

# Lecture Three

Stories of Complex Sociotechnical Systems:  
Measurement, Mechanisms, and Meaning  
Lipari Summer School, Summer, 2012

Prof. Peter Dodds

Department of Mathematics & Statistics | Center for Complex Systems |  
Vermont Advanced Computing Center | University of Vermont



Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License.

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

Positivity Bias

References



1 of 83

## Press...

- ▶ "Social Scientists waded into the Tweet stream" by Greg Miller, *Science*, **333**, 1814–1815, 2011 [15]
- ▶ "Does a Nation's Mood Lurk in Its Songs and Blogs?" by Benedict Carey *New York Times*, August 2009. (田)
- ▶ More here: <http://www.uvm.edu/~pdodds/research/> (田)



Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

Positivity Bias

References



4 of 83

## Outline

### Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

### Analysis

Songs  
Blogs  
Tweets

### Positivity Bias

### References

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

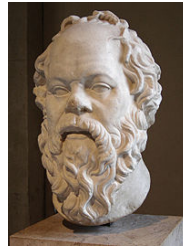
Positivity Bias

References



2 of 83

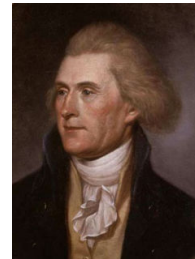
## Happiness:



Socrates et al.:  
eudaimonia [8]



Bentham:  
hedonistic calculus



Jefferson:  
... the pursuit of happiness

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

Positivity Bias

References



6 of 83

## Papers and so on:



"Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter"  
Dodds et al., PLoS ONE, 2011 [7]  
Much better version here:  
<http://arxiv.org/abs/1101.5120> (田)

- ▶ "Positivity of the English Language"  
Kloumann et al., PLoS ONE, 2012 [11]
- ▶ "Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents"  
Dodds and Danforth, *Journal of Happiness Studies*, 2009 [6]
- ▶ language assessment by Mechanical Turk (labMT 1.0)
- ▶ <http://www.onehappybird.com> (田)

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

Positivity Bias

References



3 of 83

## Early drafts:

that among these are:  
Life, Liberty, and ?? Money?  
~~Libations~~ ... ~~Alcohol~~  
~~Property~~ ~~Foot-the-ball~~  
~~Beer~~  
Happiness

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

Positivity Bias

References



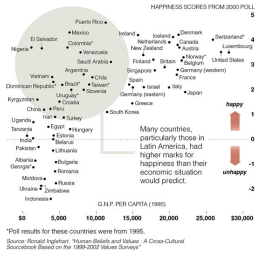
7 of 83

## Desiring happiness—not just for boffins:

- ▶ Average people routinely report being happy is what they want most in life [12, 13, 5]
- ▶ And it matters: “Happy people live longer: . . .” Survey by Diener and Chan. [5]

### A Plateau of Happiness

A country's wealth may not always dictate the happiness of its people. As part of the World Values Survey project, inhabitants of different countries and territories were asked how happy or satisfied they were. Below is a sampling of happiness rankings, along with economic status.



### National indices of well-being:

- ▶ Bhutan
- ▶ France
- ▶ Australia

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



UNIVERSITY OF VERMONT  
8 of 83

## Emotional content

### So how does one measure

1. happiness?
2. levels of other emotional states?

### Just ask people how happy they are.

- ▶ Experience sampling [2, 4, 3] (Csikszentmihalyi et al.)
- ▶ Day reconstruction [9] (Kahneman et al.)

### But self-reporting has some drawbacks:

- ▶ relies on memory and self-perception
- ▶ induces misreporting [14]
- ▶ costly

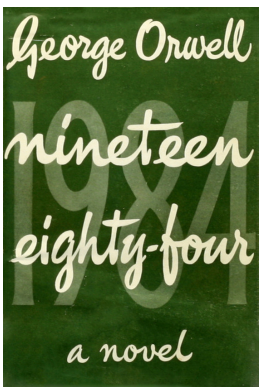
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References

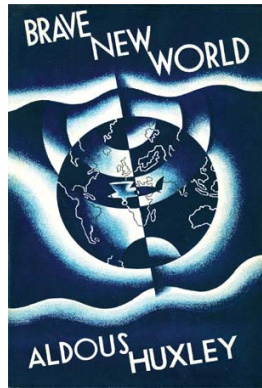


UNIVERSITY OF VERMONT  
12 of 83

## An easy knock:



Science = Orwell



Policy = Brave New World

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



UNIVERSITY OF VERMONT  
10 of 83

## Happiness, attention, and doing:

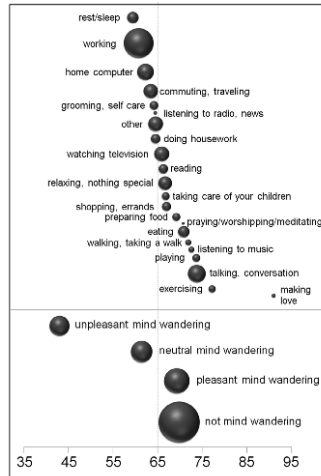


Fig. 1. Mean happiness reported during each activity (top) and while mind wandering to unpleasant topics, neutral topics, pleasant topics or not mind wandering (bottom). Dashed line indicates mean of happiness across all samples. Bubble area indicates the frequency of occurrence. The largest bubble (“not mind wandering”) corresponds to 53.1% of the samples, and the smallest bubble (“praying/worshipping/meditating”) corresponds to 0.1% of the samples.

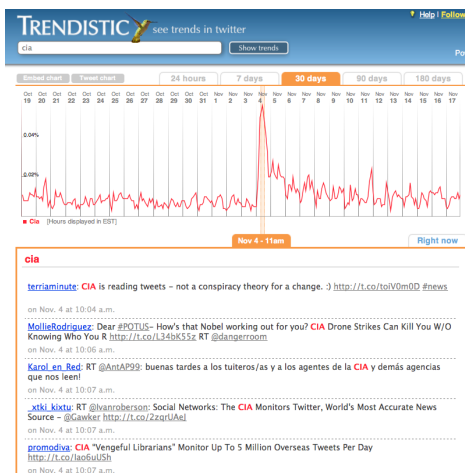
Killingsworth and Gilbert, Science, 2010 [10]

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



UNIVERSITY OF VERMONT  
13 of 83



- ▶ See story [here](#) (📄) for example [slate].

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



UNIVERSITY OF VERMONT  
11 of 83

## We'd like to build an 'hedonometer':



- ▶ An instrument to 'remotely-sense' emotional states and levels, in real time or post hoc.

### Ideally:

- ▶ Transparent
- ▶ Fast
- ▶ Based on written expression
- ▶ Uses human evaluation
- ▶ Non-reactive
- ▶ Complementary to self-reported measures
- ▶ Improvable

### Some possibilities:

- ▶ Natural language processing (e.g., OpinionFinder)
- ▶ Declared mood levels in blogs (e.g., Livejournal) [16]

Complex Sociotechnical Systems

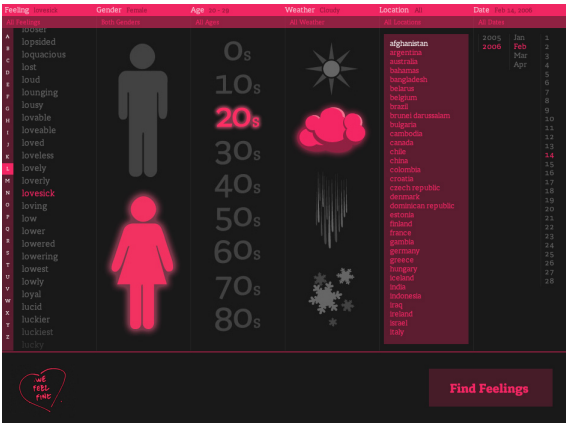
Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



UNIVERSITY OF VERMONT  
14 of 83

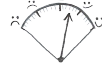


wefeelfine.org:



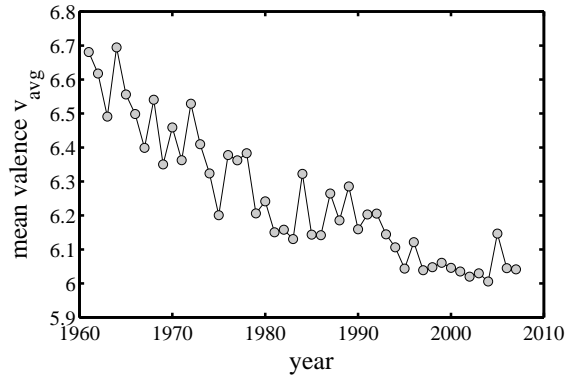
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



22 of 83

Song Lyrics—average happiness (valence)



Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



26 of 83

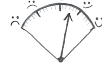
More data sets:



- 6. New York Times (20 years)
- 7. Gutenberg.org
- 8. Google Books: <http://ngrams.googlelabs.com/> (田)
- 9. ...

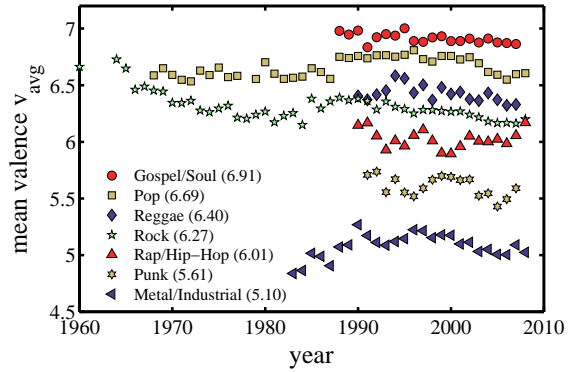
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



23 of 83

Song Lyrics—average happiness of genres:



Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



27 of 83

Some numbers:

Counts	Song lyrics	Song titles
All words	58,610,849	60,867,223
ANEW words	3,477,575 (5.9%)	5,612,708 (9.2%)
Individuals	~ 20,000	~ 632,000

Counts	blogs	SOTU
All words	155,667,394	1,796,763
ANEW words	8,581,226 (5.5%)	61,926 (3.5%)
Individuals	~ 2,335,000	43

Counts	Twitter
All words	~ 100 billion
Tweets	~ 10 billion
Individuals	~ 100 million

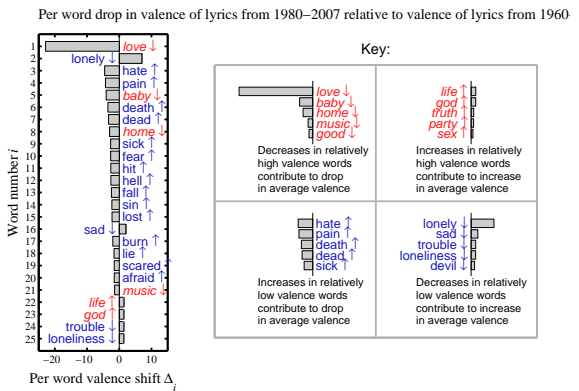
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



24 of 83

Happiness Word Shift Graph:



Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



28 of 83

### Top 16 of $\approx 20,000$ artists:

Rank	Artist	Valence
1	All-4-One	7.15
2	Luther Vandross	7.12
3	S Club 7	7.05
4	K Ci & JoJo	7.04
5	Perry Como	7.04
6	Diana Ross & The Supremes	7.03
7	Buddy Holly	7.02
8	Faith Evans	7.01
9	The Beach Boys	7.01
10	Jon B	6.98
11	Dru Hill	6.96
12	Earth Wind & Fire	6.95
13	Ashanti	6.95
14	Otis Redding	6.93
15	Faith Hill	6.93
16	NSync	6.93

(criteria:  $\geq 50$  songs and  $\geq 1000$  ANEW words)

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



29 of 83

Text:	$h_{avg}$	Words with a similar score:
Soul/Gospel lyrics [6]	6.9	chocolate (6.88), leisurely (6.88), penthouse (6.81)
Pop lyrics [6]	6.7	dream (6.73), honey (6.73), sugar (6.74)
Dante's Paradise	6.5	muffin (6.57), rabbit (6.57), smooth (6.58)
Tweets, 9/9/2008 to 12/31/2010	6.4	thought (6.39), face (6.39), blond (6.42)
Rock lyrics [6]	6.3	church (6.28), tree (6.32), air (6.34)
Enron Emails [11]	6.2	clouds (6.18), alert (6.20), computer (6.24)
State of the Union Messages [6]	6.1	grass (6.12), idol (6.12), bottle (6.15)
New York Times (1987-2007) [17]	6.0	hotel (6.00), tennis (6.02), wonder (6.03)
Blogs [6]	5.8	owl (5.80), whistle (5.81), humble (5.86)
Dante's Inferno	5.5	glacier (5.50), repentant (5.53), mischief (5.57)
Heavy Metal lyrics [6]	5.4	lamp (5.41), elevator (5.44), truck (5.47)

### Bottom 16 of $\approx 20,000$ artists:

Rank	Artist	Valence
1	Slayer	4.80
2	Misfits	4.88
3	Staind	4.93
4	Slipknot	4.98
5	Darkthrone	4.98
6	Death	5.02
7	Black Label Society	5.05
8	Pig	5.08
9	Voivod	5.14
10	Fear Factory	5.15
11	Iced Earth	5.16
12	Simple Plan	5.16
13	Machine Head	5.17
14	Metallica	5.19
15	Dimmu Borgir	5.20
16	Mudvayne	5.21

(criteria:  $\geq 50$  songs and  $\geq 1000$  ANEW words)

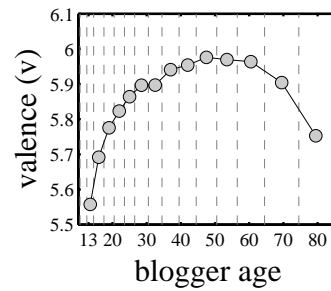
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



30 of 83

### Blogs



► Average happiness as a function of the age bloggers report they will turn in the year of their posting.

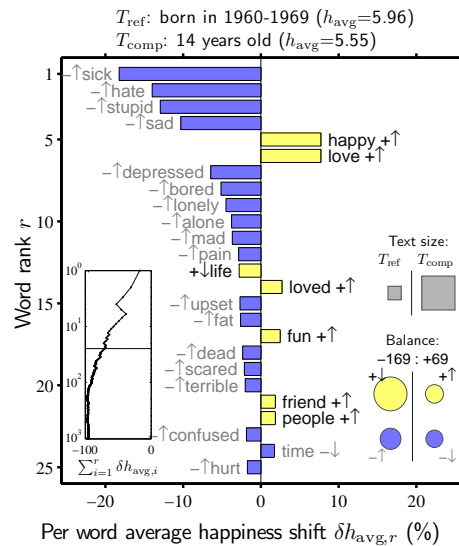
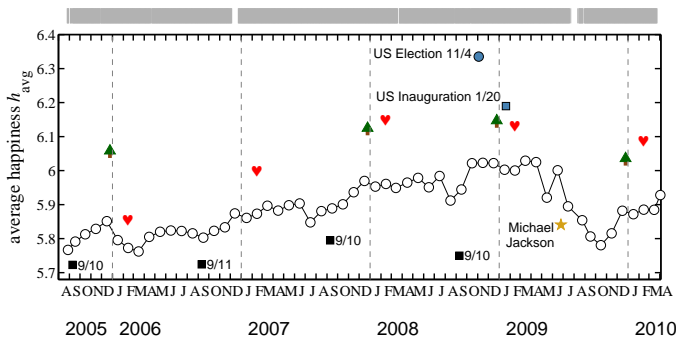
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



34 of 83

### Blogs—Overall trend

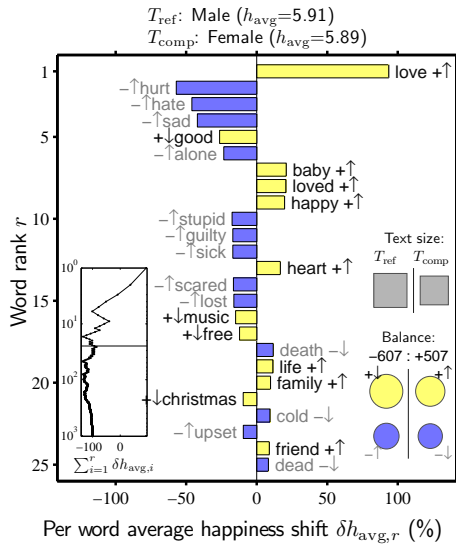


Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



35 of 83



Complex Sociotechnical Systems

Measuring Happiness  
 Some motivation  
 Measuring emotional content  
 Data sets  
 Analysis  
 Songs  
 Blogs  
 Tweets  
 Positivity Bias  
 References



UNIVERSITY OF VERMONT  
 ↻ 🔍 36 of 83

## Words most correlated with obesity levels in cities:

Word	$h_{avg}$	$r_s$	$p$ -value
stomach	5.40	0.37	1.98894e-07
mcdonalds	5.98	0.30	2.60824e-05
hungry	3.38	0.27	0.000206297
wings	6.52	0.25	0.000388915
ham	5.66	0.24	0.000763101
starving	2.58	0.22	0.00272286
spaghetti	0.00	0.20	0.00689403
ihop	0.00	0.19	0.0100034
noodles	0.00	0.18	0.0106139
ketchup	0.00	0.18	0.0145088
fat	3.24	0.18	0.0148845
sprite	0.00	0.17	0.0175705
cookin	0.00	0.17	0.0182976
heartburn	0.00	0.17	0.0200551
sugar	6.74	0.15	0.0329359
kool-aid	0.00	0.15	0.0354226
miller	5.36	0.15	0.036325
honey	7.44	0.15	0.0395531
candy	7.52	0.15	0.0398618

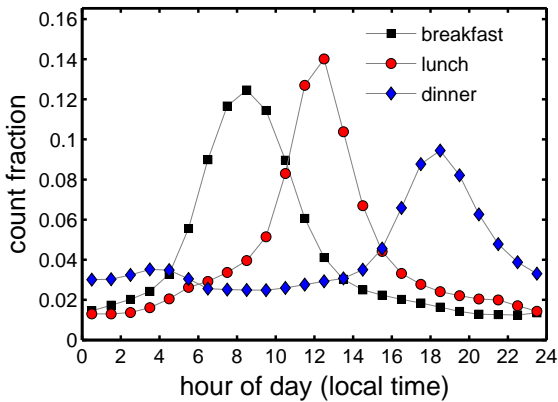
Complex Sociotechnical Systems

Measuring Happiness  
 Some motivation  
 Measuring emotional content  
 Data sets  
 Analysis  
 Songs  
 Blogs  
 Tweets  
 Positivity Bias  
 References



UNIVERSITY OF VERMONT  
 ↻ 🔍 40 of 83

## Twitter—living in the now:



Complex Sociotechnical Systems

Measuring Happiness  
 Some motivation  
 Measuring emotional content  
 Data sets  
 Analysis  
 Songs  
 Blogs  
 Tweets  
 Positivity Bias  
 References



UNIVERSITY OF VERMONT  
 ↻ 🔍 38 of 83

## Words most anti-correlated with obesity levels in cities:

Word	$h_{avg}$	$r_s$	$p$ -value
brunch	6.32	-0.41	6.37431e-09
bar	5.82	-0.35	5.54374e-07
banana	6.86	-0.35	5.67492e-07
barista	0.00	-0.35	7.29324e-07
delicious	7.92	-0.34	1.09807e-06
dinner	7.40	-0.34	1.35413e-06
coffee	7.18	-0.34	2.04145e-06
espresso	0.00	-0.33	4.45903e-06
cocktails	0.00	-0.32	4.96518e-06
booze	0.00	-0.32	6.38461e-06
mimosa	0.00	-0.31	1.24472e-05
spiced	0.00	-0.31	1.52074e-05
veggie	0.00	-0.31	1.60439e-05
sushi	5.40	-0.31	1.71997e-05
wines	6.28	-0.31	1.7432e-05
tofu	0.00	-0.31	1.86278e-05
panini	0.00	-0.31	1.86719e-05
gnocchi	0.00	-0.30	2.51419e-05
clams	0.00	-0.30	2.52124e-05
cafe	5.80	-0.30	2.6263e-05
cafe	5.80	-0.30	2.6263e-05
bento	0.00	-0.30	2.6349e-05
huevos	0.00	-0.30	3.19903e-05

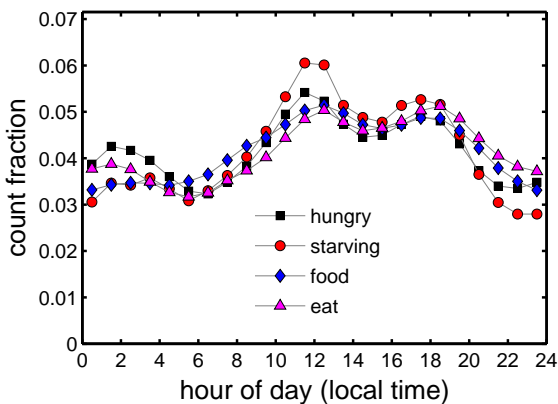
Complex Sociotechnical Systems

Measuring Happiness  
 Some motivation  
 Measuring emotional content  
 Data sets  
 Analysis  
 Songs  
 Blogs  
 Tweets  
 Positivity Bias  
 References



UNIVERSITY OF VERMONT  
 ↻ 🔍 41 of 83

## Twitter—living in the now:



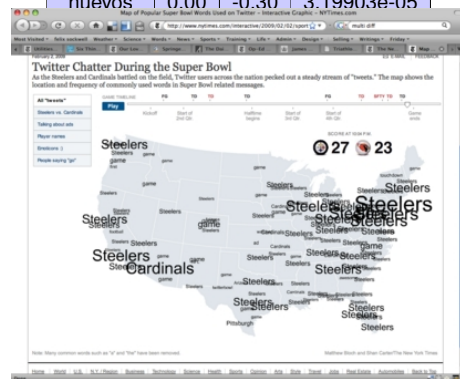
Complex Sociotechnical Systems

Measuring Happiness  
 Some motivation  
 Measuring emotional content  
 Data sets  
 Analysis  
 Songs  
 Blogs  
 Tweets  
 Positivity Bias  
 References



UNIVERSITY OF VERMONT  
 ↻ 🔍 39 of 83

## Twitter—living in the now:



Complex Sociotechnical Systems

Measuring Happiness  
 Some motivation  
 Measuring emotional content  
 Data sets  
 Analysis  
 Songs  
 Blogs  
 Tweets  
 Positivity Bias  
 References



UNIVERSITY OF VERMONT  
 ↻ 🔍 42 of 83

Tweeting the Superbowl (田) [NY Times]



valence rank	word	valence	std dev	twitter rank	g-books rank	nyt rank	lyrics rank
10193	violence	1.86	1.05	4299	1724	1238	2016
10194	cruel	1.84	1.15	2963	—	—	1447
10195	cry	1.84	1.28	1028	3075	—	226
10196	failed	1.84	1.00	2645	1618	1276	2920
10197	sickness	1.84	1.18	4735	—	—	3782
10198	abused	1.83	1.31	—	—	—	4589
10199	tortured	1.82	1.42	—	—	—	4693
10200	fatal	1.80	1.53	—	4089	—	3724
10201	killings	1.80	1.54	—	—	4914	—
10202	murdered	1.80	1.63	—	—	—	4796
10203	war	1.80	1.41	468	175	291	462
10204	kills	1.78	1.23	2459	—	—	2857
10205	jail	1.76	1.02	1642	—	2573	1619
10206	terror	1.76	1.00	4625	4117	4048	2370
10207	die	1.74	1.19	418	730	2605	143
10208	killing	1.70	1.36	1507	4428	1672	998
10209	arrested	1.64	1.01	2435	4474	1435	—
10210	deaths	1.64	1.14	—	—	2974	—
10211	raped	1.64	1.43	—	—	—	4528
10212	torture	1.58	1.05	3175	—	—	3126
10213	died	1.56	1.20	1223	866	208	826
10214	kill	1.56	1.05	798	2727	2572	430
10215	killed	1.56	1.23	1137	1603	814	1273
10216	cancer	1.54	1.07	946	1894	796	3802
10217	death	1.54	1.28	509	307	373	433
10218	murder	1.48	1.01	2762	3110	1541	1059
10219	terrorism	1.48	0.91	—	—	3192	—
10220	rape	1.44	0.79	3133	—	4115	2977
10221	suicide	1.30	0.84	2124	4707	3319	2107
10222	terrorist	1.30	0.91	3576	—	3026	—

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis  
Songs  
Blogs  
Tweets

Positivity Bias

References



49 of 83

std dev rank	word	valence	std dev	twitter rank	g-books rank	nyt rank	lyrics rank
1	ff@king	4.64	2.93	448	—	—	620
2	fx *kin	3.86	2.74	1077	—	—	688
3	fx *ked	3.56	2.71	1840	—	—	904
4	pussy	4.80	2.66	2019	—	—	949
5	whiskey	5.72	2.64	—	—	—	2208
6	slut	3.57	2.63	—	—	—	4071
7	cigarettes	3.31	2.60	—	—	—	3279
8	fx *k	4.14	2.58	322	—	—	185
9	mortality	4.38	2.55	—	3960	—	—
10	cigarette	3.09	2.52	—	—	—	2678
11	motherf* *kers	2.51	2.47	—	—	—	1466
12	churches	5.70	2.46	—	2281	—	—
13	motherf* *king	2.64	2.46	—	—	—	2910
14	capitalism	5.16	2.45	—	4648	—	—
15	porn	4.18	2.43	1801	—	—	—
16	summer	6.40	2.39	896	1226	721	590
17	beer	5.92	2.39	839	4924	3960	1413
18	execution	3.10	2.39	—	2975	—	—
19	wines	6.28	2.37	—	—	3316	—
20	zombies	4.00	2.37	4708	—	—	—
21	aids	4.28	2.35	2983	3996	1197	—
22	capitalist	4.84	2.34	—	4694	—	—
23	revenge	3.71	2.34	—	—	—	2766
24	mcdonalds	5.98	2.33	3831	—	—	—
25	beatles	6.44	2.33	3797	—	—	—
26	islam	4.68	2.33	—	4514	—	—
27	pay	5.30	2.32	627	769	460	499
28	alcohol	5.20	2.32	2787	2617	3752	3600
29	muthaf* *kin	3.00	2.31	—	—	—	4107
30	christ	6.16	2.31	2509	909	4238	1526

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis  
Songs  
Blogs  
Tweets

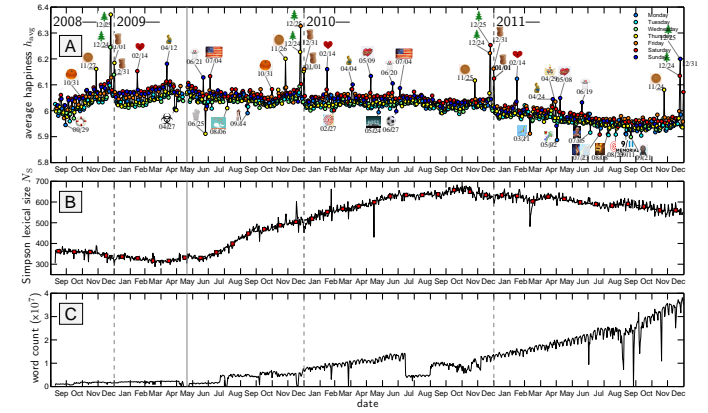
Positivity Bias

References

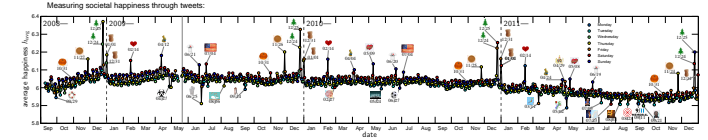


50 of 83

Twitter—overall time series:

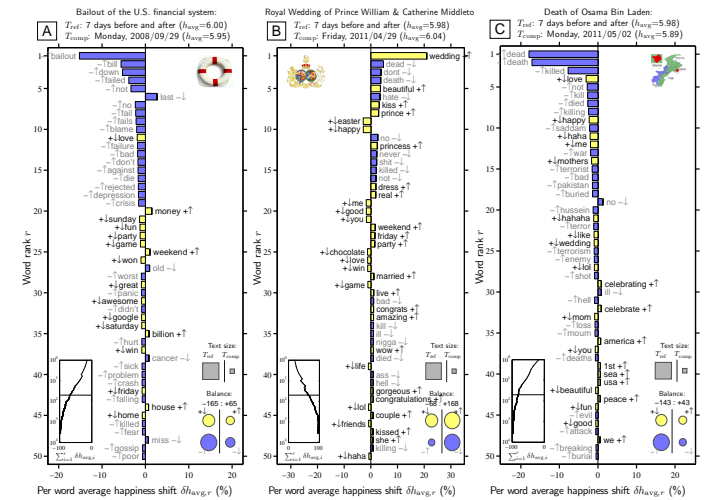
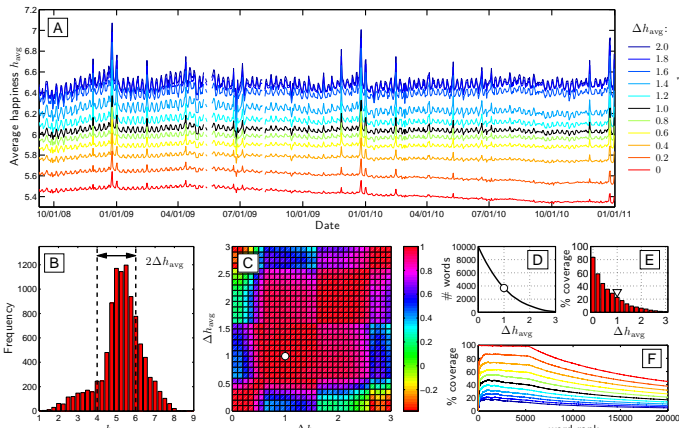


Twitter—overall time series:



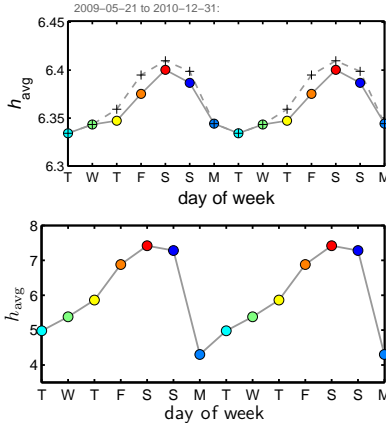
- Global happiness spikes = predictable rituals.
- Global sadness spikes = unpredictable, exogenous shocks.
- No accidental happiness outbreaks.

The very surprising tunable hedonometer:





# Twitter—weekly time series:

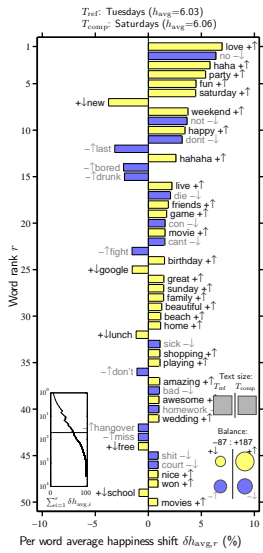
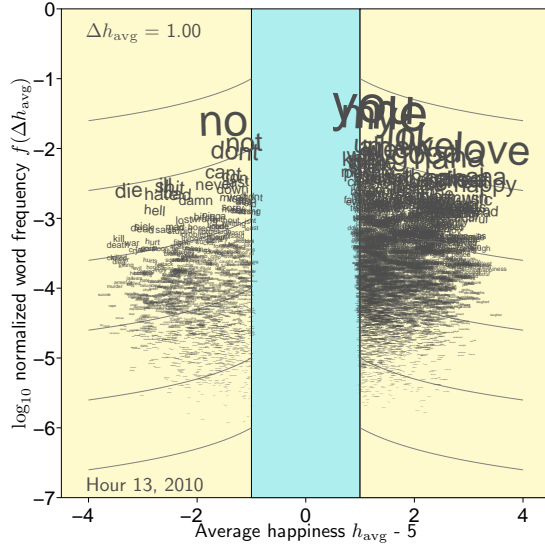


Complex Sociotechnical Systems

- Measuring Happiness
  - Some motivation
  - Measuring emotional content
  - Data sets
- Analysis
  - Songs
  - Blogs
  - Tweets
- Positivity Bias
- References



55 of 83



Complex Sociotechnical Systems

- Measuring Happiness
  - Some motivation
  - Measuring emotional content
  - Data sets
- Analysis
  - Songs
  - Blogs
  - Tweets
- Positivity Bias
- References



56 of 83

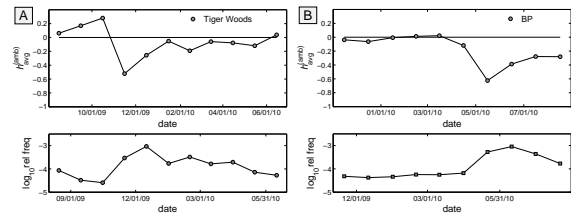
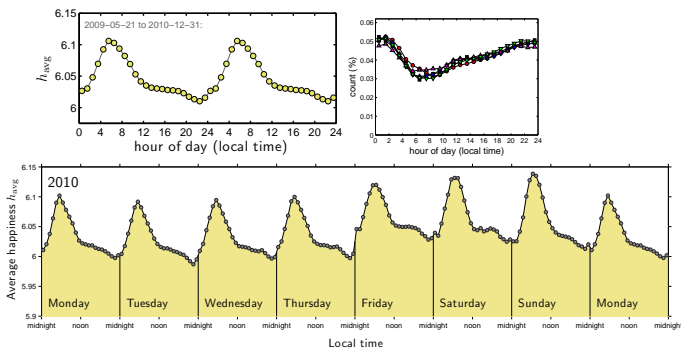
Complex Sociotechnical Systems

- Measuring Happiness
  - Some motivation
  - Measuring emotional content
  - Data sets
- Analysis
  - Songs
  - Blogs
  - Tweets
- Positivity Bias
- References



59 of 83

# The daily unravelling of the human mind:

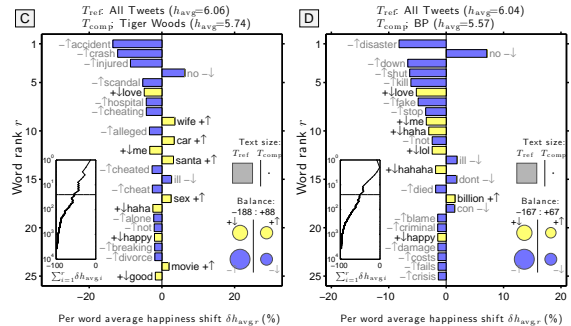


Complex Sociotechnical Systems

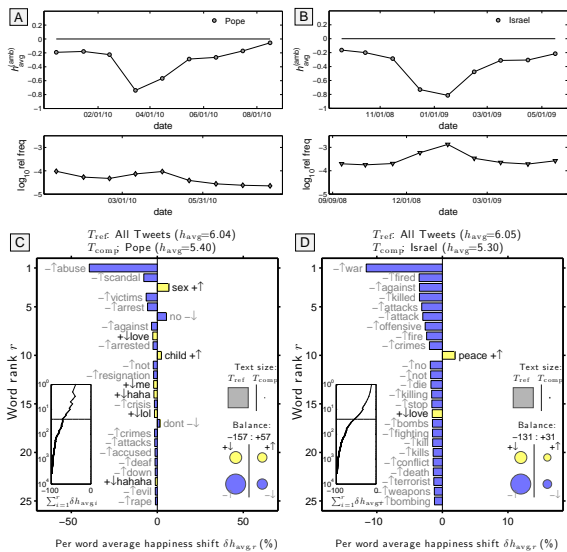
- Measuring Happiness
  - Some motivation
  - Measuring emotional content
  - Data sets
- Analysis
  - Songs
  - Blogs
  - Tweets
- Positivity Bias
- References



60 of 83



# Happiness in Manhattan (just for fun):



Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

Positivity Bias

References



61 of 83

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

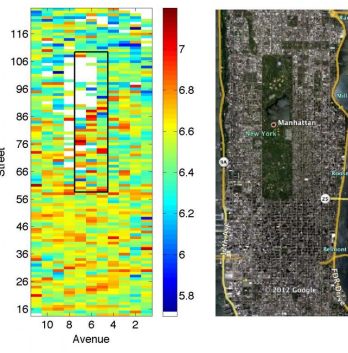
Songs  
Blogs  
Tweets

Positivity Bias

References



64 of 83



See Blog post on onehappybird (田)

# Twitter—location:

Word	$h_{avg}^{(20)}$	Total Tweets	Total ANEW	Word	$h_{avg}^{(20)}$	Total Tweets	Total ANEW
1. love	+1.12	46,687,176 (6)	85,269,499 (5)	91. me	-0.06	144,342,098 (4)	88,888,051 (4)
2. happy	+1.12	16,611,964 (18)	21,442,329 (8)	92. ?	-0.07	2,633,238 (52)	67,479,049 (1)
3. win	+1.26	7,981,856 (26)	14,640,728 (20)	93. commute	-0.09	90,126 (94)	90,092 (92)
4. kiss	+1.21	1,697,405 (59)	3,162,330 (48)	94. guy	-0.09	2,737,909 (47)	1,697,177 (57)
5. cash	+1.21	1,279,236 (63)	2,468,406 (54)	95. right	-0.10	19,166,480 (10)	15,850,283 (19)
6. vacation	+1.11	934,501 (67)	1,783,270 (56)	96. school	-0.11	5,364,217 (24)	6,924,193 (34)
7. Christmas	+1.03	4,887,968 (35)	10,645,630 (25)	97. Republican	-0.13	229,773 (86)	188,338 (85)
8. Cool	+0.85	8,576,384 (25)	17,867,708 (18)	98. they	-0.14	27,442,369 (8)	27,510,189 (11)
9. party	+0.93	6,438,886 (29)	12,090,597 (23)	99. winter	-0.19	1,255,945 (64)	1,217,225 (64)
10. sex	+0.89	3,551,767 (39)	7,687,972 (31)	00. kiss	-0.19	2,056,408 (55)	2,091,540 (53)
11. Valentine	+0.85	247,288 (84)	464,914 (75)	01. Jon Stewart	-0.20	52,084 (97)	33,086 (96)
12. family	+0.79	5,014,816 (32)	10,629,391 (26)	02. gas	-0.22	1,022,879 (95)	812,029 (68)
13. sun	+0.65	2,385,348 (52)	4,602,627 (44)	03. no	-0.22	95,129,093 (5)	38,894,616 (6)
14. life	+0.59	11,698,454 (17)	27,770,708 (16)	04. Democrat	-0.23	98,188 (93)	75,249 (93)
15. hope	+0.48	11,833,337 (18)	22,952,366 (13)	05. left	-0.27	4,893,634 (34)	4,611,878 (43)
16. heaven	+0.43	741,878 (71)	1,435,792 (59)	06. Senate	-0.29	447,732 (78)	316,835 (80)
17. ?	+0.42	10,470,483 (20)	6,787,478 (35)	07. election	-0.30	590,184 (73)	375,055 (78)
18. income	+0.36	510,425 (76)	418,161 (77)	08. Sarah Palin	-0.34	225,577 (87)	150,096 (88)
19. friends	+0.33	7,669,719 (27)	7,541,106 (29)	09. Obama	-0.35	2,981,150 (44)	1,998,326 (54)
20. snow	+0.32	2,596,165 (48)	5,011,795 (40)	10. economy	-0.36	698,678 (75)	469,834 (76)
21. :)	+0.32	1,680,165 (60)	1,102,512 (67)	11. Congress	-0.36	391,510 (79)	279,695 (81)
22. night	+0.29	17,889,345 (12)	17,806,796 (17)	12. drug	-0.39	519,906 (77)	499,091 (74)
23. vegan	+0.28	183,889 (90)	178,670 (86)	13. Muslim	-0.42	215,900 (88)	146,266 (89)
24. Jesus	+0.27	2,027,720 (59)	1,673,992 (58)	14. George Bush	-0.43	32,341 (98)	23,102 (98)
25. girl	+0.25	10,070,132 (22)	19,886,691 (14)	15. climate	-0.44	564,177 (80)	229,129 (83)
26. USA	+0.23	2,577,172 (54)	1,294,588 (95)	16. Pope	-0.51	152,220 (91)	135,955 (90)
27. you	+0.22	173,276,993 (3)	145,464,084 (2)	17. oil	-0.53	1,477,355 (62)	1,148,990 (66)
28. sur	+0.21	11,846,465 (10)	14,437,899 (21)	18. ?	-0.54	5,175,518 (31)	4,702,352 (42)
29. ?	+0.20	2,618,940 (48)	1,475,221 (60)	19. Glenn Beck	-0.54	113,991 (92)	101,090 (91)
30. health	+0.20	2,575,543 (59)	4,950,292 (41)	20. Islam	-0.54	187,223 (89)	79,311 (94)
31. tomorrow	+0.20	10,379,617 (21)	8,899,406 (28)	21. ?	-0.65	341,141 (83)	244,215 (82)
32. !	+0.19	3,463,257 (40)	1,385,072 (62)	22. ?	-0.70	2,907,145 (45)	1,891,225 (55)
33. summer	+0.13	2,988,785 (43)	2,554,459 (50)	23. flu	-0.75	901,403 (68)	639,000 (79)
34. ?	+0.13	26,122,914 (7)	25,513,287 (7)	24. min	-0.78	3,233,466 (41)	6,959,903 (28)
35. today	+0.13	25,588,506 (9)	23,619,518 (12)	25. BP	-0.78	582,167 (74)	326,100 (79)
36. man	+0.12	15,826,341 (14)	29,558,118 (9)	26. mosque	-0.79	69,812 (95)	46,736 (95)
37. woman	+0.10	2,543,096 (51)	5,603,347 (39)	27. dark	-0.85	1,577,553 (61)	3,233,911 (47)
38. Stephen Colbert	+0.10	23,778 (99)	14,697 (99)	28. Lehman Brothers	-1.08	5,500 (100)	4,280 (100)
39. :)	+0.10	943,413 (66)	516,171 (73)	29. Goldman Sachs	-1.08	52,703 (96)	30,769 (97)
40. ?	+0.09	39,655,724 (1)	142,219,359 (1)	30. Afghanistan	-1.13	673,519 (83)	172,697 (97)
41. coffee	+0.04	2,800,972 (46)	2,399,867 (52)	91. Iraq	-1.37	238,931 (85)	213,425 (84)
42. church	+0.03	1,812,253 (58)	3,452,171 (45)	92. cold	-1.39	3,670,447 (36)	7,015,518 (32)
43. work	+0.02	18,415,618 (11)	16,191,892 (18)	93. gun	-1.81	680,903 (72)	1,263,217 (63)
44. !	+0.02	90,960,343 (2)	282,865,043 (1)	94. hate	-2.43	9,652,881 (25)	18,158,870 (15)
45. yes	+0.02	11,593,356 (19)	7,499,840 (30)	95. hell	-2.49	6,266,102 (30)	11,056,735 (24)
46. them	+0.01	15,352,295 (13)	14,838,889 (22)	96. sick	-2.53	3,376,068 (37)	6,783,965 (36)
47. hot	-0.01	7,122,144 (28)	6,286,163 (37)	97. sad	-2.56	3,563,745 (38)	6,951,686 (33)
48. boy	-0.01	4,933,333 (33)	9,670,512 (27)	98. war	-2.63	1,955,991 (57)	3,417,388 (46)
49. yesterday	-0.01	3,077,761 (42)	2,852,023 (49)	99. depressed	-2.64	280,872 (82)	541,394 (72)
50. Michael Jackson	-0.02	825,979 (70)	371,442 (71)	100. headache	-2.83	256,600 (69)	1,146,064 (16)

Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

Positivity Bias

References

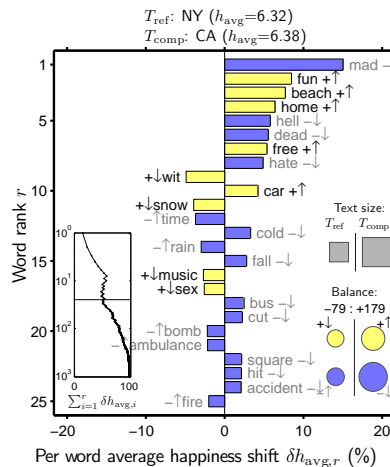


63 of 83

# Twitter—location:

Text element and context correlate in happiness scores:

- ▶ Compare ambient happiness with text element happiness.
- ▶ Spearman correlation coefficient:  $r_s \approx 0.79$ ,  $p$ -value  $< 10^{-10}$ .
- ▶ An on-average result: says nothing about any individual sentence.
- ▶ Extra random piece: stemming fails.



Complex Sociotechnical Systems

Measuring Happiness

Some motivation  
Measuring emotional content  
Data sets

Analysis

Songs  
Blogs  
Tweets

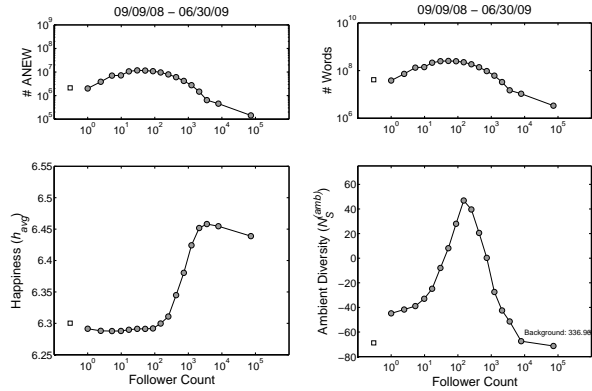
Positivity Bias

References



66 of 83

# Twitter—popularity based on follower count:



► Dunbar's number  $\approx 150$ .

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets

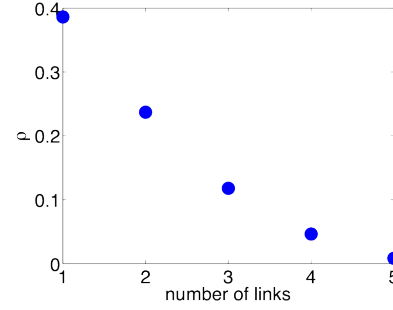
Analysis  
Songs  
Blogs  
Tweets

Positivity Bias

References



# Twitter—interactions:



- Decay in happiness correlation in social network.
- $\rho$  = Spearman's correlation coefficient.

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets

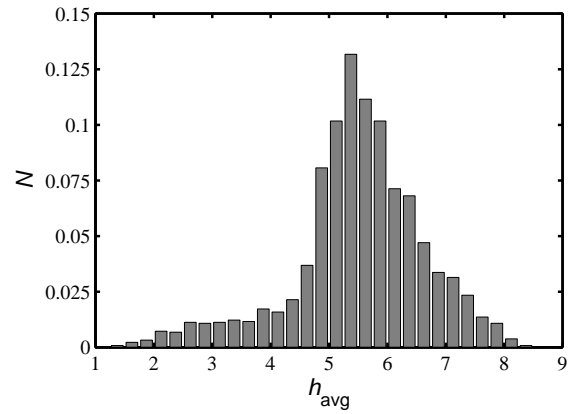
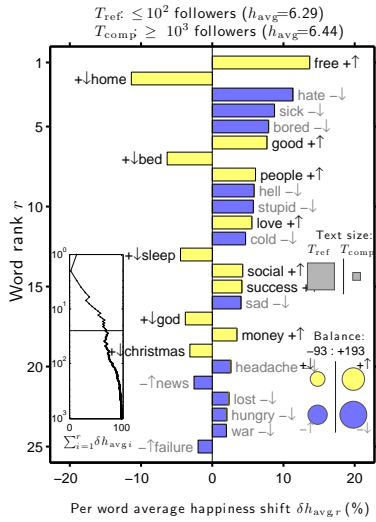
Analysis  
Songs  
Blogs  
Tweets

Positivity Bias

References



# Positive bias in the English language:



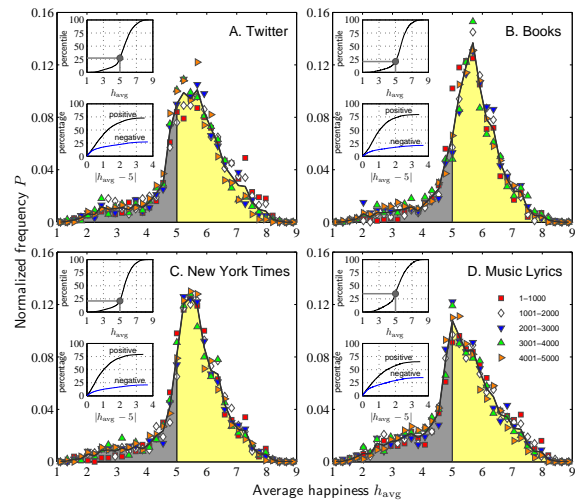
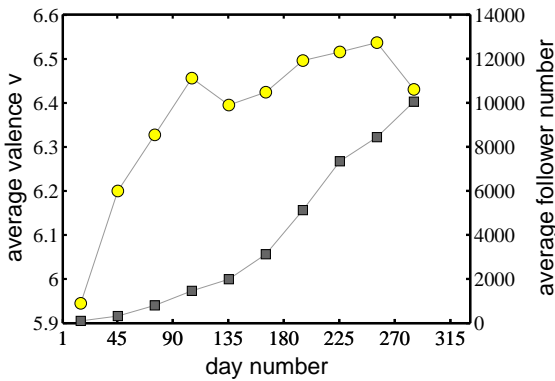
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets

Analysis  
Songs  
Blogs  
Tweets

Positivity Bias

References



Complex Sociotechnical Systems

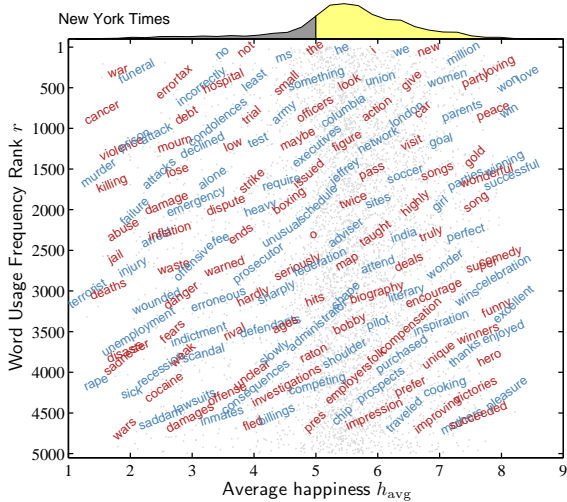
Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets

Analysis  
Songs  
Blogs  
Tweets

Positivity Bias

References





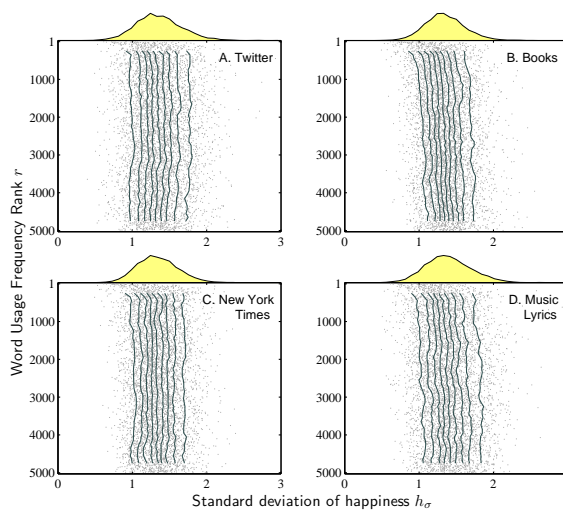
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets

Positivity Bias  
References



73 of 83



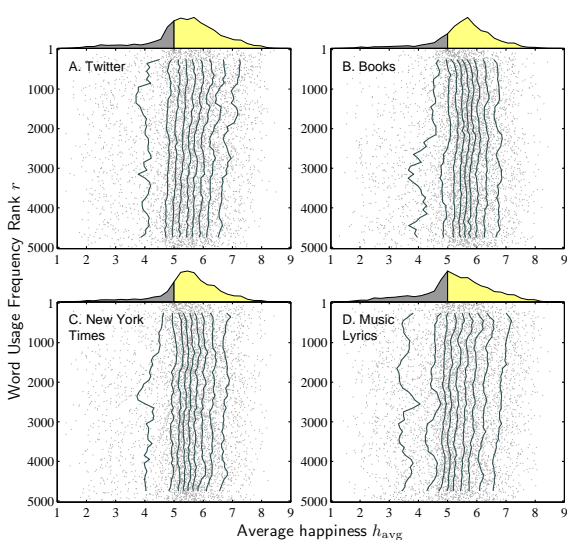
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets

Positivity Bias  
References



76 of 83



Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets

Positivity Bias  
References



74 of 83

Random other things (now and next):

- ▶ Gross National Happiness Index, [hedonometer.org](http://hedonometer.org) (in development)
- ▶ Prediction . . .
- ▶ Scores for letters, phonemes, as a function of tense.
- ▶ Fifteen additional languages being scored on Mechanical Turk
- ▶ How does happiness vary with proximity to nature? to Walmart?
- ▶ Emotional contagion.
- ▶ Quantifying metaphor and narrative and stories . . .

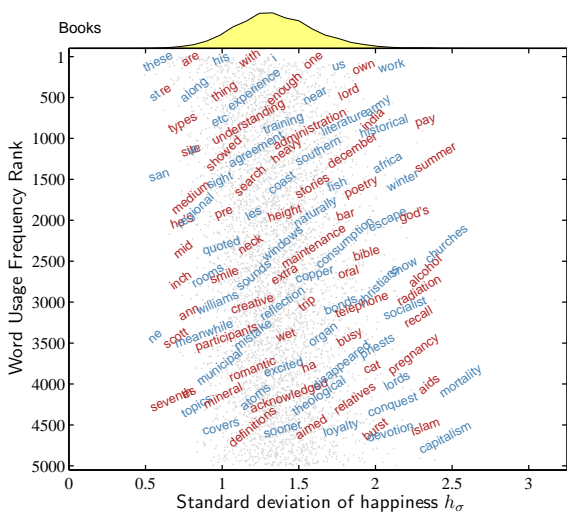
Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets

Positivity Bias  
References



77 of 83



Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets

Positivity Bias  
References



75 of 83

References I

[1] M. Bradley and P. Lang. Affective norms for english words (anew): Stimuli, instruction manual and affective ratings. Technical report c-1, University of Florida, Gainesville, FL, 1999. pdf (E)

[2] T. Conner Christensen, L. Feldman Barrett, E. Bliss-Moreau, K. Lebo, and C. Kaschub. A practical guide to experience-sampling procedures. *Journal of Happiness Studies*, 4:53–78, 2003.

[3] M. Csikszentmihalyi. *Flow*. Harper & Row, New York, 1990.

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets

Positivity Bias  
References



78 of 83

## References II

- [4] M. Csikszentmihalyi, R. Larson, and S. Prescott. The ecology of adolescent activity and experience. [\*Journal of Youth and Adolescence\*, 6:281–294, 1977.](#)
- [5] E. Diener and M. Y. Chan. Happy people live longer: Subjective well-being contributes to health and longevity. [\*Applied Psychology: Health and Well-Being\*, 3:1–43, 2011.](#) pdf (田)
- [6] P. S. Dodds and C. M. Danforth. Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. [\*Journal of Happiness Studies\*, 2009.](#) doi:10.1007/s10902-009-9150-9. pdf (田)

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



79 of 83

## References V

- [13] S. Lyubomirsky. [\*The How of Happiness\*. The Penguin Press, New York, 2007.](#)
- [14] C. Martinelli and S. W. Parker. Deception and misreporting in a social program. forthcoming in [\*Journal of the European Economic Association\*, 2007.](#) pdf (田)
- [15] G. Miller. Social Scientists waded into the Tweet stream. [\*Science Magazine\*, 333:1814–1815, 2011.](#) pdf (田)
- [16] G. Mishne and M. de Rijke. Capturing global mood levels using blog posts. [\*AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs\*, 2005.](#)

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



82 of 83

## References III

- [7] P. S. Dodds, K. D. Harris, I. M. Kloumann, C. A. Bliss, and C. M. Danforth. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. [\*PLoS ONE\*, 6:e26752, 2011.](#) Draft version available at <http://arxiv.org/abs/1101.5120v4>. Accessed October 24, 2011. pdf (田)
- [8] W. T. Jones. [\*The Classical Mind\*. Harcourt, Brace, Jovanovich, New York, 1970.](#)
- [9] D. Kahneman, A. B. Krueger, D. A. Schkade, N. Schwarz, and A. A. Stone. A survey method for characterizing daily life experience: The day reconstruction method. [\*Science\*, 306\(5702\):1776–1780, 2004.](#) pdf (田)

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



80 of 83

## References VI

- [17] E. Sandhaus. [\*The New York Times Annotated Corpus\*. Linguistic Data Consortium, Philadelphia, 2008.](#)

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



83 of 83

## References IV

- [10] M. A. Killingsworth and D. T. Gilbert. A wandering mind is an unhappy mind. [\*Science Magazine\*, 330:932, 2010.](#) pdf (田)
- [11] I. M. Kloumann, C. M. Danforth, K. D. Harris, C. A. Bliss, and P. S. Dodds. Positivity of the English language. [\*PLoS ONE\*, 7:e29484, 2012.](#) Draft version available at <http://arxiv.org/abs/1108.5192>. Accessed October 24, 2011. pdf (田)
- [12] R. Layard. [\*Happiness\*. The Penguin Press, London, 2005.](#)

Complex Sociotechnical Systems

Measuring Happiness  
Some motivation  
Measuring emotional content  
Data sets  
Analysis  
Songs  
Blogs  
Tweets  
Positivity Bias  
References



81 of 83