# Lecture Two

## 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








# Outline





# Papers and so on:





Login | Create Account | Feedback

PLoS one
accelerating the publication of peer-reviewed science

Search articles... | Advanced Search | Browse | RSS

Home | Browse Articles | About | For Readers | For Authors and Reviewers | Journals | Hubs | PLoS.org

RESEARCH ARTICLE

**Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter**

Article | Metrics | Related Content | Comments: 4

Peter Sheridan Dodds[1,2,3*], Kameron Decker Harris[1], Isabel M. Kloumann[1,4], Catherine A. Bliss[1], Christopher M. Danforth[1,2,3*]

1 Department of Mathematics and Statistics, University of Vermont, Burlington, Vermont, United States of America, 2 Center for Complex Systems, University of Vermont, Burlington, Vermont, United States of America, 3 Vermont Advanced Computing Center, University of Vermont, Burlington, Vermont, United States of America, 4 Department of Physics, University of Vermont, Burlington, Vermont, United States of America

To add a note, highlight some text. Hide notes
Make a general comment

Jump to
Abstract
Introduction
Results and Discussion
Methods
Supporting Information
Acknowledgments
Author Contributions
References

Abstract Top

Individual happiness is a fundamental societal metric. Normally measured through self-report, happiness has often been indirectly characterized and overshadowed by more readily quantifiable economic indicators such as gross domestic product. Here, we examine expressions made on the online, global microblog and social networking service Twitter, uncovering and explaining temporal variations in happiness and information levels over timescales ranging from hours to years. Our data set comprises over 46 billion words contained in nearly 4.6 billion expressions posted over a 33 month span by over 63 million unique users. In measuring happiness, we construct a tunable, real-time, remote-sensing, and non-invasive, text-based hedonometer. In building our metric, made available with this paper, we conducted a survey to obtain happiness evaluations of over 10,000 individual words, representing a tenfold size improvement over similar existing word sets. Rather than being ad hoc, our word list is chosen solely by frequency of usage, and we show how a highly robust and tunable metric can be constructed and defended.

Download: PDF | Citation | XML
Print article
EzReprint New & improved!

Metrics
Total Article Views: 13,921
Cited in: CrossRef (0), PubMed Central (0), Scopus (1), Web of Science® (0)
Social Networks: CiteULike (1), Facebook (93)
More

Related Content
Related Subject Categories: Computer Science, Physics
Related Articles on the Web: Google Scholar, PubMed
More

Share this Article
Email this article

"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 (⊞)

- "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 (⊞)

# Press...



- “Social Scientists wade 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. (⊞)

- More here: http://www.uvm.edu/∼pdodds/research/ (⊞)

# Outline

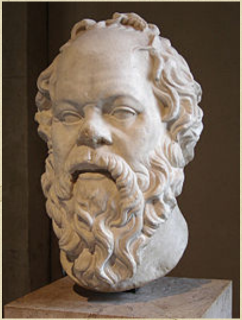




# Happiness:







Socrates et al.:
eudaimonia [8]

Bentham:
hedonistic calculus

Jefferson:
. . . the pursuit of happiness

# Early drafts:

that among these are:

Life, ✓

Liberty, ✓ and ?? ~~Money~~?

~~Libations~~ ... ~~Alcohol~~

~~Property~~ ~~Foot-the-ball~~

~~Beer~~

Happiness ✓✓✓

## Desiring happiness—not just for boffins:

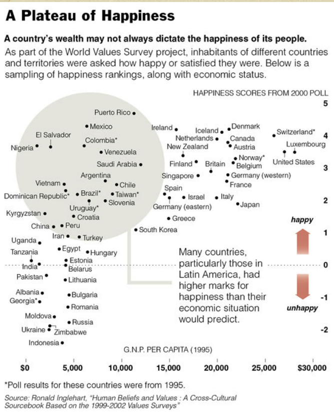
- Average people routinely report being happy is what they want most in life [12, 13, 5]
- And it matters: "Happy people live longer:. . . " Survey by Diener and Chan. [5]

## 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]

## 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]




## National indices of well-being:

- Bhutan
- France
- Australia

# Outline

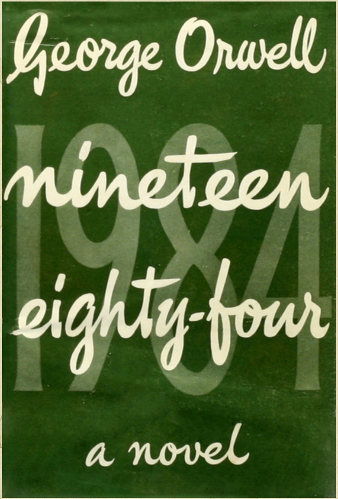
Measuring Happiness
- Some motivation
- Measuring emotional content
- Data sets

Analysis
- Songs
- Blogs
- Tweets

Positivity Bias

References

# An easy knock:

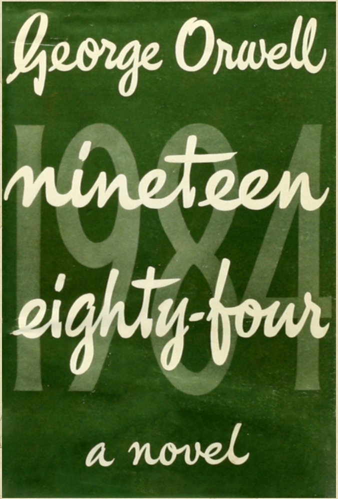



Science = Orwell

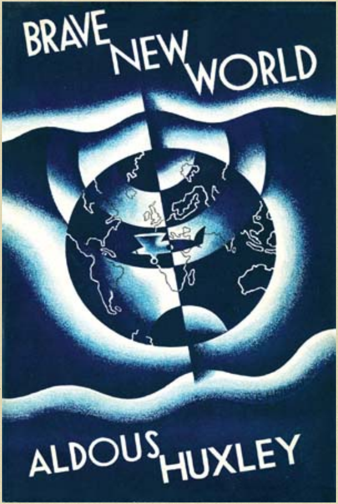
# An easy knock:



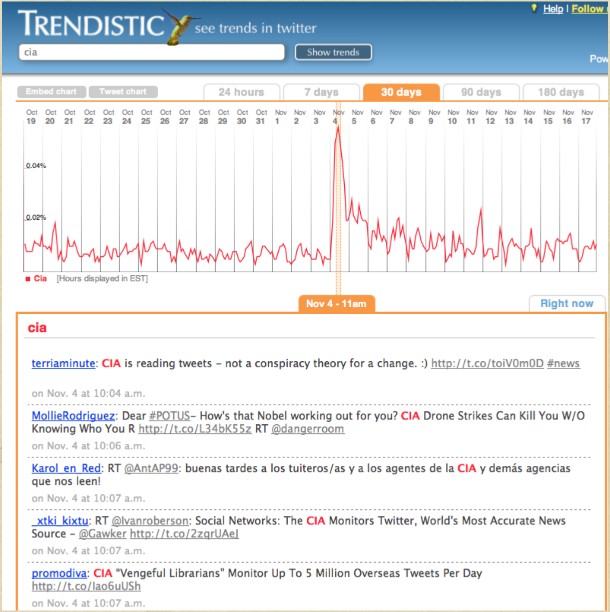





Science = Orwell




Policy = Brave New World




- See story here (⊞) for example [slate].

# 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

# 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

# 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

# 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

# Happiness, attention, and doing:

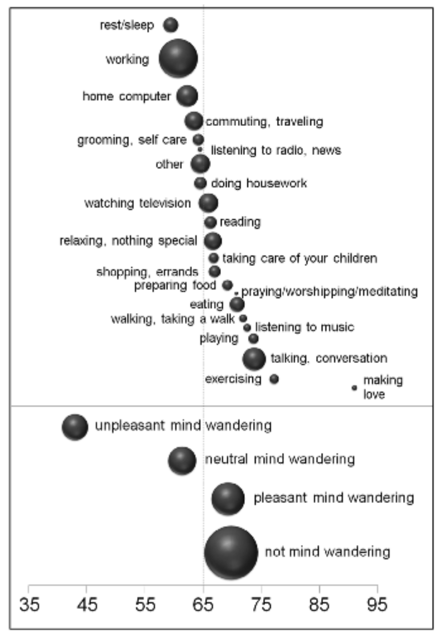
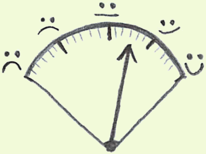



**Fig. 1.** Mean happiness reported during each activity (**top**) and while mind wandering to unpleasant topics, neutral topics, pleasant topics or not mind wandering (**bottom**). Dashed line indicates mean of happiness across all samples. Bubble area indicates the frequency of occurrence. The largest 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]

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



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

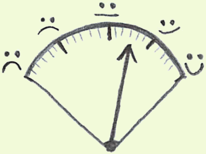
## Ideally:

- 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]

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



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

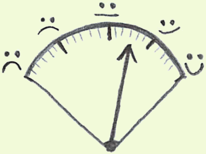
## Ideally:

- 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]

## 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]

# ANEW study

- **ANEW** = “Affective Norms for English Words”
- Study: participants shown lists of isolated words
- Asked to grade each word’s valence, arousal, and dominance level
- Integer scale of 1–9
- $N$ =1034 words—previously identified as bearing emotional weight
- Participants = College students (*cough*)
- Results published by Bradley and Lang (1999) [1]

# ANEW study

- ANEW = "Affective Norms for English Words"
- Study: participants shown lists of isolated words
- Asked to grade each word's valence, arousal, and dominance level
- Integer scale of 1–9
- $N$ =1034 words—previously identified as bearing emotional weight
- Participants = College students (*cough*)
- Results published by Bradley and Lang (1999) [1]





# ANEW study

- ANEW = “Affective Norms for English Words”
- Study: participants shown lists of isolated words
- Asked to grade each word’s valence, arousal, and dominance level
- Integer scale of 1–9
- $N$ =1034 words—previously identified as bearing emotional weight
- Participants = College students (*cough*)
- Results published by Bradley and Lang (1999) [1]

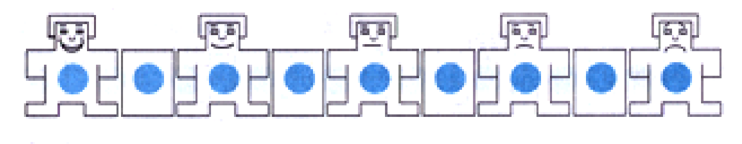




# ANEW study—three 1–9 scales:



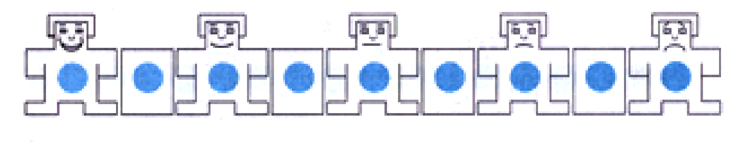
valence:



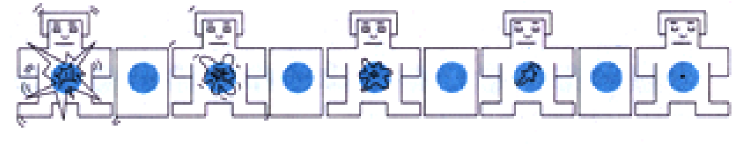
# ANEW study—three 1–9 scales:



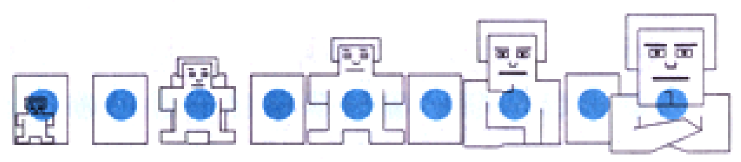
valence:



arousal:



dominance:



# ANEW study words—examples


love/paradise/triumphant
glory/luxury/trophy
optimism/pancakes/church
engine/paper/street
derelict/neurotic/vanity
fault/corrupt/lawsuit
trauma/hostage/disgusted
funeral/rape/suicide

ANEW = "Affective Norms for English Words" [1]

# Analysing text:




## Lyrics for Michael Jackson's Billie Jean

"She was more like a beauty queen
from a movie scene.
⋮
And mother always told me,
be careful who you love.
And be careful of what you do
'cause the lie becomes the truth.

Billie Jean is not my lover,
She's just a girl who claims
that I am the one.
⋮

| ANEW words | $v_k$ | $f_k$ |
|---|---|---|
| $k$=1. love | 8.72 | 1 |
| 2. mother | 8.39 | 1 |
| 3. baby | 8.22 | 3 |
| 4. beauty | 7.82 | 1 |
| 5. truth | 7.80 | 1 |
| 6. people | 7.33 | 2 |
| 7. strong | 7.11 | 1 |
| 8. young | 6.89 | 2 |
| 9. girl | 6.87 | 4 |
| 10. movie | 6.86 | 1 |
| 11. perfume | 6.76 | 1 |
| 12. queen | 6.44 | 1 |
| 13. name | 5.55 | 1 |
| 14. lie | 2.79 | 1 |

$$v_{\text{text}} = \frac{\sum_k v_k f_k}{\sum_k f_k}$$

$$v_{\text{Billie Jean}} = 7.1$$

$$v_{\text{Thriller}} = 6.3$$

$$v_{\text{Michael Jackson}} = 6.4$$

# Outline

# Data sets:

## Texts:

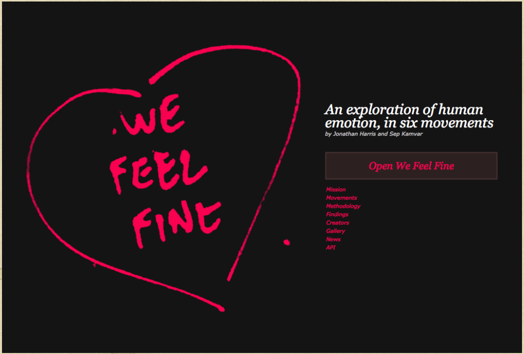
1. Song lyrics (1960–2007)
2. Song titles (1960–2008)
3. State of the Union (SOTU) Addresses (1790–2008)

## Sources:

- hotlyrics.com (⊞)
- freedb.com (⊞)
- American Presidency Project: www.presidency.ucsb.edu (⊞).

# Data sets:

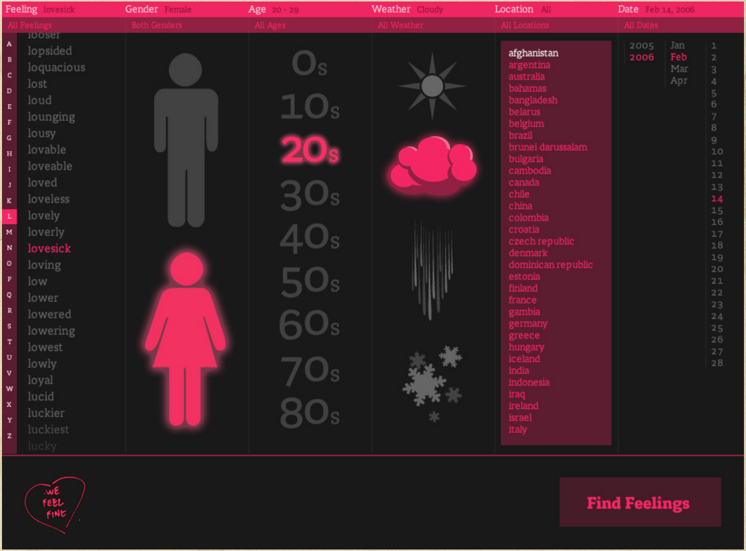
4. Blog phrases containing “I feel...”, “I am feeling”, etc., taken from wefeelfine.org (⊞) (API, 2005–2010)




- Created by Jonathan Harris & Sep Kamvar



# wefeelfine.org:




Complex Sociotechnical Systems

Measuring Happiness

Some motivation

Measuring emotional content



Analysis

Songs

Blogs

Tweets

Positivity Bias

References

The UNIVERSITY of VERMONT

22 of 83

# More data sets:

5. twitter
6. New York Times (20 years)
7. Gutenberg.org
8. Google Books: http://ngrams.googlelabs.com/ (⊞)
9. . . .





The UNIVERSITY of VERMONT

# Some numbers:





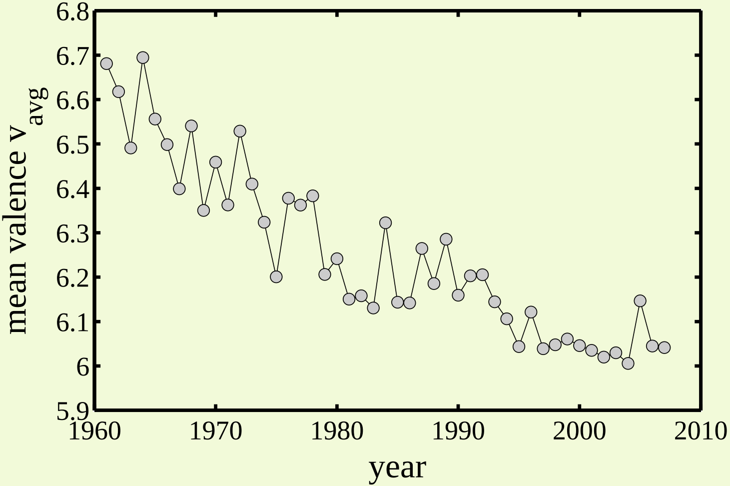
| Counts | Song lyrics | Song titles |
|---|---|---|
| All words | 58,610,849 | 60,867,223 |
| ANEW words | 3,477,575 (5.9%) | 5,612,708 (9.2%) |
| Individuals | $\sim$ 20,000 | $\sim$ 632,000 |

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

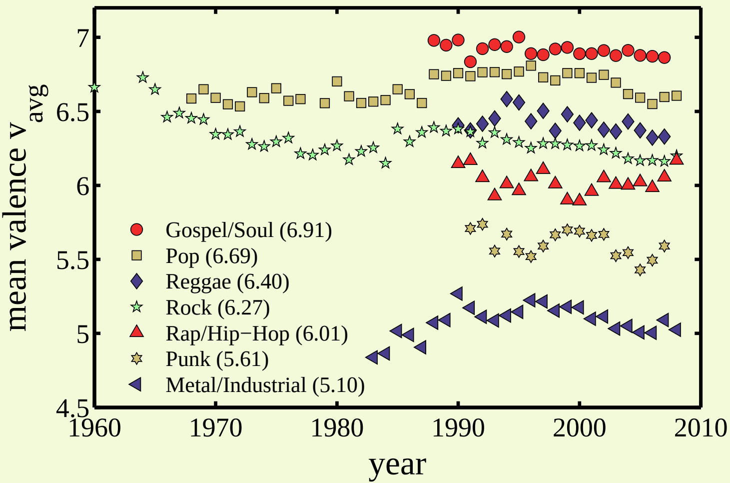
| Counts | Twitter |
|---|---|
| All words | $\sim$ 100 billion |
| Tweets | $\sim$ 10 billion |
| Individuals | $\sim$ 100 million |

# Outline

# Song Lyrics—average happiness (valence)



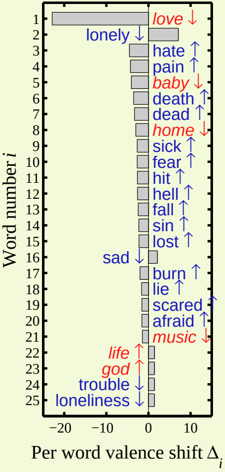
# Song Lyrics—average happiness of genres:

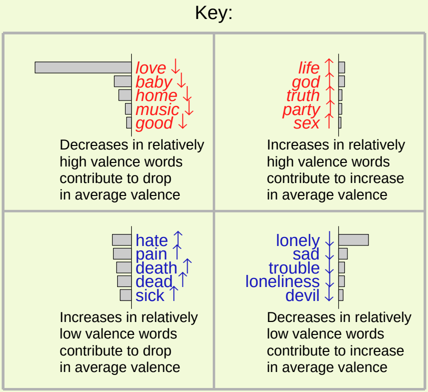


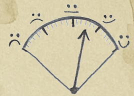
# Happiness Word Shift Graph:



Per word drop in valence of lyrics from 1980−2007 relative to valence of lyrics from 1960



# Top 16 of $\simeq$ 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)

# Bottom 16 of $\simeq$ 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)

# Outline

Measuring Happiness

- Some motivation
- Measuring emotional content
- Data sets

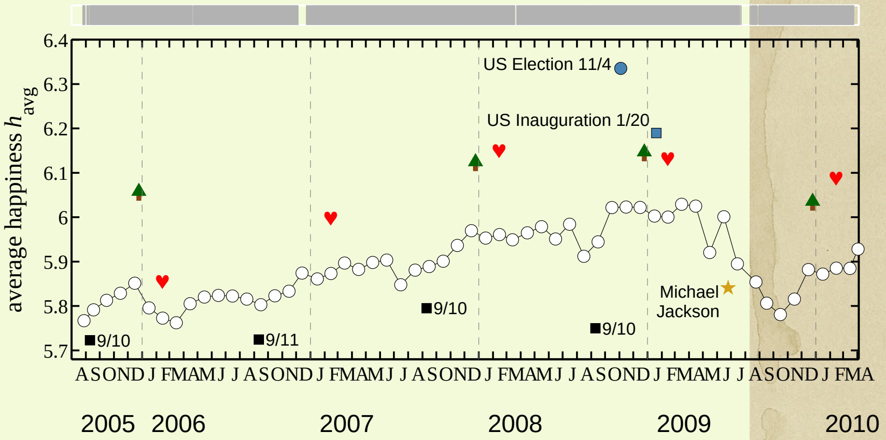
Analysis

- Songs
- Blogs
- Tweets

Positivity Bias

References

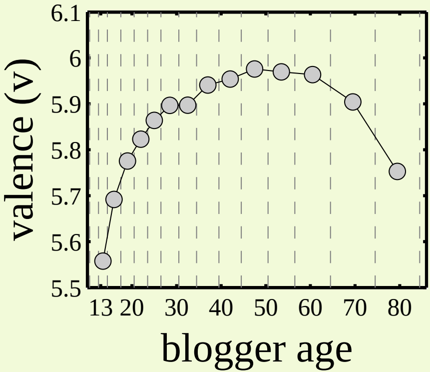
# Blogs—Overall trend



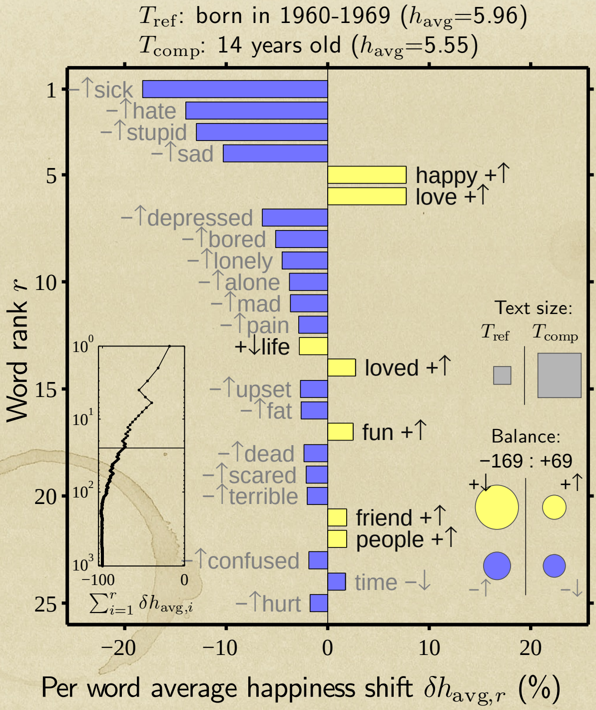
| Text: | $h_{avg}$ | Words with a similar score: |
|---|---|---|
| Soul/Gospel lyrics [6] | 6.9 | chocolate (6.88), leisurely (6.88), penthouse (6.81) |
| Pop lyrics [6] | 6.7 | dream (6.73), honey (6.73), sugar (6.74) |
| Dante's Paradise | 6.5 | muffin (6.57), rabbit (6.57), smooth (6.58) |
| Tweets, 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 (⊞) | 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) |

# Blogs

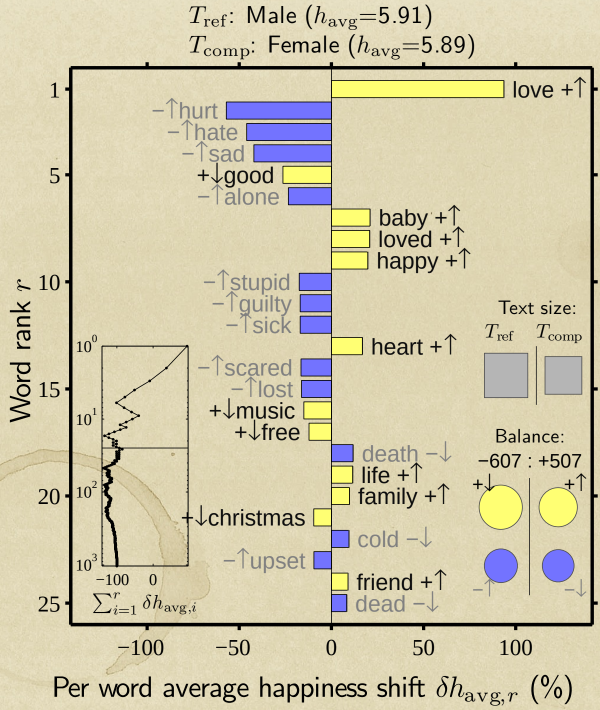





- Average happiness as a function of the age bloggers report they will turn in the year of their posting.

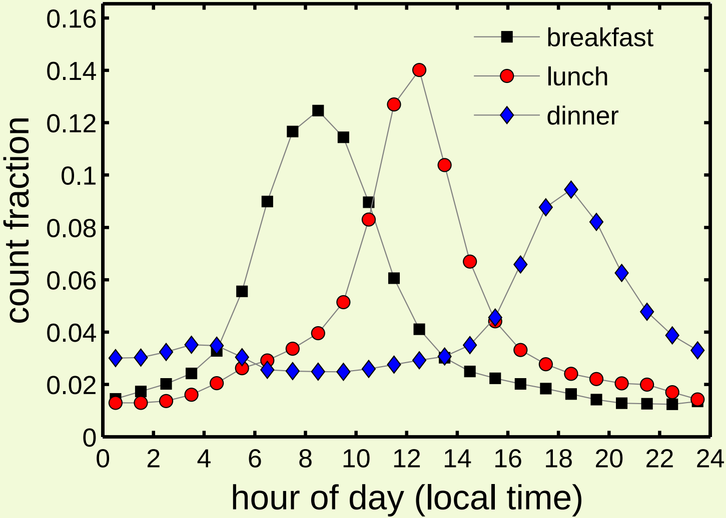


$T_{\rm ref}$: born in 1960-1969 ($h_{\rm avg}$=5.96)
$T_{\rm comp}$: 14 years old ($h_{\rm avg}$=5.55)
−↑sick
−↑hate
−↑stupid
−↑sad
happy +↑
love +↑
−↑depressed
−↑bored
−↑lonely
−↑alone
−↑mad
−↑pain
+↓life
loved +↑
−↑upset
−↑fat
fun +↑
−↑dead
−↑scared
−↑terrible
friend +↑
people +↑
−↑confused
time −↓
−↑hurt
Text size:
$T_{\rm ref}$
$T_{\rm comp}$
Balance:
−169 : +69
+↓
+↑
−↑
−↓
Word rank $r$
1
5
10
15
20
25
$10^0$
$10^1$
$10^2$
$10^3$
−100
0
$\sum_{i=1}^{r} \delta h_{{\rm avg},i}$
−20
−10
0
10
20
Per word average happiness shift $\delta h_{{\rm avg},r}$ (%)

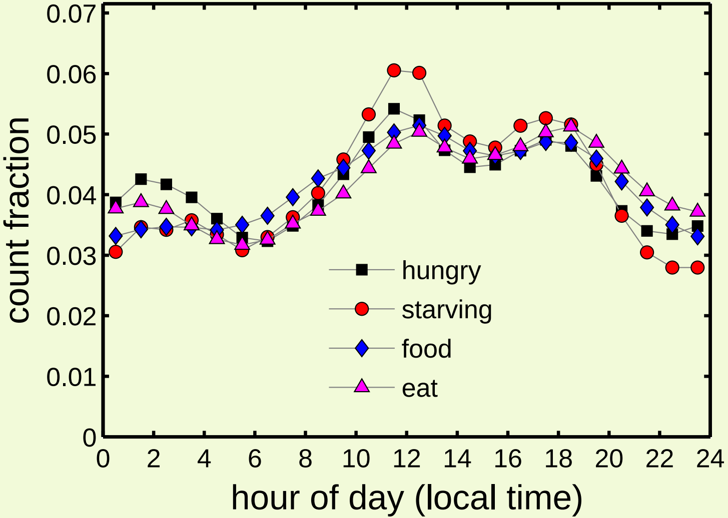

$T_{\text{ref}}$: Male ($h_{\text{avg}}$=5.91)
$T_{\text{comp}}$: Female ($h_{\text{avg}}$=5.89)
love +↑
−↑hurt
−↑hate
−↑sad
+↓good
−↑alone
baby +↑
loved +↑
happy +↑
−↑stupid
−↑guilty
−↑sick
heart +↑
−↑scared
−↑lost
+↓music
+↓free
death −↓
life +↑
family +↑
+↓christmas
cold −↓
−↑upset
friend +↑
dead −↓
Text size:
$T_{\text{ref}}$ $T_{\text{comp}}$
Balance:
−607 : +507
+↓ +↑
−↑ −↓
Word rank $r$
1
5
10
15
20
25
−100
−50
0
50
100
Per word average happiness shift $\delta h_{\text{avg},r}$ (%)
$10^0$
$10^1$
$10^2$
$10^3$
−100 0
$\sum_{i=1}^{r} \delta h_{\text{avg},i}$

# Outline

# Twitter—living in the now:



# Twitter—living in the now:



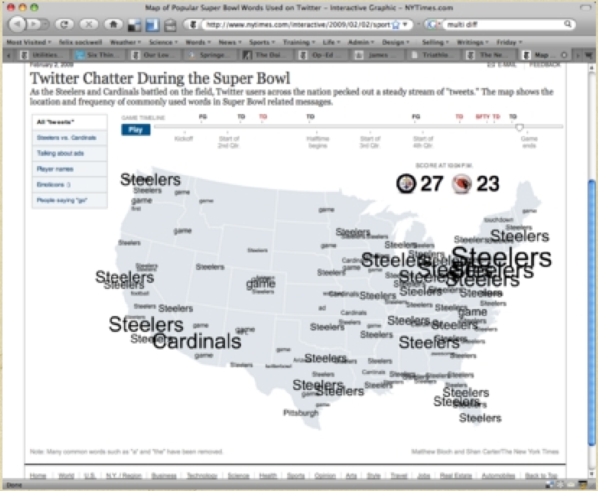
# Words most correlated with obesity levels in cities:

| 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 |

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

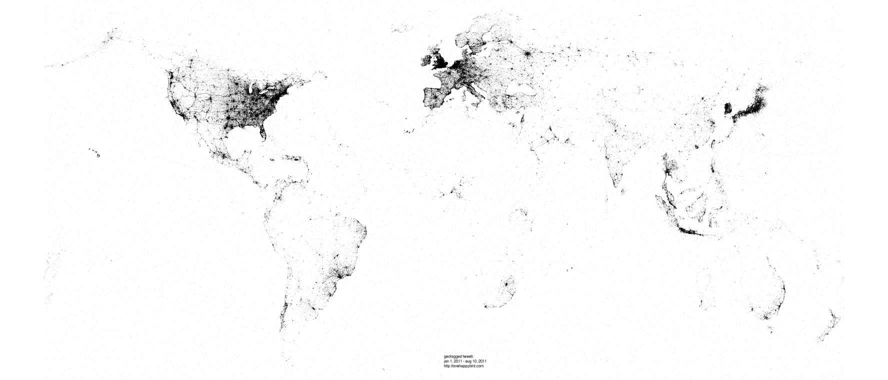
| | | | |
|---|---|---|---|
| brunch | 6.32 | -0.41 | 6.37431e-09 |
| bar | 5.82 | -0.35 | 5.54374e-07 |
| banana | 6.86 | -0.35 | 5.67492e-07 |
| barista | 0.00 | -0.35 | 7.29324e-07 |
| delicious | 7.92 | -0.34 | 1.09807e-06 |
| dinner | 7.40 | -0.34 | 1.35413e-06 |
| coffee | 7.18 | -0.34 | 2.04145e-06 |
| 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 |

# Twitter—living in the now:

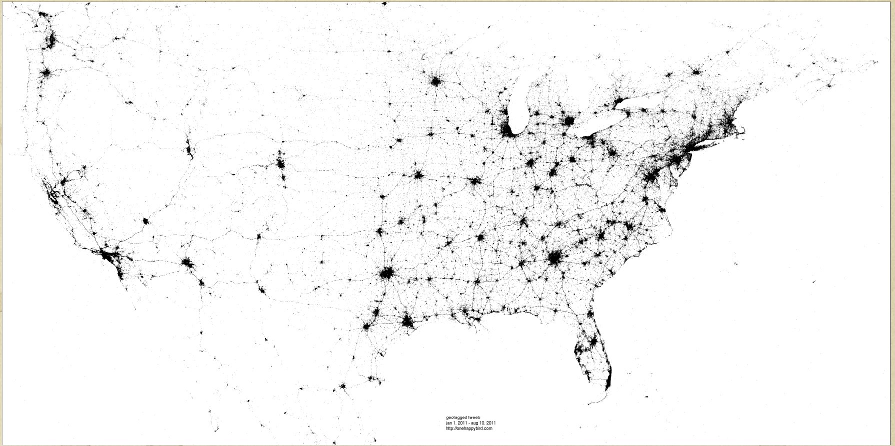



Tweeting the Superbowl (⊞) [NY Times]

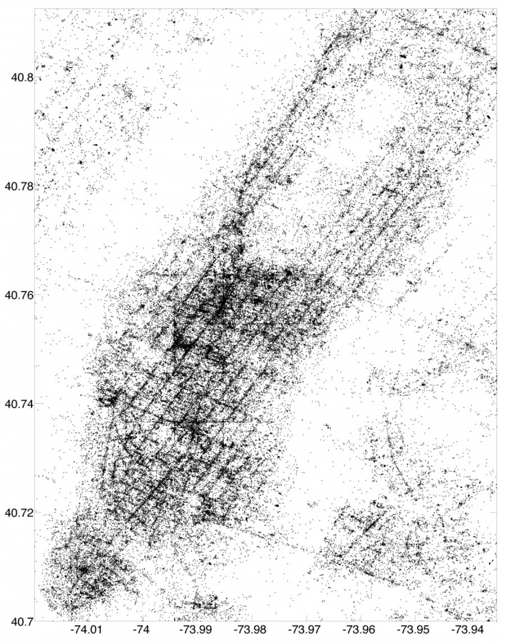
{"in_reply_to_user_id_str":null,"id_str":"122740606790410240","in_reply_to_user_id":null,"text":"RT @outrage_molly: #occ
upydenver we will be marching down E Colfax first with a shout out to @occupythehood","created_at":"Sat Oct 08 18:30:15
+0000 2011","contributors":null,"retweeted_status":{"in_reply_to_user_id_str":null,"id_str":"122739523930169344","in_rep
ly_to_user_id":null,"text":"#occupydenver we will be marching down E Colfax first with a shout out to @occupythehood","c
reated_at":"Sat Oct 08 18:25:57 +0000 2011","contributors":null,"geo":null,"favorited":false,"source":"\u003Ca href=\"ht
tp:\/\/ubersocial.com\" rel=\"nofollow\"\u003EUberSocial for BlackBerry\u003C\/a\u003E","retweet_count":1,"in_reply_to_s
creen_name":null,"coordinates":null,"entities":{"hashtags":[{"text":"occupydenver","indices":[0,13]}],"urls":[],"user_me
ntions":[{"id_str":"379936422","indices":[74,88],"name":"Occupy  The Hood ","id":379936422,"screen_name":"OccupyTheHood"
}]},"retweeted":false,"in_reply_to_status_id":null,"in_reply_to_status_id_str":null,"place":null,"user":{"id_str":"16538
6225","show_all_inline_media":false,"contributors_enabled":false,"following":null,"profile_background_image_url_https":"
https:\/\/si0.twimg.com\/profile_background_images\/123611013\/web_graphic2-1.jpg","created_at":"Sun Jul 11 12:17:25 +00
00 2010","profile_background_color":"d9571f","profile_image_url":"http:\/\/a0.twimg.com\/profile_images\/1069877121\/mol
lykate3_normal.jpg","profile_background_tile":false,"favourites_count":24,"follow_request_sent":null,"time_zone":"Mounta
in Time (US & Canada)","profile_sidebar_fill_color":"d9a4db","url":null,"description":"May the road rise with you. Anger
 is an energy. Rise. ~PiL\r | Public Policy grad student. Views are mine. Go 99%. ","default_profile":false,"geo_enabled
":false,"profile_sidebar_border_color":"52014a","followers_count":462,"is_translator":false,"profile_image_url_https":"h
ttps:\/\/si0.twimg.com\/profile_images\/1069877121\/mollykate3_normal.jpg","listed_count":6,"profile_use_background_imag
e":true,"friends_count":622,"location":"Denver","profile_text_color":"333333","protected":false,"lang":"en","verified":f
alse,"profile_background_image_url":"http:\/\/a2.twimg.com\/profile_background_images\/123611013\/web_graphic2-1.jpg","n
ame":"Molly","notifications":null,"profile_link_color":"507d4b","id":165386225,"default_profile_image":false,"statuses_c
ount":2399,"utc_offset":-25200,"screen_name":"outrage_molly"},"id":122739523930169344,"truncated":false},"geo":null,"fav
orited":false,"source":"web","retweet_count":1,"in_reply_to_screen_name":null,"coordinates":null,"entities":{"hashtags":
[{"text":"occupydenver","indices":[19,32]}],"urls":[],"user_mentions":[{"id_str":"165386225","indices":[3,17],"name":"Mo
lly","id":165386225,"screen_name":"outrage_molly"},{"id_str":"379936422","indices":[93,107],"name":"Occupy  The Hood ","
id":379936422,"screen_name":"OccupyTheHood"}]},"retweeted":false,"in_reply_to_status_id":null,"in_reply_to_status_id_str
":null,"place":null,"user":{"id_str":"362705928","show_all_inline_media":false,"contributors_enabled":false,"following":
null,"profile_background_image_url_https":"https:\/\/si0.twimg.com\/images\/themes\/theme1\/bg.png","created_at":"Fri Au
g 26 21:10:18 +0000 2011","profile_background_color":"C0DEED","profile_image_url":"http:\/\/a3.twimg.com\/profile_images
\/1514944353\/284683_780784519822_42112934_38358736_4257516_n_normal.jpg","profile_background_tile":false,"favourites_co
unt":1,"follow_request_sent":null,"time_zone":"Mountain Time (US & Canada)","profile_sidebar_fill_color":"DDEEF6","url":
null,"description":"From The Ruins Are Born The Spark","default_profile":true,"geo_enabled":false,"profile_sidebar_borde
r_color":"C0DEED","followers_count":154,"is_translator":false,"profile_image_url_https":"https:\/\/si0.twimg.com\/profil
e_images\/1514944353\/284683_780784519822_42112934_38358736_4257516_n_normal.jpg","listed_count":11,"profile_use_backgro
und_image":true,"friends_count":70,"location":"","profile_text_color":"333333","protected":false,"lang":"en","verified":
false,"profile_background_image_url":"http:\/\/a0.twimg.com\/images\/themes\/theme1\/bg.png","name":"DonRhetorick","noti
fications":null,"profile_link_color":"0084B4","id":362705928,"default_profile_image":false,"statuses_count":1975,"utc_of
fset":-25200,"screen_name":"DonRhetorick"},"id":122740606790410240,"truncated":false}




http:/www.onehappybird.com (⊞)

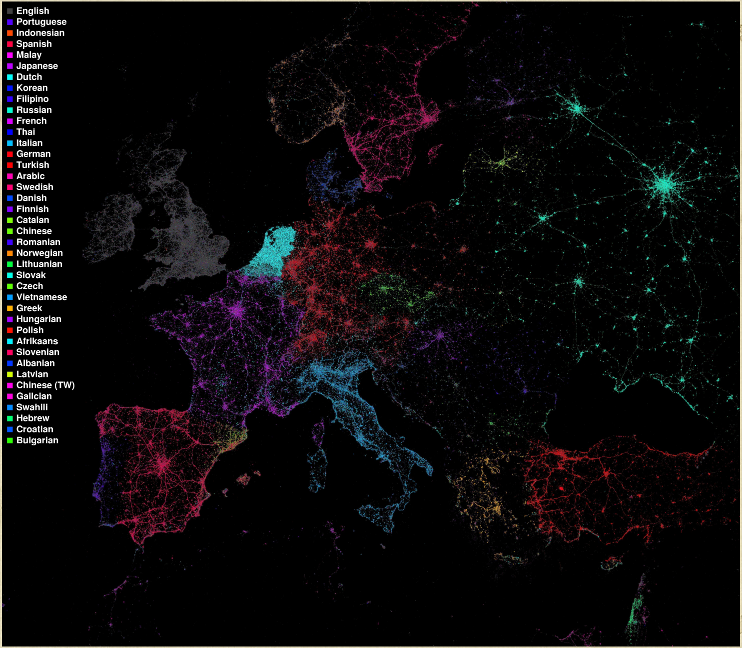



http:/www.onehappybird.com (⊞)

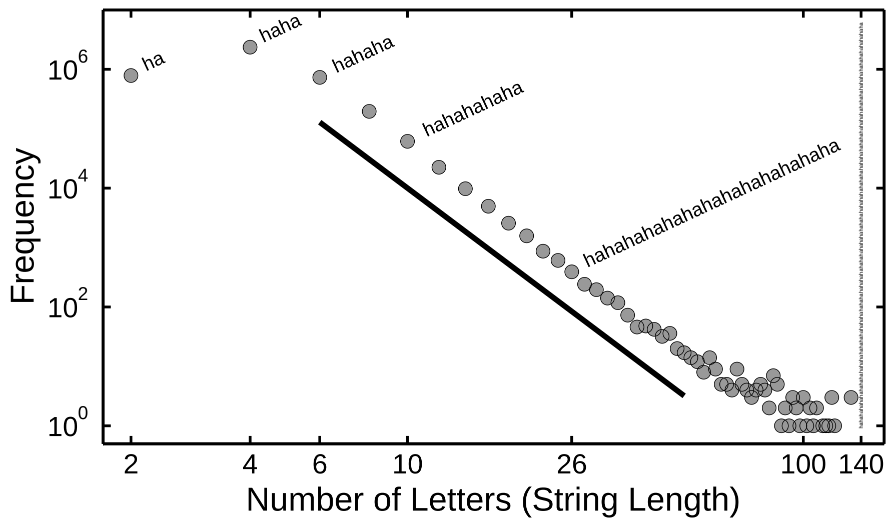



http://www.onehappybird.com (⊞)

http://flowingdata.com/2011/10/27/language-communities-of-twitter/ (田)



# The happiest distribution:



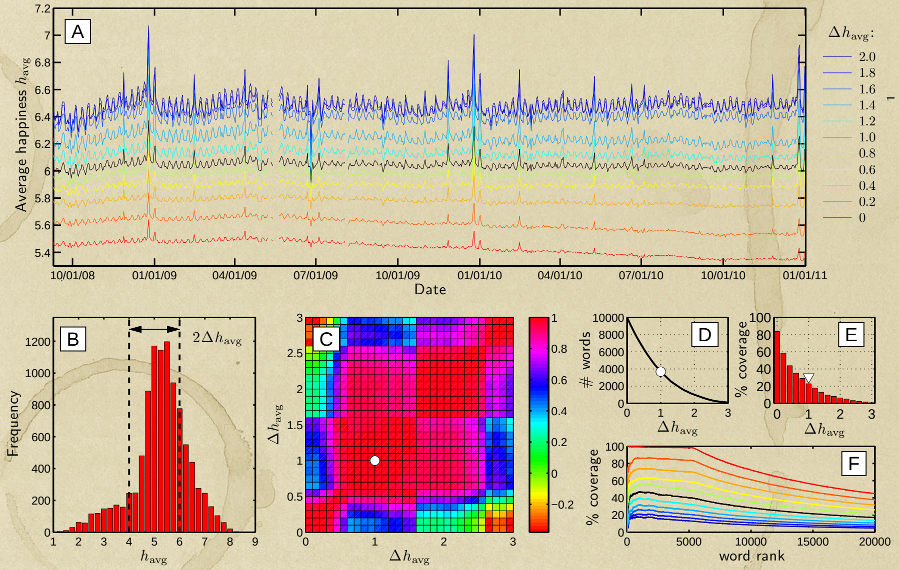
# labMT 1.0:
# language assessment by Mechanical Turk

| valence rank | word | valence | std dev | twitter rank | g-books rank | nyt rank | lyrics rank |
|---|---|---|---|---|---|---|---|
| 1 | laughter | 8.50 | 0.93 | 3600 | – | – | 1728 |
| 2 | happiness | 8.44 | 0.97 | 1853 | 2458 | – | 1230 |
| 3 | love | 8.42 | 1.11 | 25 | 317 | 328 | 23 |
| 4 | happy | 8.30 | 0.99 | 65 | 1372 | 1313 | 375 |
| 5 | laughed | 8.26 | 1.16 | 3334 | 3542 | – | 2332 |
| 6 | laugh | 8.22 | 1.37 | 1002 | 3998 | 4488 | 647 |
| 7 | laughing | 8.20 | 1.11 | 1579 | – | – | 1122 |
| 8 | excellent | 8.18 | 1.10 | 1496 | 1756 | 3155 | – |
| 9 | laughs | 8.18 | 1.16 | 3554 | – | – | 2856 |
| 10 | joy | 8.16 | 1.06 | 988 | 2336 | 2723 | 809 |
| 11 | successful | 8.16 | 1.08 | 2176 | 1198 | 1565 | – |
| 12 | win | 8.12 | 1.08 | 154 | 3031 | 776 | 694 |
| 13 | rainbow | 8.10 | 0.99 | 2726 | – | – | 1723 |
| 14 | smile | 8.10 | 1.02 | 925 | 2666 | 2898 | 349 |
| 15 | won | 8.10 | 1.22 | 810 | 1167 | 439 | 1493 |
| 16 | pleasure | 8.08 | 0.97 | 1497 | 1526 | 4253 | 1398 |
| 17 | smiled | 8.08 | 1.07 | – | 3537 | – | 2248 |
| 18 | rainbows | 8.06 | 1.36 | – | – | – | 4216 |
| 19 | winning | 8.04 | 1.05 | 1876 | – | 1426 | 3646 |
| 20 | celebration | 8.02 | 1.53 | 3306 | – | 2762 | 4070 |
| 21 | enjoyed | 8.02 | 1.53 | 1530 | 2908 | 3502 | – |
| 22 | healthy | 8.02 | 1.06 | 1393 | 3200 | 3292 | 4619 |
| 23 | music | 8.02 | 1.12 | 132 | 875 | 167 | 374 |
| 24 | celebrating | 8.00 | 1.14 | 2550 | – | – | – |
| 25 | congratulations | 8.00 | 1.63 | 2246 | – | – | – |
| 26 | weekend | 8.00 | 1.29 | 317 | – | 833 | 2256 |
| 27 | celebrate | 7.98 | 1.15 | 1606 | – | 3574 | 2108 |
| 28 | comedy | 7.98 | 1.15 | 1444 | – | 2566 | – |
| 29 | jokes | 7.98 | 0.98 | 2812 | – | – | 3808 |
| 30 | rich | 7.98 | 1.32 | 1625 | 1221 | 1469 | 890 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

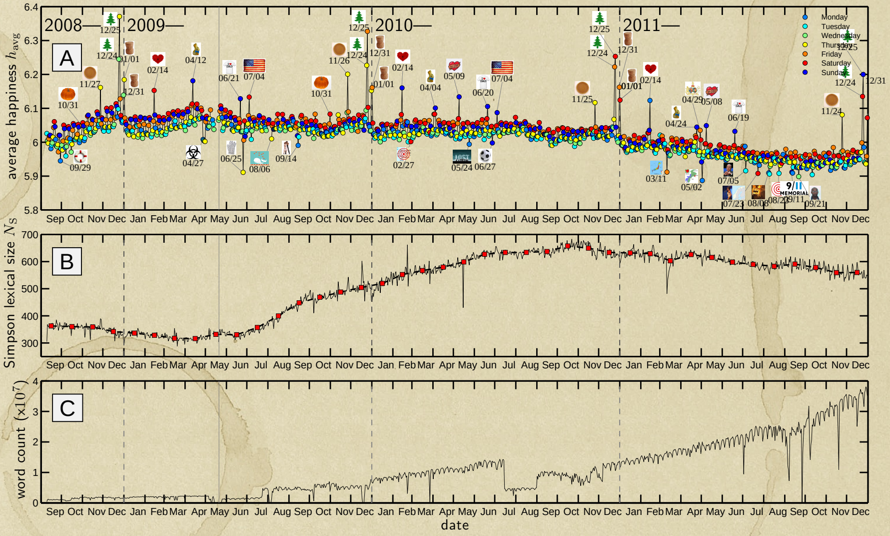
| valence rank | word | valence | std dev | twitter rank | g-books rank | nyt rank | lyrics rank |
|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 10193 | violence | 1.86 | 1.05 | 4299 | 1724 | 1238 | 2016 |
| 10194 | cruel | 1.84 | 1.15 | 2963 | – | – | 1447 |
| 10195 | cry | 1.84 | 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 | 1884 | 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 | – |

| std dev rank | word | valence | std dev | twitter rank | g-books rank | nyt rank | lyrics rank |
|---|---|---|---|---|---|---|---|
| 1 | f↯@king | 4.64 | 2.93 | 448 | – | – | 620 |
| 2 | f☆☆kin | 3.86 | 2.74 | 1077 | – | – | 688 |
| 3 | f☆☆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 | f☆☆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 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

# The very surprising tunable hedonometer:

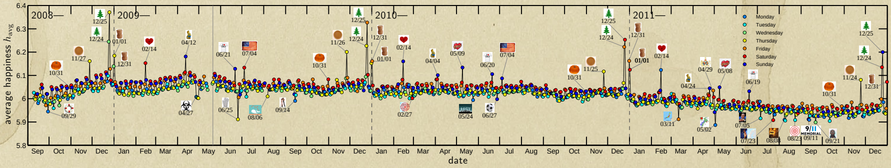


# Twitter—overall time series:



# Twitter—overall time series:

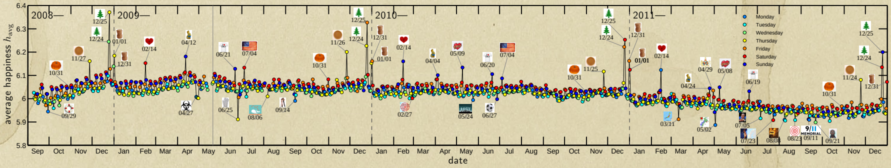
Measuring societal happiness through tweets:




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

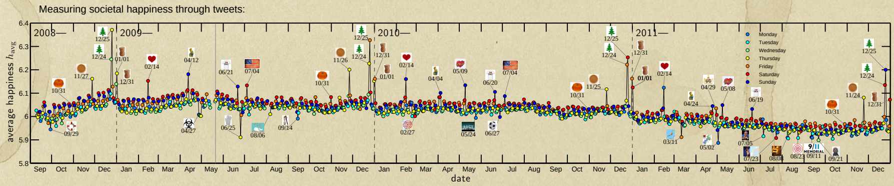
# Twitter—overall time series:

Measuring societal happiness through tweets:

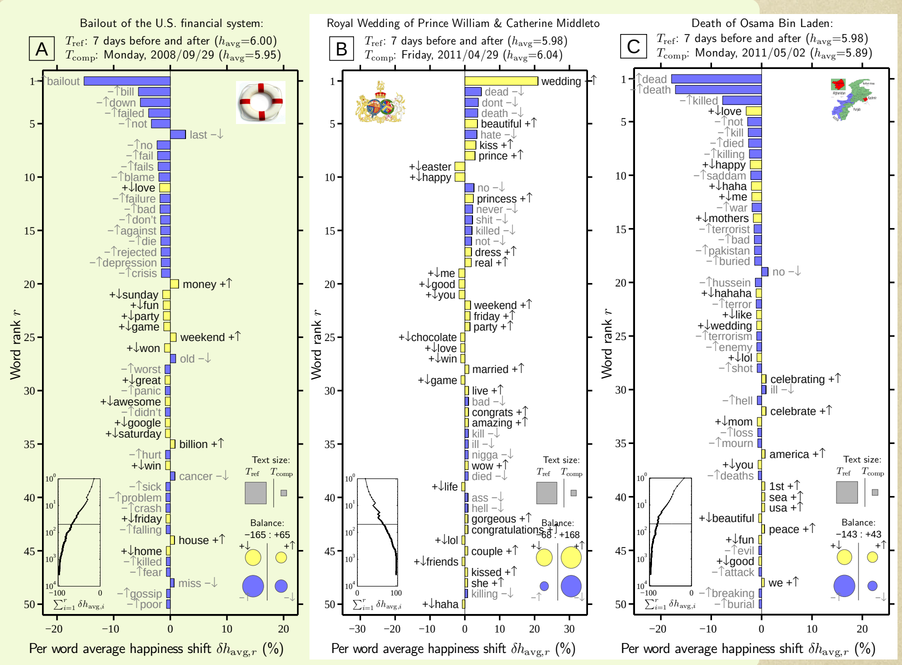



- Global happiness spikes = predictable rituals.
- Global sadness spikes = unpredictable, exogeneous shocks.
- No accidental happiness outbreaks.

# Twitter—overall time series:

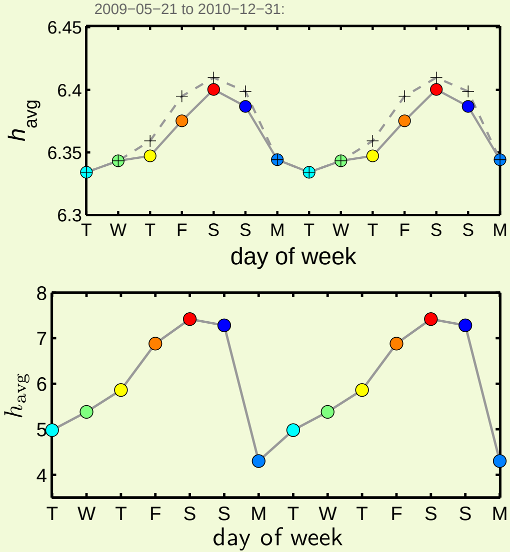


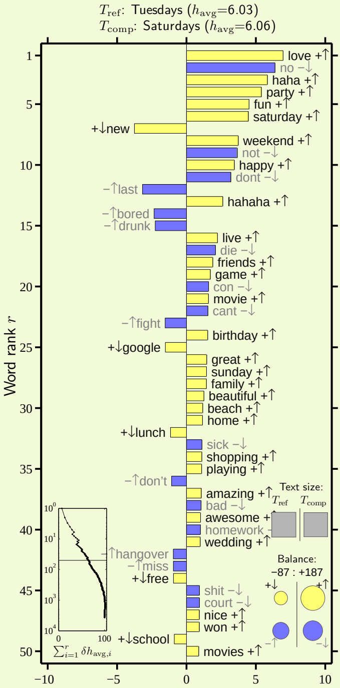

- Global happiness spikes = predictable rituals.
- Global sadness spikes = unpredictable, exogeneous shocks.
- No accidental happiness outbreaks.



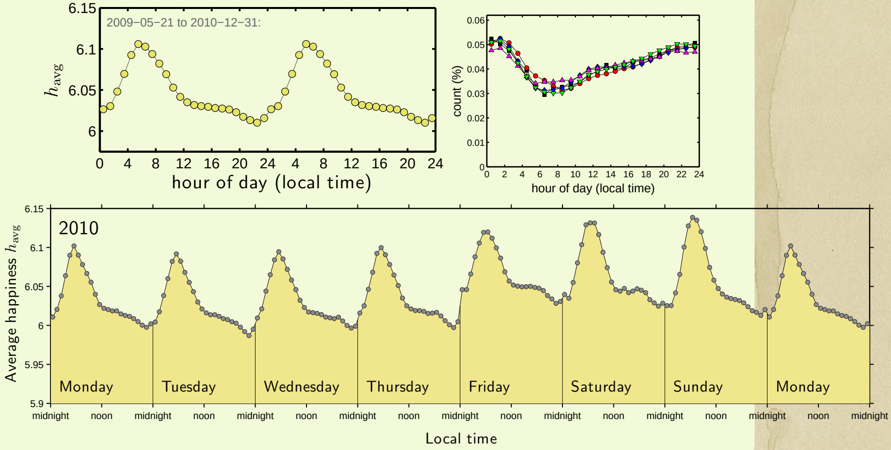
Bailout of the U.S. financial system:
A
$T_{\rm ref}$: 7 days before and after ($h_{\rm avg}$=6.00)
$T_{\rm comp}$: Monday, 2008/09/29 ($h_{\rm avg}$=5.95)
−↑bailout
−↑bill
−↑down
−↑failed
−↑not
last −↓
−↑no
−↑fail
−↑fails
−↑blame
+↓love
−↑failure
−↑bad
−↑don't
−↑against
−↑die
−↑rejected
−↑depression
−↑crisis
money +↑
+↓sunday
+↓fun
+↓party
+↓game
weekend +↑
+↓won
old −↓
−↑worst
+↓great
−↑panic
+↓awesome
−↑didn't
+↓google
+↓saturday
billion +↑
−↑hurt
+↓win
cancer −↓
−↑sick
−↑problem
−↑crash
+↓friday
−↑falling
house +↑
+↓home
−↑killed
−↑fear
miss −↓
−↑gossip
−↑poor
Text size:
$T_{\rm ref}$ $T_{\rm comp}$
Balance:
−165 : +65
+↓ +↑
−↑ −↓
$\sum_{i=1}^{r} \delta h_{{\rm avg},i}$
Word rank $r$
Per word average happiness shift $\delta h_{{\rm avg},r}$ (%)
Royal Wedding of Prince William & Catherine Middleto
B
$T_{\rm ref}$: 7 days before and after ($h_{\rm avg}$=5.98)
$T_{\rm comp}$: Friday, 2011/04/29 ($h_{\rm avg}$=6.04)
wedding +↑
dead −↓
dont −↓
death −↓
beautiful +↑
hate −↓
kiss +↑
prince +↑
+↓easter
+↓happy
no −↓
princess +↑
never −↓
shit −↓
killed −↓
not −↓
dress +↑
real +↑
+↓me
+↓good
+↓you
weekend +↑
friday +↑
party +↑
+↓chocolate
+↓love
+↓win
married +↑
+↓game
live +↑
bad −↓
congrats +↑
amazing +↑
kill −↓
ill −↓
nigga −↓
wow +↑
died −↓
+↓life
ass −↓
hell −↓
gorgeous +↑
congratulations +↑
+↓lol
couple +↑
+↓friends
kissed +↑
she +↑
killing −↓
+↓haha
Text size:
$T_{\rm ref}$ $T_{\rm comp}$
Balance:
−68 : +168
Word rank $r$
Per word average happiness shift $\delta h_{{\rm avg},r}$ (%)
Death of Osama Bin Laden:
C
$T_{\rm ref}$: 7 days before and after ($h_{\rm avg}$=5.98)
$T_{\rm comp}$: Monday, 2011/05/02 ($h_{\rm avg}$=5.89)
−↑dead
−↑death
−↑killed
+↓love
−↑not
−↑kill
−↑died
−↑killing
+↓happy
−↑saddam
+↓haha
+↓me
−↑war
+↓mothers
−↑terrorist
−↑bad
−↑pakistan
−↑buried
no −↓
−↑hussein
+↓hahaha
−↑terror
+↓like
+↓wedding
−↑terrorism
−↑enemy
+↓lol
−↑shot
celebrating +↑
ill −↓
−↑hell
celebrate +↑
+↓mom
−↑loss
−↑mourn
america +↑
+↓you
−↑deaths
1st +↑
sea +↑
usa +↑
+↓beautiful
peace +↑
+↓fun
−↑evil
+↓good
−↑attack
we +↑
−↑breaking
−↑burial
Text size:
$T_{\rm ref}$ $T_{\rm comp}$
Balance:
−143 : +43
Word rank $r$
Per word average happiness shift $\delta h_{{\rm avg},r}$ (%)

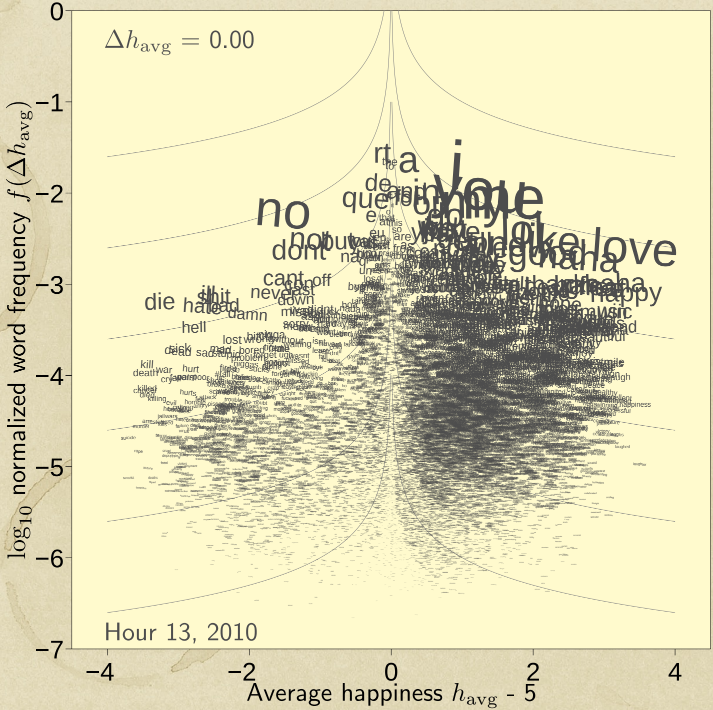
# Twitter—weekly time series:



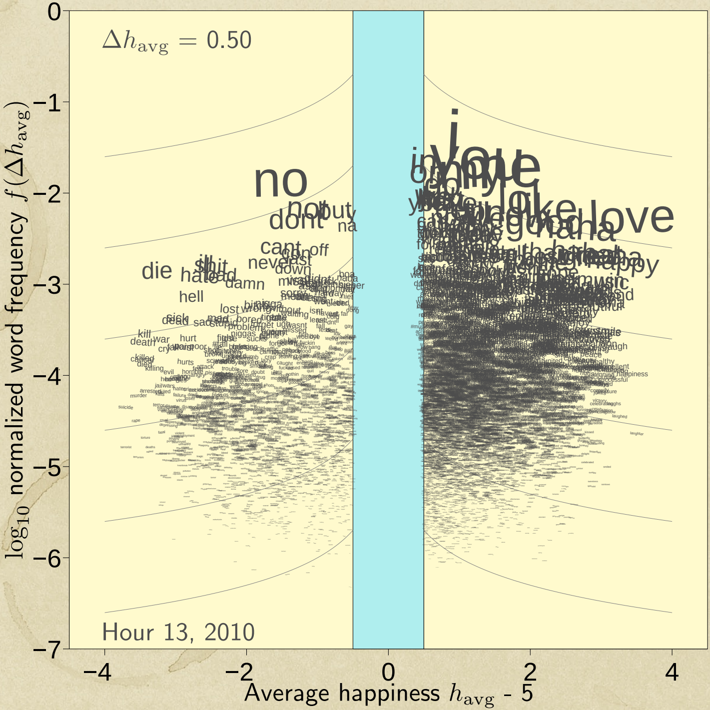

$T_{\rm ref}$: Tuesdays ($h_{\rm avg}$=6.03)
$T_{\rm comp}$: Saturdays ($h_{\rm avg}$=6.06)
Word rank $r$
love +↑
no −↓
haha +↑
party +↑
fun +↑
saturday +↑
+↓new
weekend +↑
not −↓
happy +↑
dont −↓
−↑last
hahaha +↑
−↑bored
−↑drunk
live +↑
die −↓
friends +↑
game +↑
con −↓
movie +↑
cant −↓
−↑fight
birthday +↑
+↓google
great +↑
sunday +↑
family +↑
beautiful +↑
beach +↑
home +↑
+↓lunch
sick −↓
shopping +↑
playing +↑
−↑don't
amazing +↑
bad −↓
awesome +↑
homework −↓
wedding +↑
−↑hangover
−↑miss
+↓free
shit −↓
court −↓
nice +↑
won +↑
+↓school
movies +↑
Text size:
$T_{\rm ref}$
$T_{\rm comp}$
Balance:
−87 : +187
+↓
+↑
−↑
−↓
$\sum_{i=1}^{r} \delta h_{{\rm avg},i}$

# The daily unravelling of the human mind:

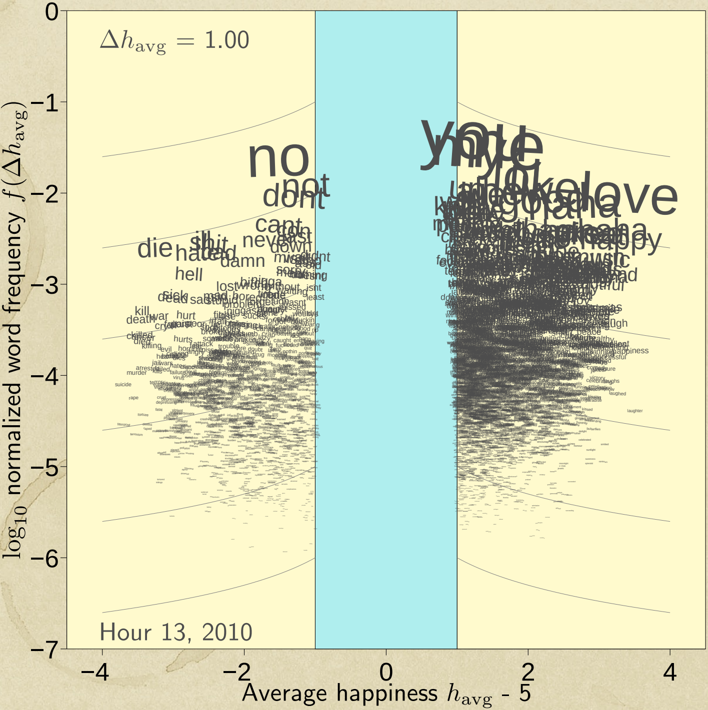




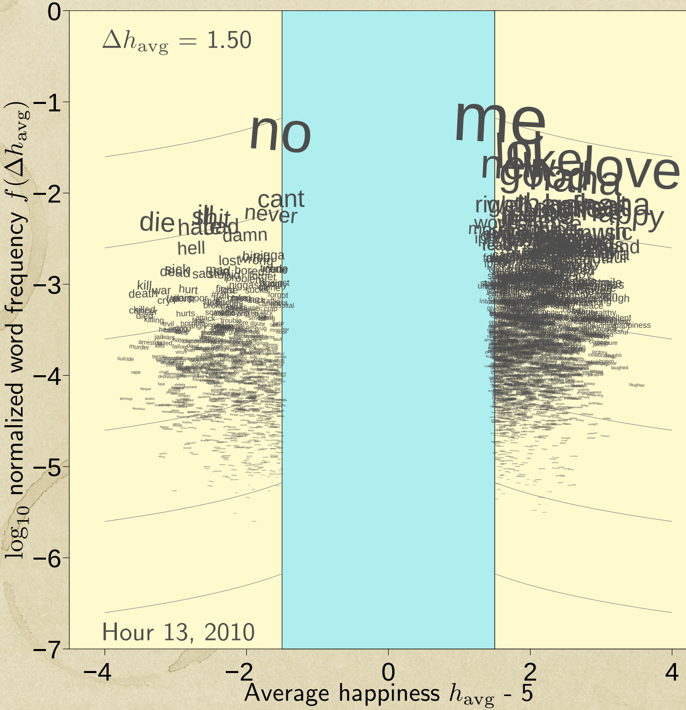
$\Delta h_{\text{avg}} = 0.00$
Hour 13, 2010
$\log_{10}$ normalized word frequency $f(\Delta h_{\text{avg}})$
Average happiness $h_{\text{avg}}$ - 5
0
−1
−2
−3
−4
−5
−6
−7
−4
−2
0
2
4



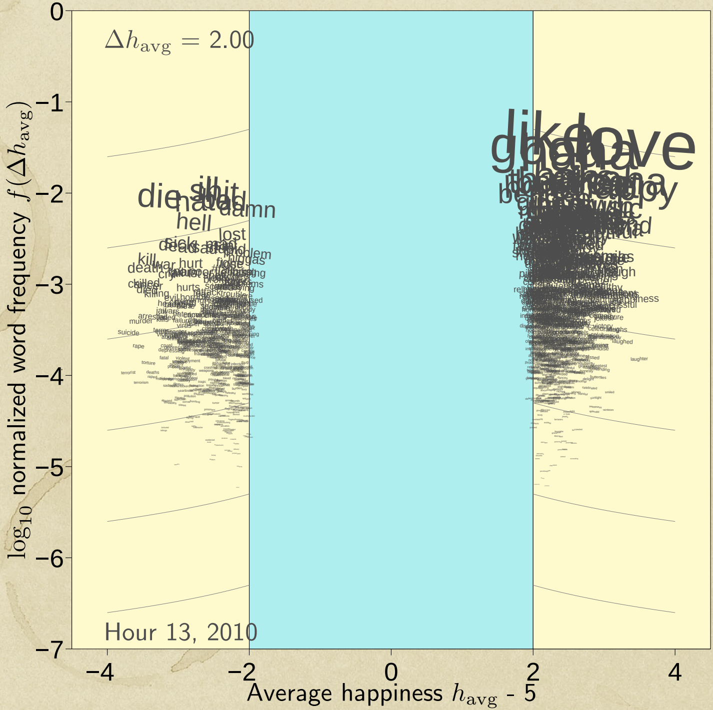
$\Delta h_{\text{avg}} = 0.50$
Hour 13, 2010
$\log_{10}$ normalized word frequency $f(\Delta h_{\text{avg}})$
Average happiness $h_{\text{avg}}$ - 5
0
−1
−2
−3
−4
−5
−6
−7
−4
−2
0
2
4
no
not
but
dont
na
cant
off
never
least
down
die
hate
bad
damn
hell
lost
sick
dead
kill
death
war
hurt
love
lol
like
haha
happy
music



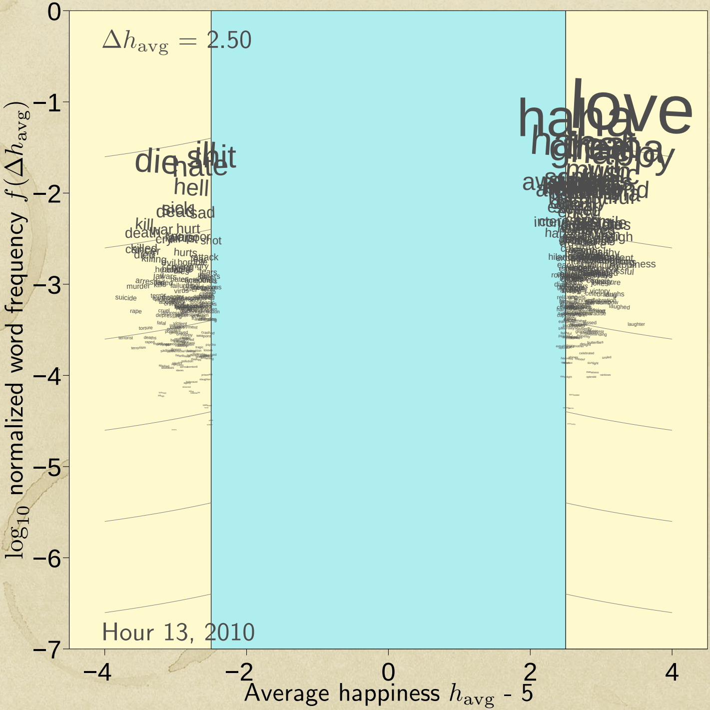
$\Delta h_{\text{avg}} = 1.00$
Hour 13, 2010
$\log_{10}$ normalized word frequency $f(\Delta h_{\text{avg}})$
Average happiness $h_{\text{avg}}$ - 5
no
not
dont
cant
never
die
hated
damn
hell
you
love
lol
haha
happy



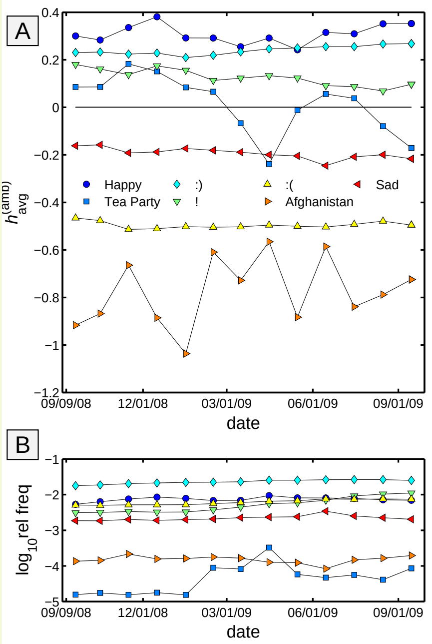
$\Delta h_{\rm avg} = 1.50$
Hour 13, 2010
no
me
love
cant
never
die
damn
hell
0
−1
−2
−3
−4
−5
−6
−7
−4
−2
0
2
4
$\log_{10}$ normalized word frequency $f(\Delta h_{\rm avg})$
Average happiness $h_{\rm avg}$ - 5

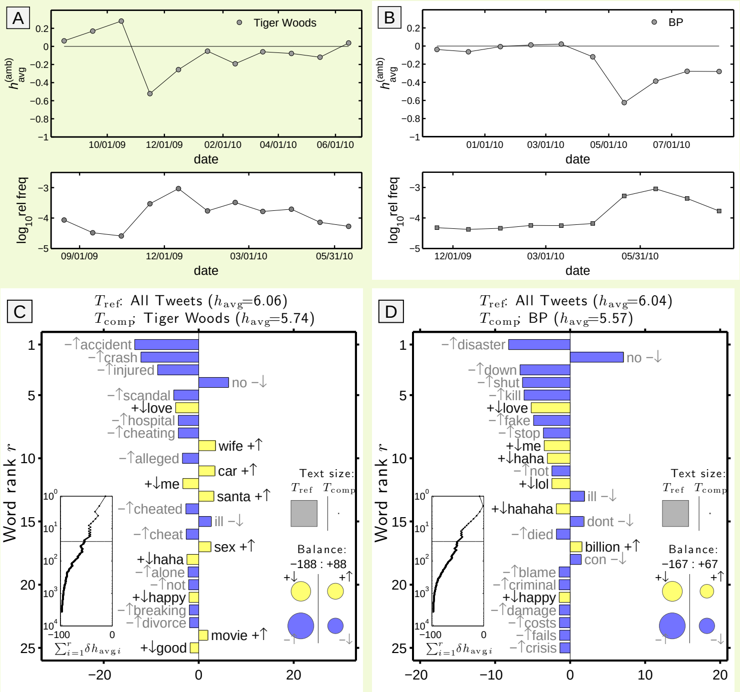


$\Delta h_{\text{avg}} = 2.00$
Hour 13, 2010
$\log_{10}$ normalized word frequency $f(\Delta h_{\text{avg}})$
Average happiness $h_{\text{avg}}$ - 5
0
−1
−2
−3
−4
−5
−6
−7
−4
−2
0
2
4
love
die
damn
hell
lost

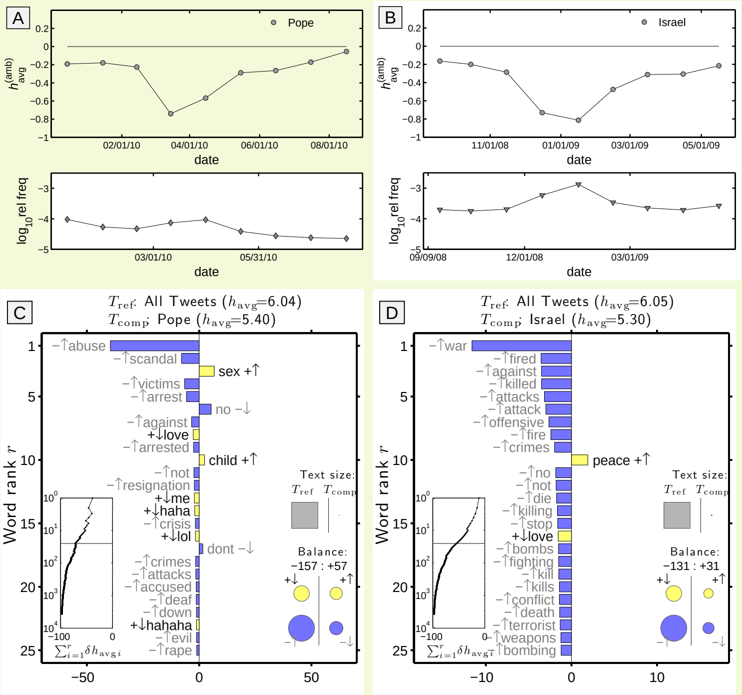


$\Delta h_{\mathrm{avg}} = 2.50$
Hour 13, 2010
love
haha
die
shit
hate
ill
hell
sick
dead
sad
kill
war
hurt
death
shot
killed
cancer
killing
hurts
attack
evil
arrested
murder
kills
suicide
rape
torture
terrorist
terrorism
fatal
depressing
laughter
laughed
laughs
happiness
celebrated
butterflies
rainbows
0
−1
−2
−3
−4
−5
−6
−7
−4
−2
0
2
4
$\log_{10}$ normalized word frequency $f(\Delta h_{\mathrm{avg}})$
Average happiness $h_{\mathrm{avg}}$ - 5


A
B
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
−1.2
$h_{avg}^{(amp)}$
Happy
:)
:(
Sad
Tea Party
!
Afghanistan
09/09/08
12/01/08
03/01/09
06/01/09
09/01/09
date
−1
−2
−3
−4
−5
$\log_{10}$rel freq


A
Tiger Woods
$h_{avg}^{(amb)}$
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
10/01/09
12/01/09
02/01/10
04/01/10
06/01/10
date
$\log_{10}$ rel freq
−3
−4
−5
09/01/09
12/01/09
03/01/10
05/31/10
date
B
BP
$h_{avg}^{(amb)}$
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
01/01/10
03/01/10
05/01/10
07/01/10
date
$\log_{10}$ rel freq
−3
−4
−5
12/01/09
03/01/10
05/31/10
date
C
$T_{ref}$: All Tweets ($h_{avg}$=6.06)
$T_{comp}$: Tiger Woods ($h_{avg}$=5.74)
Word rank $r$
−↑accident
−↑crash
−↑injured
no −↓
−↑scandal
+↓love
−↑hospital
−↑cheating
wife +↑
−↑alleged
car +↑
+↓me
santa +↑
−↑cheated
ill −↓
−↑cheat
sex +↑
+↓haha
−↑alone
−↑not
+↓happy
−↑breaking
−↑divorce
movie +↑
+↓good
Text size:
$T_{ref}$ $T_{comp}$
Balance:
−188 : +88
$\sum_{i=1}^{r} \delta h_{avg\,i}$
D
$T_{ref}$: All Tweets ($h_{avg}$=6.04)
$T_{comp}$: BP ($h_{avg}$=5.57)
Word rank $r$
−↑disaster
no −↓
−↑down
−↑shut
−↑kill
+↓love
−↑fake
−↑stop
+↓me
+↓haha
−↑not
+↓lol
ill −↓
+↓hahaha
dont −↓
−↑died
billion +↑
con −↓
−↑blame
−↑criminal
+↓happy
−↑damage
−↑costs
−↑fails
−↑crisis
Text size:
$T_{ref}$ $T_{comp}$
Balance:
−167 : +67
$\sum_{i=1}^{r} \delta h_{avg\,i}$




A
Pope
$h^{(amb)}_{avg}$
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
02/01/10
04/01/10
06/01/10
08/01/10
date
$\log_{10}$rel freq
−3
−4
−5
03/01/10
05/31/10
date
B
Israel
$h^{(amb)}_{avg}$
11/01/08
01/01/09
03/01/09
05/01/09
date
$\log_{10}$rel freq
09/09/08
12/01/08
03/01/09
date
C
$T_{ref}$: All Tweets ($h_{avg}$=6.04)
$T_{comp}$: Pope ($h_{avg}$=5.40)
−↑abuse
−↑scandal
sex +↑
−↑victims
−↑arrest
no −↓
−↑against
+↓love
−↑arrested
child +↑
−↑not
−↑resignation
+↓me
+↓haha
−↑crisis
+↓lol
dont −↓
−↑crimes
−↑attacks
−↑accused
−↑deaf
−↑down
+↓hahaha
−↑evil
−↑rape
Word rank $r$
Text size:
$T_{ref}$ $T_{comp}$
Balance:
−157 : +57
+↓ +↑
−↑ −↓
$\sum_{i=1}^{r}\delta h_{avg\,i}$
−50
0
50
D
$T_{ref}$: All Tweets ($h_{avg}$=6.05)
$T_{comp}$: Israel ($h_{avg}$=5.30)
−↑war
−↑fired
−↑against
−↑killed
−↑attacks
−↑attack
−↑offensive
−↑fire
−↑crimes
peace +↑
−↑no
−↑not
−↑die
−↑killing
−↑stop
+↓love
−↑bombs
−↑fighting
−↑kill
−↑kills
−↑conflict
−↑death
−↑terrorist
−↑weapons
−↑bombing
Word rank $r$
Text size:
$T_{ref}$ $T_{comp}$
Balance:
−131 : +31
−10
0
10

| Word | $h_{\rm avg}^{\rm (amb)}$ | Total Tweets | Total ANEW | Word | $h_{\rm avg}^{\rm (amb)}$ | Total Tweets | Total ANEW |
|---|---|---|---|---|---|---|---|
| 1. love | +1.42 | 46,687,476 (6) | 85,269,499 (5) | 51. me | -0.06 | 144,342,098 (4) | 88,088,051 (4) |
| 2. happy | +1.32 | 16,541,968 (13) | 32,442,529 (8) | 52. ? | -0.07 | 2,333,283 (53) | 674,679 (69) |
| 3. win | +1.26 | 7,981,856 (26) | 14,640,728 (20) | 53. commute | -0.09 | 90,126 (94) | 90,092 (92) |
| 4. kiss | +1.21 | 1,697,405 (59) | 3,162,330 (48) | 54. gay | -0.09 | 2,727,309 (47) | 1,697,177 (57) |
| 5. cash | +1.21 | 1,279,236 (63) | 2,468,496 (51) | 55. right | -0.10 | 19,166,480 (10) | 15,850,283 (19) |
| 6. vacation | +1.11 | 934,501 (67) | 1,783,270 (56) | 56. school | -0.11 | 9,264,217 (24) | 6,924,193 (34) |
| 7. Christmas | +1.03 | 4,887,968 (35) | 10,645,630 (25) | 57. Republican | -0.13 | 229,773 (86) | 188,338 (85) |
| 8. God | +0.95 | 8,576,364 (25) | 17,867,768 (16) | 58. they | -0.16 | 27,442,360 (8) | 27,150,189 (11) |
| 9. party | +0.93 | 6,438,886 (29) | 12,090,597 (23) | 59. winter | -0.19 | 1,255,945 (64) | 1,217,225 (64) |
| 10. sex | +0.89 | 3,551,767 (39) | 7,087,972 (31) | 60. lose | -0.19 | 2,056,468 (55) | 2,091,540 (53) |
| 11. Valentine | +0.85 | 247,288 (84) | 464,914 (75) | 61. Jon Stewart | -0.20 | 52,084 (97) | 33,086 (96) |
| 12. family | +0.79 | 5,014,816 (32) | 10,629,361 (26) | 62. gas | -0.22 | 1,022,879 (65) | 812,029 (68) |
| 13. sun | +0.65 | 2,385,348 (52) | 4,602,627 (44) | 63. no | -0.22 | 95,129,093 (5) | 38,894,616 (6) |
| 14. life | +0.50 | 14,006,454 (17) | 27,770,768 (10) | 64. Democrat | -0.23 | 93,193 (93) | 75,450 (93) |
| 15. hope | +0.48 | 11,833,337 (18) | 22,952,366 (13) | 65. left | -0.27 | 4,893,634 (34) | 4,611,878 (43) |
| 16. heaven | +0.43 | 741,878 (71) | 1,485,702 (59) | 66. Senate | -0.29 | 447,732 (78) | 316,835 (80) |
| 17. :) | +0.42 | 10,470,483 (20) | 6,787,678 (35) | 67. election | -0.30 | 560,184 (75) | 375,055 (78) |
| 18. income | +0.36 | 510,425 (76) | 418,161 (77) | 68. Sarah Palin | -0.34 | 225,577 (87) | 150,096 (88) |
| 19. friends | +0.33 | 7,669,719 (27) | 7,541,106 (29) | 69. Obama | -0.35 | 2,981,150 (44) | 1,998,326 (54) |
| 20. snow | +0.32 | 2,596,165 (49) | 5,011,785 (40) | 70. economy | -0.36 | 608,878 (73) | 460,834 (76) |
| 21. :-) | +0.32 | 1,680,165 (60) | 1,102,512 (67) | 71. Congress | -0.36 | 391,510 (79) | 279,695 (81) |
| 22. night | +0.29 | 17,089,505 (12) | 17,606,796 (17) | 72. drugs | -0.39 | 509,606 (77) | 469,091 (74) |
| 23. vegan | +0.28 | 183,889 (90) | 178,676 (86) | 73. Muslim | -0.42 | 215,300 (88) | 146,506 (89) |
| 24. Jesus | +0.27 | 2,027,720 (56) | 1,673,992 (58) | 74. George Bush | -0.43 | 32,341 (98) | 23,102 (98) |
| 25. girl | +0.25 | 10,070,132 (22) | 19,886,691 (14) | 75. climate | -0.44 | 364,177 (80) | 229,129 (83) |
| 26. USA | +0.23 | 2,157,172 (54) | 1,204,585 (65) | 76. Pope | -0.51 | 152,320 (91) | 135,955 (90) |
| 27. you | +0.22 | 173,276,993 (3) | 145,464,084 (2) | 77. oil | -0.53 | 1,377,355 (62) | 1,148,990 (66) |
| 28. our | +0.21 | 14,062,465 (16) | 14,437,899 (21) | 78. I feel | -0.54 | 5,173,513 (31) | 4,702,352 (42) |
| 29. ;) | +0.20 | 2,618,940 (48) | 1,475,221 (60) | 79. Glenn Beck | -0.54 | 113,991 (92) | 101,090 (91) |
| 30. health | +0.20 | 2,575,543 (50) | 4,950,202 (41) | 80. Islam | -0.54 | 187,223 (89) | 70,311 (94) |
| 31. tomorrow | +0.20 | 10,379,637 (21) | 8,899,406 (28) | 81. :-( | -0.65 | 341,141 (81) | 244,215 (82) |
| 32. ! | +0.16 | 3,463,257 (40) | 1,385,072 (62) | 82. :( | -0.70 | 2,907,145 (45) | 1,891,225 (55) |
| 33. summer | +0.13 | 2,998,785 (43) | 2,554,459 (50) | 83. flu | -0.75 | 901,403 (68) | 639,000 (70) |
| 34. we | +0.13 | 39,132,934 (7) | 34,513,587 (7) | 84. rain | -0.78 | 3,233,464 (41) | 5,959,903 (38) |
| 35. today | +0.13 | 25,588,506 (9) | 23,619,518 (12) | 85. BP | -0.78 | 582,167 (74) | 326,100 (79) |
| 36. man | +0.12 | 15,856,341 (14) | 29,558,118 (9) | 86. mosque | -0.79 | 69,812 (95) | 46,736 (95) |
| 37. woman | +0.10 | 2,543,036 (51) | 5,603,347 (39) | 87. dark | -0.95 | 1,577,553 (61) | 3,233,911 (47) |
| 38. Stephen Colbert | +0.10 | 23,778 (99) | 14,697 (99) | 88. Lehman Brothers | -1.08 | 8,500 (100) | 4,280 (100) |
| 39. ;-) | +0.10 | 943,413 (66) | 516,171 (73) | 89. Goldman Sachs | -1.08 | 52,703 (96) | 30,769 (97) |
| 40. RT | +0.06 | 339,055,724 (1) | 142,219,359 (3) | 90. Afghanistan | -1.15 | 273,519 (83) | 172,637 (87) |
| 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 | 1,812,251 (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) | 16,191,802 (18) | 93. gun | -1.81 | 680,903 (72) | 1,263,217 (63) |
| 44. I | +0.02 | 307,960,343 (2) | 282,865,043 (1) | 94. hate | -2.43 | 9,652,881 (23) | 18,158,870 (15) |
| 45. yes | +0.02 | 11,593,356 (19) | 7,499,840 (30) | 95. hell | -2.49 | 6,266,162 (30) | 11,056,735 (24) |
| 46. them | 0.00 | 15,352,295 (15) | 14,398,889 (22) | 96. sick | -2.55 | 3,576,058 (37) | 6,783,395 (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 | 1,955,901 (57) | 3,417,588 (46) |
| 49. yesterday | -0.01 | 3,077,761 (42) | 2,852,623 (49) | 99. depressed | -2.64 | 280,872 (82) | 541,394 (72) |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |

## Text element and context correlate in happiness scores:

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

## Text element and context correlate in happiness scores:

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

## Text element and context correlate in happiness scores:

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

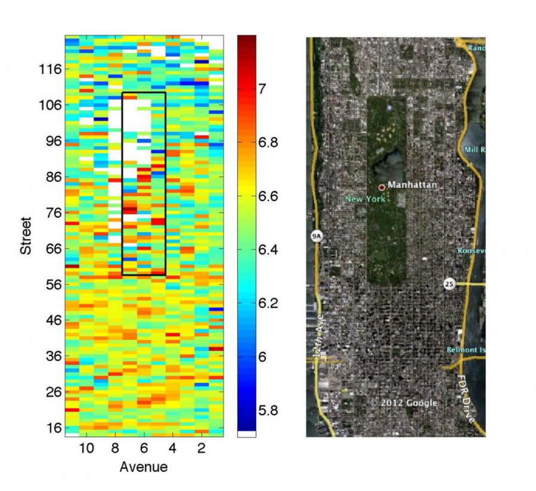
## Text element and context correlate in happiness scores:

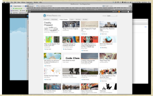
- Compare ambient happiness with text element happiness.
- Spearman correlation coefficient: $r_s \simeq 0.79$, $p$-value $< 10^{-10}$.
- An on-average result: says nothing about any individual sentence.
- Extra random piece: stemming fails.

# Happiness in Manhattan (just for fun):



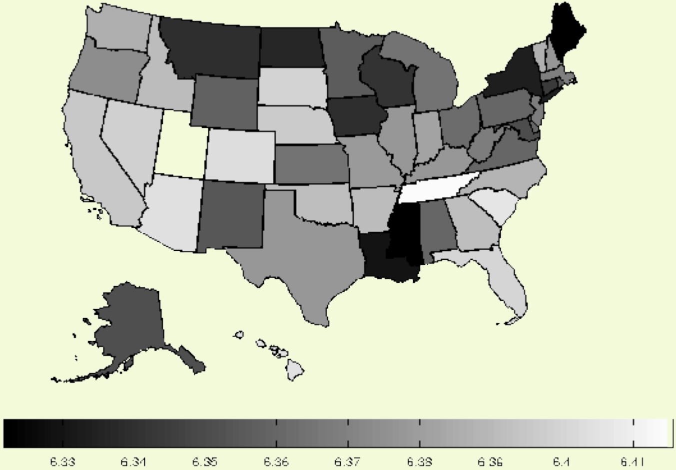




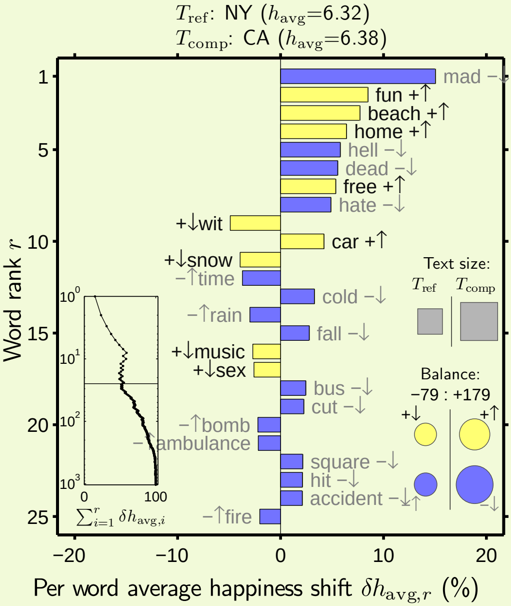



See Blog post on onehappybird (⊞)

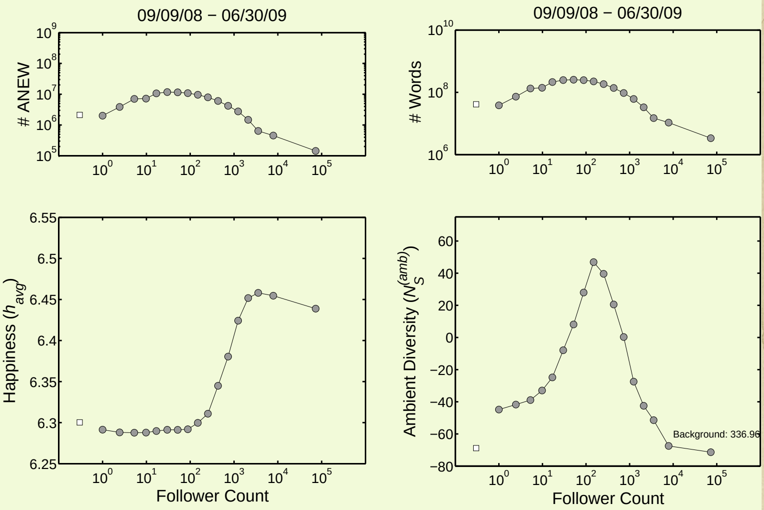
# Twitter—location:



# Twitter—location:

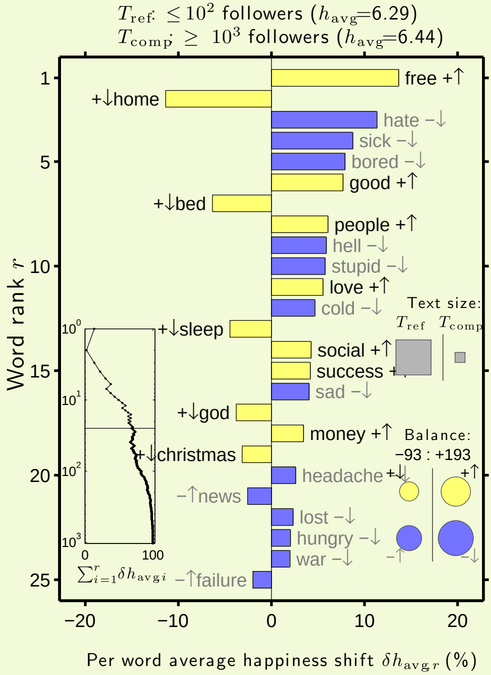


# Twitter—popularity based on follower count:

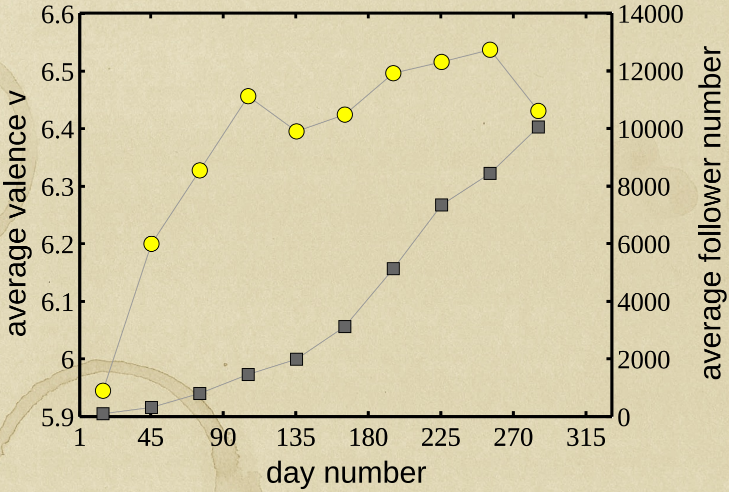



- Dunbar's number $\simeq$ 150.



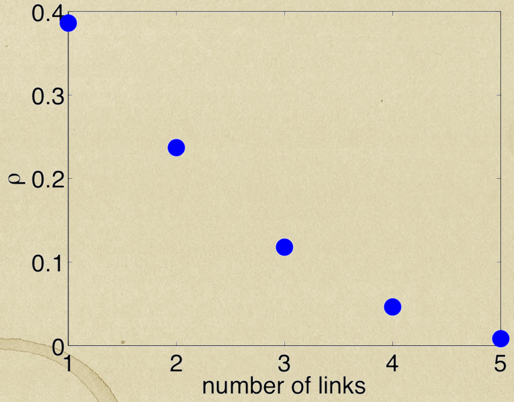


$T_{\rm ref}$: $\leq 10^2$ followers ($h_{\rm avg}$=6.29)
$T_{\rm comp}$: $\geq$ $10^3$ followers ($h_{\rm avg}$=6.44)
free +↑
+↓home
hate −↓
sick −↓
bored −↓
good +↑
+↓bed
people +↑
hell −↓
stupid −↓
love +↑
cold −↓
+↓sleep
social +↑
success +↑
sad −↓
+↓god
money +↑
+↓christmas
headache −↓
−↑news
lost −↓
hungry −↓
war −↓
−↑failure
Text size:
$T_{\rm ref}$ $T_{\rm comp}$
Balance:
−93 : +193
+↓ +↑
−↑ −↓
Word rank $r$
Per word average happiness shift $\delta h_{{\rm avg},r}$ (%)
$\sum_{i=1}^{r} \delta h_{{\rm avg}\,i}$



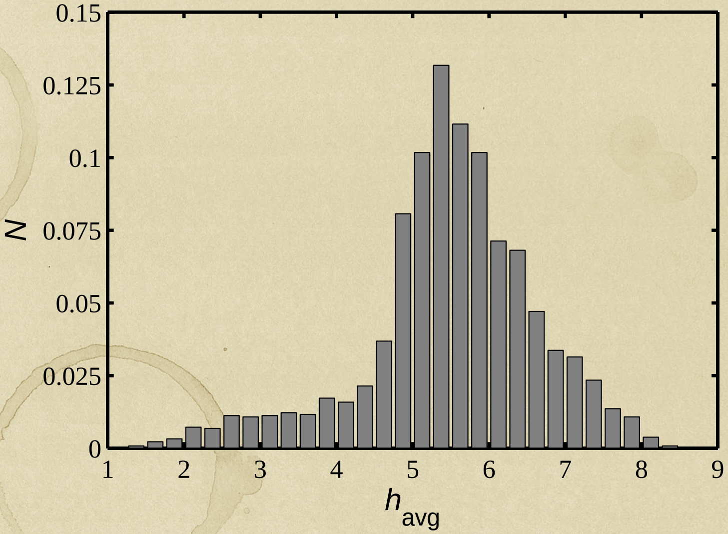
average valence v
average follower number
day number
6.6
6.5
6.4
6.3
6.2
6.1
6
5.9
14000
12000
10000
8000
6000
4000
2000
0
1
45
90
135
180
225
270
315

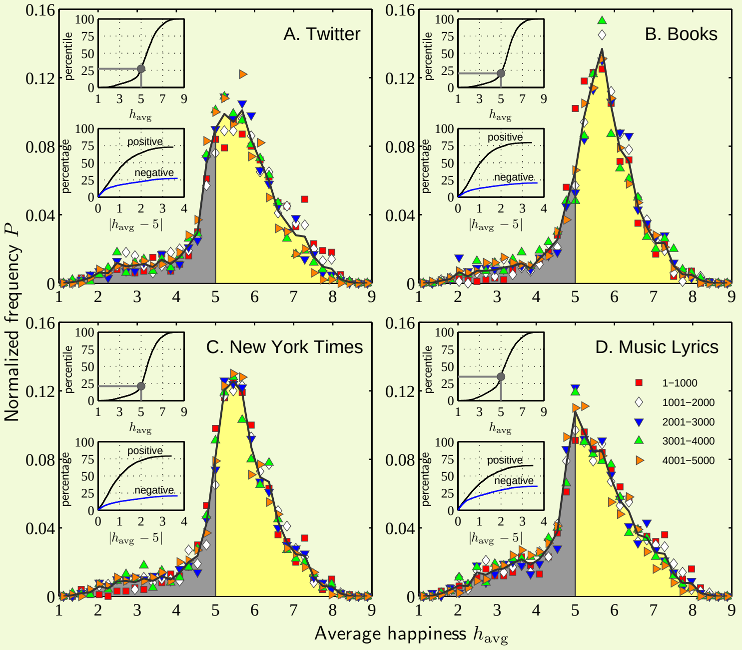
# Twitter—interactions:

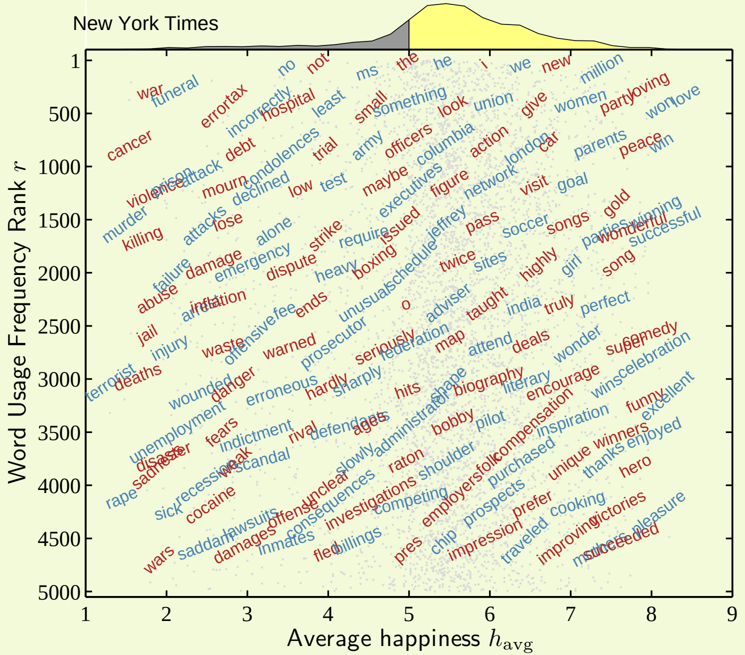


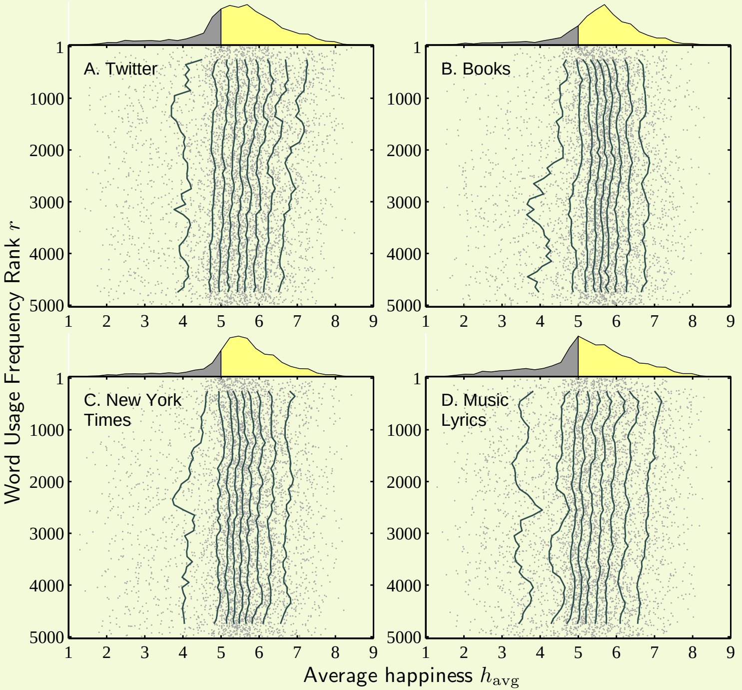

- Decay in happiness correlation in social network.
- $\rho$ = Spearman's correlation coefficient.

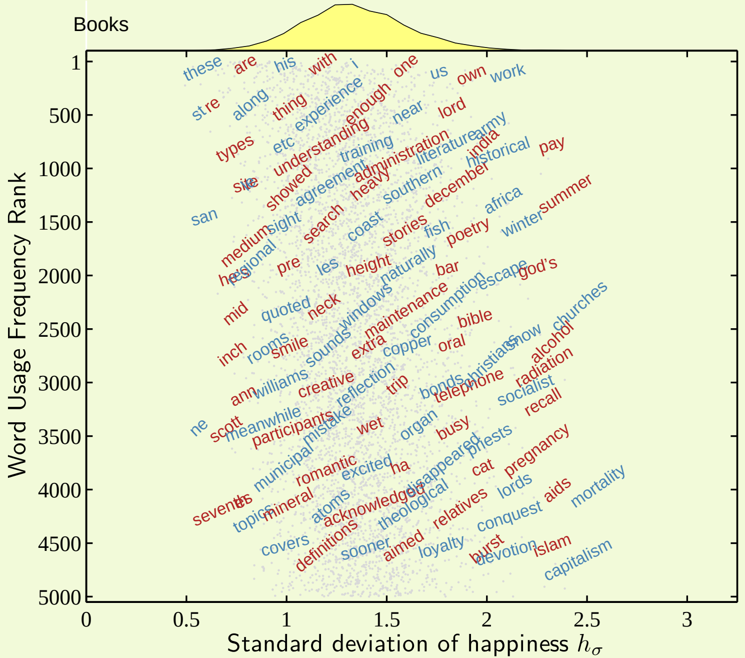
# Positive bias in the English language:

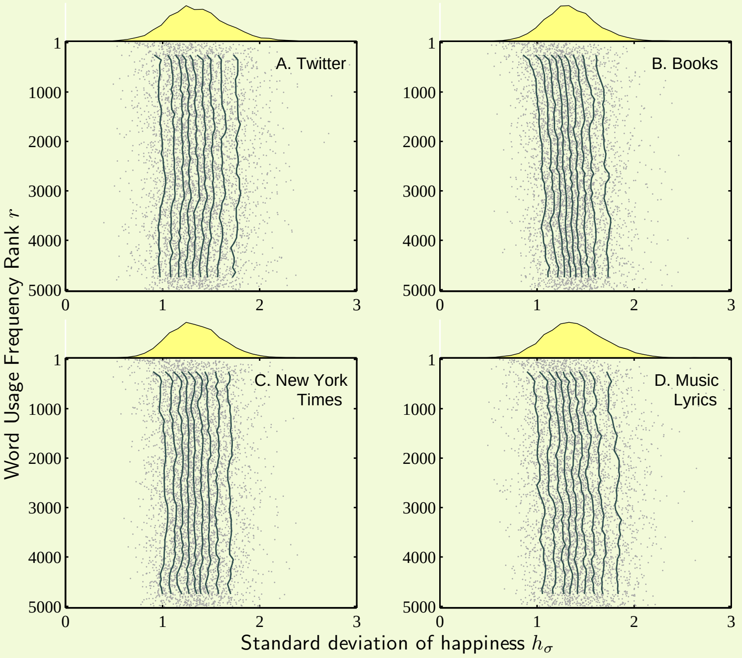



A. Twitter
B. Books
C. New York Times
D. Music Lyrics
Normalized frequency $P$
Average happiness $h_{\text{avg}}$
percentile
percentage
$h_{\text{avg}}$
$|h_{\text{avg}} - 5|$
positive
negative
1–1000
1001–2000
2001–3000
3001–4000
4001–5000


New York Times
Word Usage Frequency Rank $r$
Average happiness $h_{\text{avg}}$
1
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
1 2 3 4 5 6 7 8 9


A. Twitter
B. Books
C. New York Times
D. Music Lyrics
Word Usage Frequency Rank $r$
Average happiness $h_{\text{avg}}$
1
1000
2000
3000
4000
5000
1 2 3 4 5 6 7 8 9


Books
Word Usage Frequency Rank
Standard deviation of happiness $h_\sigma$
1
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
0
0.5
1
1.5
2
2.5
3
these
are
his
with
i
one
us
own
work
along
thing
experience
enough
near
lord
types
etc
understanding
training
administration
literature
army
india
historical
pay
showed
agreement
heavy
southern
december
san
search
coast
stories
fish
africa
summer
winter
poetry
medium
sight
regional
pre
les
height
naturally
bar
escape
god's
mid
quoted
neck
windows
maintenance
consumption
bible
churches
inch
rooms
smile
sounds
extra
copper
oral
alcohol
williams
creative
reflection
trip
telephone
radiation
socialist
recall
ne
scott
meanwhile
participants
mistake
wet
organ
busy
priests
pregnancy
municipal
romantic
excited
ha
cat
lords
mortality
topics
mineral
atoms
acknowledged
theological
relatives
conquest
aids
covers
definitions
sooner
aimed
loyalty
burst
devotion
islam
capitalism


A. Twitter
B. Books
C. New York Times
D. Music Lyrics
Word Usage Frequency Rank $r$
Standard deviation of happiness $h_\sigma$
1
1000
2000
3000
4000
5000
0
1
2
3

# Random other things (now and next):

- Gross National Happiness Index, 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 ...

# Random other things (now and next):

- Gross National Happiness Index, 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 . . .

# Random other things (now and next):

- Gross National Happiness Index, 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 . . .

# Random other things (now and next):

- Gross National Happiness Index, 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 ...

# Random other things (now and next):

- Gross National Happiness Index, 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 . . .

# Random other things (now and next):

- Gross National Happiness Index, 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 . . .

# Random other things (now and next):



- Gross National Happiness Index, 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 ...

# 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 (⊞)

[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.

# 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 (⊞)

[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 (⊞)

# 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 (⊞)

[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 (⊞)

# References IV



[10] M. A. Killingsworth and D. T. Gilbert.
A wondering mind is an unhappy mind.
Science Magazine, 330:932, 2010. pdf (⊞)

[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 (⊞)

[12] R. Layard.
Happiness.
The Penguin Press, London, 2005.

# 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 (⊞)

[15] G. Miller.
Social Scientists wade into the Tweet stream.
Science Magazine, 333:1814–1815, 2011. pdf (⊞)

[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.

# References VI



[17] E. Sandhaus.
The New York Times Annotated Corpus.
Linguistic Data Consortium, Philadelphia, 2008.