on the Hadoop Map/Reduce system (11) and in the News Feed filtering system, such that no text was seen by the researchers. As such, it was consistent with Facebook's Data Use

Outcomes:

- Percent of positive words written by a user
- Percent of negative words written by a user

Empirical expectations if emotional contagion true:

- Those in the positivity-reduced condition should be less positive compared with the control
- Those in the negativity-reduced condition should be less negative with the control

Empirical expectations if cross-emotional contagion true:

- Those in the positivity-reduced condition should express increased negativity
- Those in the negativity-reduced condition should express increased positivity

Main results:

- When positive posts were reduced, people posted fewer positive words in their status updates, and increased negative words
- When negative posts were reduced, people posted fewer negative words in their status updates, and increased positive words
- Is therefore evidence of mass emotional contagion

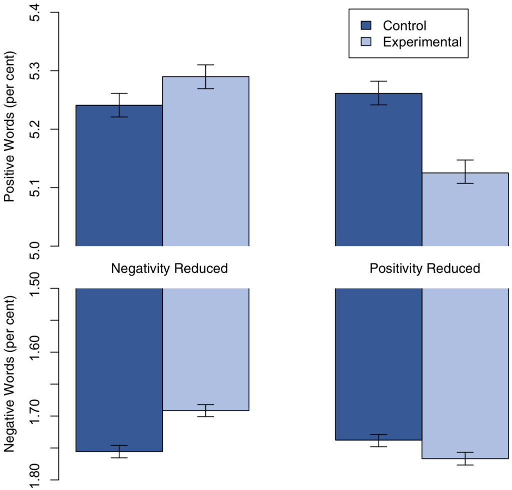

Fig. 1. Mean number of positive (*Upper*) and negative (*Lower*) emotion words (percent) generated people, by condition. Bars represent standard errors.

Implications:

- Because treatment was posts online, we know the mechanism is not interacting with negative/positive people
  - Contagion, thus does not require non-verbal behavior
- No negativity bias, because relatively equal effects for positive and negative posts
- Withdrawal effect: People are less expressive when less exposed to emotional posts
- No evidence that seeing more positive posts leads to negativity (through social comparison)
- Effects sizes are very small, but may have large aggregated consequences

Examining campaign tone with sentiment analysis

Tones from a Narrowing Race: Polling and Online Political Communication during the 2014 Scottish Referendum Campaign

EVELYNE BRIE AND YANNICK DUFRESNE*

The use of negative political communication is a predominant characteristic of modern politics. However, literature doesn't provide an answer to the following question: what explains fluctuations in the use of negative messages within political organisations during a given political campaign? The present paper examines this question in the context of the 2014 Scottish independence referendum. Data consists of all tweets distributed by the official Twitter account of both campaign organisations (@YesScotland and @UK_Together) between June 16, 2014 and September 17, 2014. Results are obtained by a non-parametric local regression and by time-series regression analyses. Our model demonstrates that having an advance in the polls had a statistically significant influence on the tweet sentiment of at least one organisation during the referendum campaign: Better Together's messages were more negative when it was ahead in the polls. Meanwhile, Yes Scotland's messages were more negative after each of the leaders' debates.

Keywords: Sentiment Analysis; Twitter; Referendum; United Kingdom; Scotland

Research motivation

- Negative political communication is tied to a variety of outcomes, including political preferences and voter turnout
- Evidence from psychology suggest that negative emotions tend to prevail over positive ones
- There are clear incentives to go negative: attract more attention and more likely to be remembered than positive ones
- Do we actually start to see campaigns go negative when their chance of winning diminishing?

Research setup: The 2014 Scottish referendum

- Two clearly defined sides: Yes Scotland & Better Together
- Previous research suggests defenders of the status quo will go negative in general
- But authors also examine whether negative sentiment increased as the fortunes of the Yes campaign diminished

Past research

- Research shows that vote choice is determined more by negative campaigning than by positive campaigning
- Most political strategists suggest that negative communication is an effective strategy
- Better Together messages concentrated on the losses associated with a YES vote; Yes Scotland messages concentrated on the gains from an independent Scotland

Reasons to study negative campaigning during referendums

1. Are, by definition, about positional issues
2. Two sides, so simpler to study
3. Dichotomous competition should favor negative campaigning, because a loss for one side is a gain for the other

Competitive context

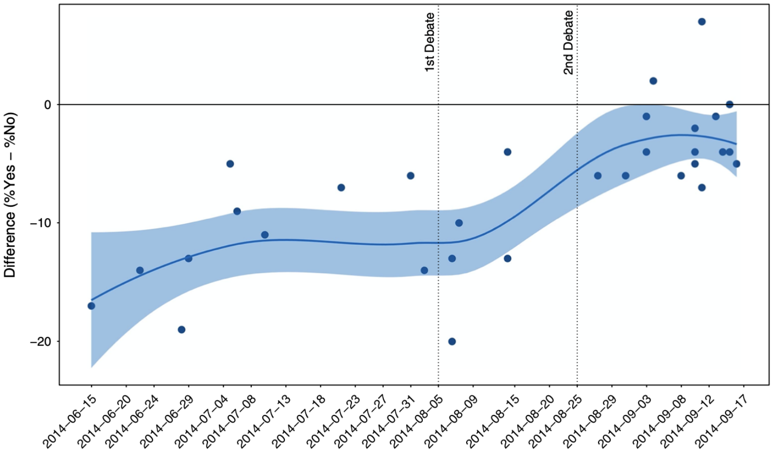

Fig. 1. The evolution of public opinion in the referendum campaign.

Data and Method

- Use tweets as a proxy for political communication during the Yes Scotland and Better Together campaigns
- Collect tweets from @YesScotland (n = 3,078) and @UK_Together (n = 1,230) during the last 3 months of the campaign
- Apply a dictionary-based sentiment analysis to all tweets from each campaign using Bing Liu, Minqing Hu, and Junsheng Cheng (2005)
- Sentiment scores are the difference between the number of positive and negative words in each tweet

Example of a negative sentiment Tweet:

@UK_Together: Brian Wilson says **breaking** away from the UK is a **risk** rural communities seem **unwilling** to take. (4 August 2014)

Example of a positive sentiment Tweet:

@YesScotland: With a Yes, we can make Scotland's **wealth** work better – creating better jobs, providing **decent** pensions and investing in childcare. (15 July 2014)

Results

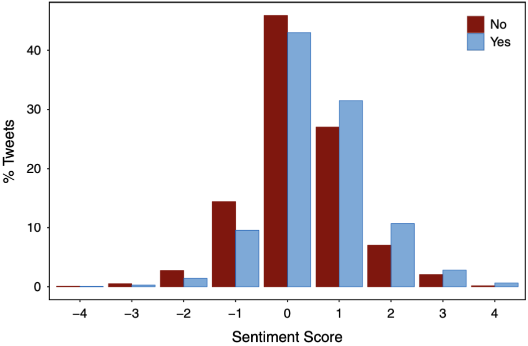

Fig. 2. Tweet sentiment distribution of all messages diffused by both official Twitter accounts during the Scottish referendum campaign.
Source: Official Twitter accounts of Better Together (N = 1,230) and of Yes Scotland (N = 3,078) (16 June to 17 September 2014).

Results

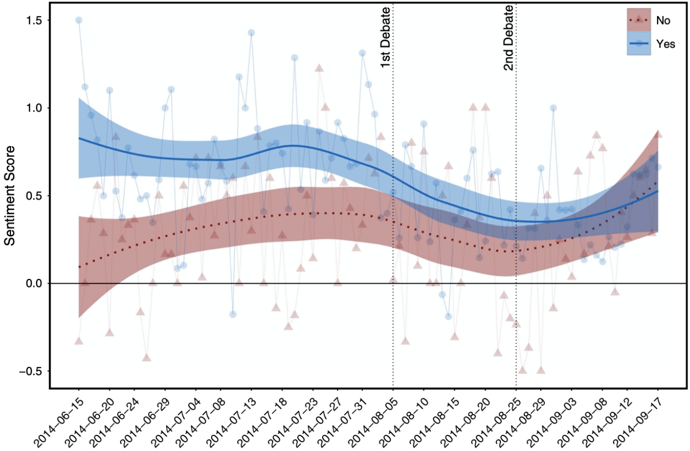

Fig. 3. The daily evolution of tweet sentiment.
Source: Official Twitter accounts of Better Together (N = 1,230) and of Yes Scotland (N = 3,078) (16 June to 17 September 2014). The first and second televised debates were respectively held on 5 and 25 August 2014.

Results

TABLE 1 *Testing the effect of public opinion and timing on sentiment score*

	Sentiment Score (Daily Average)					
	YES Campaign			NO Campaign		
Poll Lead	0.003 (0.01)		−0.01 (0.01)	−0.04* (0.01)		−0.05** (0.02)
Momentum		−0.001 (0.01)	−0.001 (0.01)		−0.01 (0.01)	0.01 (0.01)
Post-debate 1			−0.31* (0.15)			0.08 (0.19)
Post-debate 2			−0.38* (0.18)			−0.20 (0.24)
Poll Day	0.14 (0.07)	0.15* (0.07)	0.15* (0.07)	0.03 (0.09)	0.09 (0.10)	0.03 (0.10)
Sentiment Score$_{t-1}$	0.20* (0.10)	0.20* (0.10)	0.13 (0.10)	0.005 (0.10)	0.08 (0.10)	−0.05 (0.11)
Days to Referendum	0.003** (0.001)	0.003* (0.001)	−0.002 (0.003)	0.0002 (0.002)	0.001 (0.002)	−0.002 (0.004)
Constant	0.27*** (0.08)	0.27** (0.08)	0.74** (0.22)	0.29* (0.10)	0.23* (0.10)	0.45 (0.28)
N	94	94	94	94	94	94
R^2	0.18	0.18	0.23	0.11	0.03	0.15

*p < 0.05; **p < 0.01; ***p <0.001.
Source: Official Twitter accounts of *Better Together* (N = 1,230) and of *Yes Scotland* (N = 3,078), and What Scotland Thinks (16 June to 17 September 2014).

Dictionary versus Data-driven Approaches to Sentiment

Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods

Kokil Jaidka[a,b,1], Salvatore Giorgi[c], H. Andrew Schwartz[d], Margaret L. Kern[e], Lyle H. Ungar[c], and Johannes C. Eichstaedt[f,g,1]

Can sentiment analysis of social media data capture well-being?

- Collecting measures of well-being is time- and resource-intensive
- Dictionary-based approaches are a potentially simple way to measure well-being
- Jaidka et al. (2020) assess the value of dictionary and data-driven approaches to sentiment analysis

Can sentiment analysis of social media data capture well-being?

- LIWC a pre-determined set of positive and negative words
- Creators of the LabMT and ANEW dictionaries asked raters to annotate words for their valence
- Data-driven methods using machine learning to identify the assocations between linguistic information and its emotional content (whole sentences annotated by rates)

Authors assess how well these approaches predict actual well-being

- Aggregate 1.73 million responses to the Gallup-Sharecare Well-Being Index from 2009 to 2015 to get county-level measures of life satisfaction, happiness, worry, and sadness

Results

Table 2. Pearson correlations (*r*) between Twitter-based emotions and Gallup-Sharecare Well-Being Index estimates across 1,208 US counties

N = 1,208 U.S. counties	Word-level									Data-driven				
	LIWC 2015			PERMA		ANEW		LabMT		Sentence-level			Person-level	
										WWBP	Swiss Chocolate			
	Positive	Positive (modified)	Negative	Positive	Negative	Valence	Valence (modified)	Valence	Valence (modified)	Affect	Positive	Negative	WWBP Life Sat.	Direct prediction
Life Satisfaction	-.21	-.06	-.32	.22	-.37	-.03	.15	-.27	.01	.29	.24	-.29	.39	.62
Happiness	-.13	.13	-.27	.27	-.17	.04	.18	-.07	.16	.23	.24	-.30	.23	.51
Worry	.11	.01	.03	-.01	.02	.03	-.05	.02	-.04	.00	-.02	.11	-.03	.52
Sadness	.25	-.01	.22	-.19	.18	.09	-.10	.19	-.09	-.18	-.20	.33	-.23	.64

The gray column headers identify the modified LIWC (removed 3 words), LabMT (removed 15 words), and ANEW (removed 2 words) dictionaries (in the text). The color indicates the direction and magnitude of correlation; white cells are nonsignificant, and all others are $P < 0.05$ corrected for multiple comparisons.

Further results

- Data-driven approaches also better for predicting actual county-level health and socio-economic outcomes
- Correcting for sample differences between Twitter and Gallup didn't change anything
- Looked at the individual level with Facebook users who had answered the Gallup survey, and dictionary-based methods also poor

But why are dictionary methods so bad at predicting well-being?

- Examine the positive dictionary words that are negatively predictive of positive well-being
- Examine the negative dictionary words that are positively predictive of positive well-being...

..

Top right: “Positive” words that *negatively* predict happiness
Bottom left: “Negative” words that *positively* predict happiness

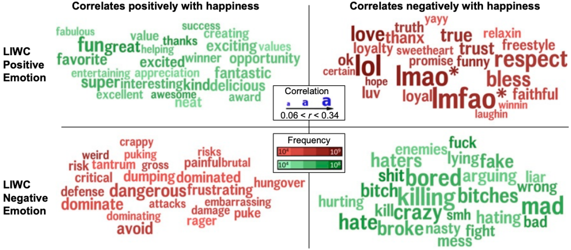

Fig. 1. Sources of error in the LIWC positive and negative emotion dictionaries. The matrix illustrates the 25 most frequent words from the two dictionaries that were correlated as expected (green indicates true LIWC positives and true negatives) or opposite to expectation (red indicates false positives and false negatives) with the Gallup happiness item. The size of the word denotes the magnitude of its correlation (0.06 $< r <$ 0.34; $P <$ 0.05 corrected for multiple comparisons). The shade indicates the normalized frequency, with darker shades reflecting higher frequencies relative to other words.

- The words “LOL”, “love”, and “good” are some of the most frequent words, accounting for about 25% of the county word occurences
- Yet these are some of those that are negatively associated with well-being (and income)
- Removing them uniformly improved convergence with the Gallup measure (gray columns in Table 2)

False LIWC words are also most frequently used in certain regions

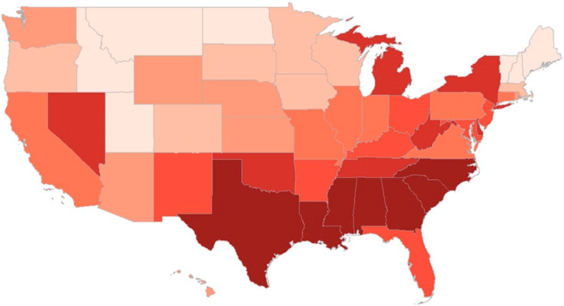

Fig. 2. The relative frequency of false LIWC positive emotion words across the United States. States with a darker shade of red had relatively higher numbers of positive emotion words that correlated negatively with county Gallup happiness (Fig. 1, *Upper Right*) at $P < 0.05$, controlling for multiple comparisons.

- Dictionary methods should be used cautiously
- Most discrepanices are driven by a few frequent words (LOL, love, good), that are now often used ironically or sarcastically
- Data-driven methods have higher convergent validity with ground-truth measures of well-being
    - But costly to implement