THE MINOR FALL, THE MAJOR LIFT:

WHAT CAN WE LEARN FROM LARGE CORPORA OF LYRICS?

Yong-Yeol (YY) Ahn, @yy

http://yongyeol.com

CNetS, IUNI, School of Informatics and Computing

Indiana University Bloomington

Warning:
violence & obscene language

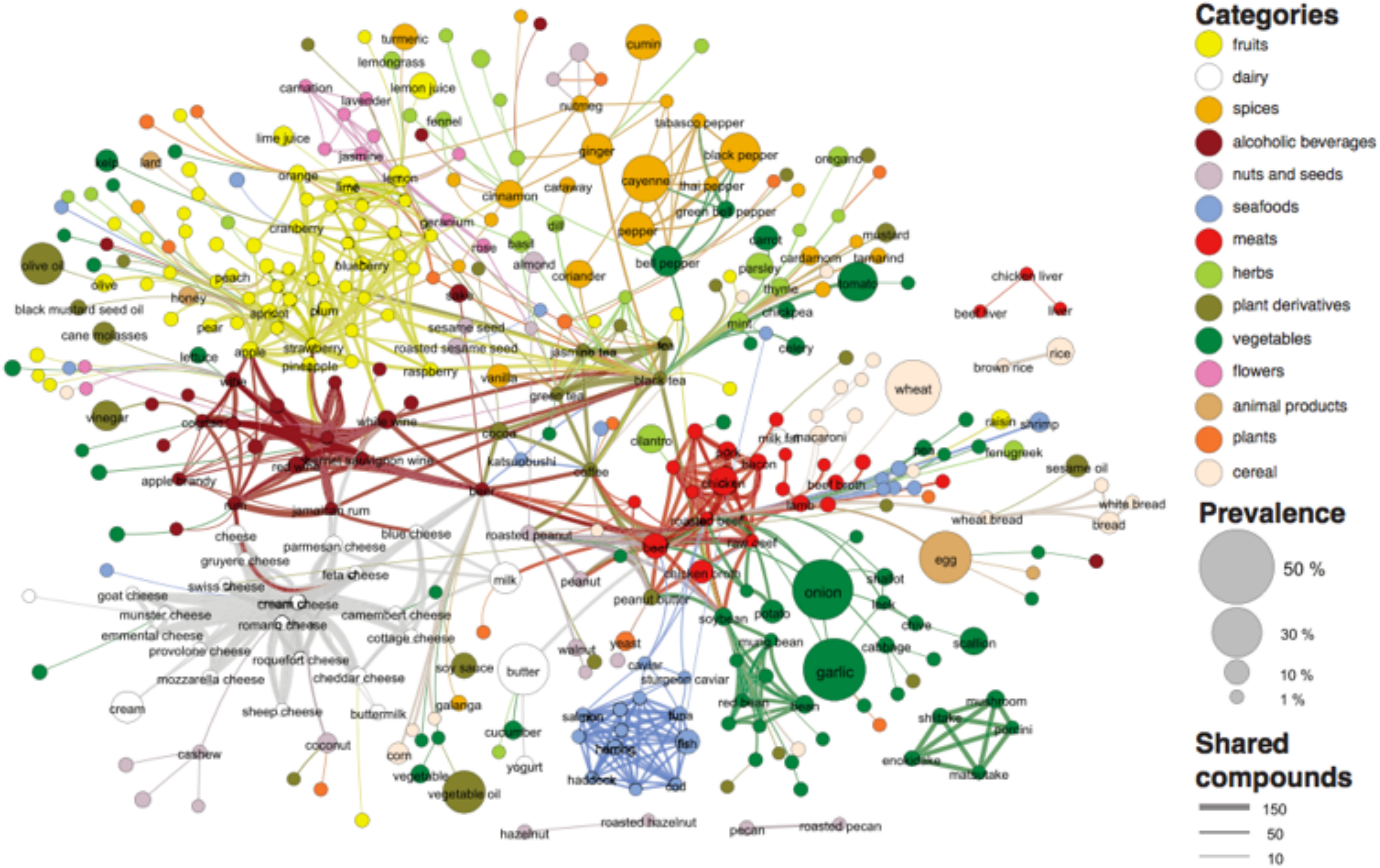
Research in Progress

Writing a paper outside your area of expertise

Research in Progress

Writing a paper outside your area of expertise

Original plan:
Ingredients, flavor compounds, and recipes —> Network



Check out CAW1 talks!

#0 Large corpora + text analysis

1.

J Happiness Stud
DOI 10.1007/s10902-009-9150-9

RESEARCH PAPER

Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents

Peter Sheridan Dodds · Christopher M. Danforth

Abstract The importance of quantifying the nature and intensity of emotional states at the level of populations is evident: we would like to know how, when, and why individuals feel as they do if we wish, for example, to better construct public policy, build more successful organizations, and, from a scientific perspective, more fully understand economic and social phenomena. Here, by incorporating direct human assessment of words, we quantify happiness levels on a continuous scale for a diverse set of large-scale texts: song titles and lyrics, weblogs, and State of the Union addresses. Our method is transparent, improvable, capable of rapidly processing Web-scale texts, and moves beyond approaches based on

1.

J Happiness Stud
DOI 10.1007/s10902-009-9150-9

RESEARCH PAPER

Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents

Peter Sheridan Dodds · Christopher M. Danforth

Abstract The importance of quantifying the nature and intensity of emotional states at the level of populations is evident: we would like to know how, when, and why individuals feel as they do if we wish, for example, to better construct public policy, build more successful organizations, and, from a scientific perspective, more fully understand economic and social phenomena. Here, by incorporating direct human assessment of words, we quantify happiness levels on a continuous scale for a diverse set of large-scale texts: song titles and lyrics, weblogs, and State of the Union addresses. Our method is transparent, improvable, capable of rapidly processing Web-scale texts, and moves beyond approaches based on

1.

J Happiness Stud
DOI 10.1007/s10902-009-9150-9

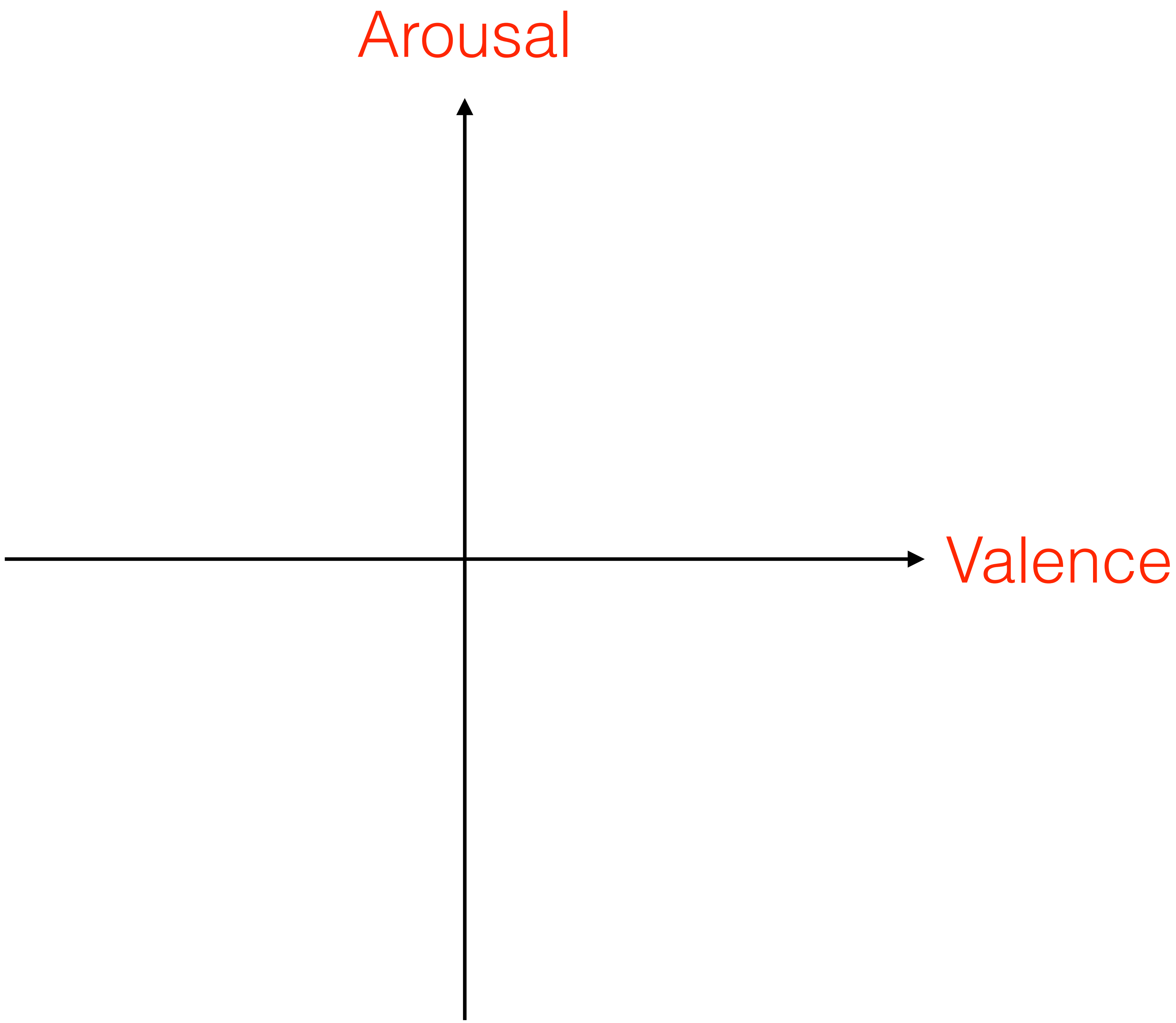
RESEARCH PAPER

Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents

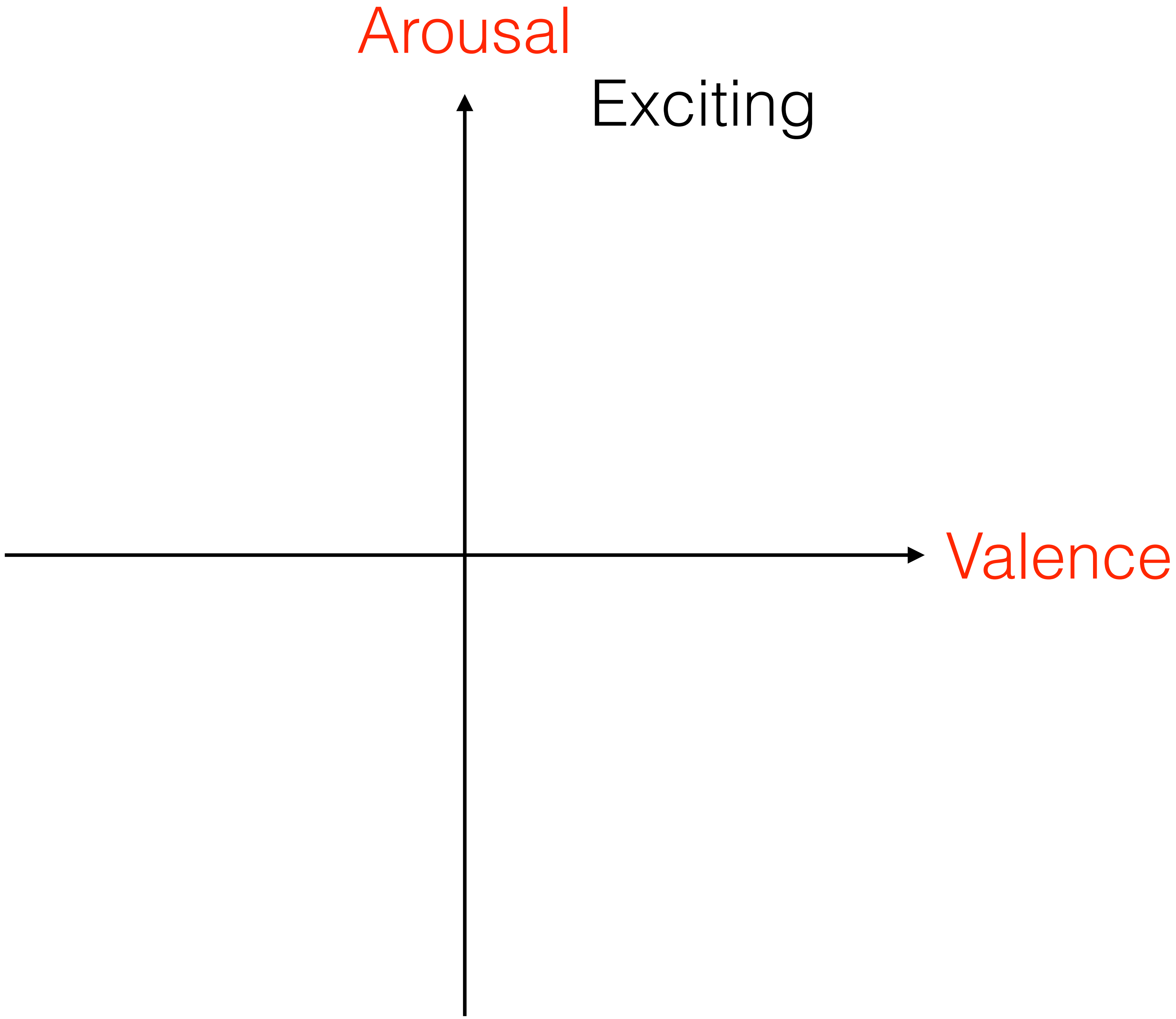
Peter Sheridan Dodds · Christopher M. Danforth

Abstract The importance of quantifying the nature and intensity of emotional states at the level of populations is evident: we would like to know how, when, and why individuals feel as they do if we wish, for example, to better construct public policy, build more successful organizations, and, from a scientific perspective, more fully understand economic and social phenomena. Here, by incorporating direct human assessment of words, we quantify happiness levels on a continuous scale for a diverse set of large-scale texts: song titles and lyrics, weblogs, and State of the Union addresses. Our method is transparent, improvable, capable of rapidly processing Web-scale texts, and moves beyond approaches based on

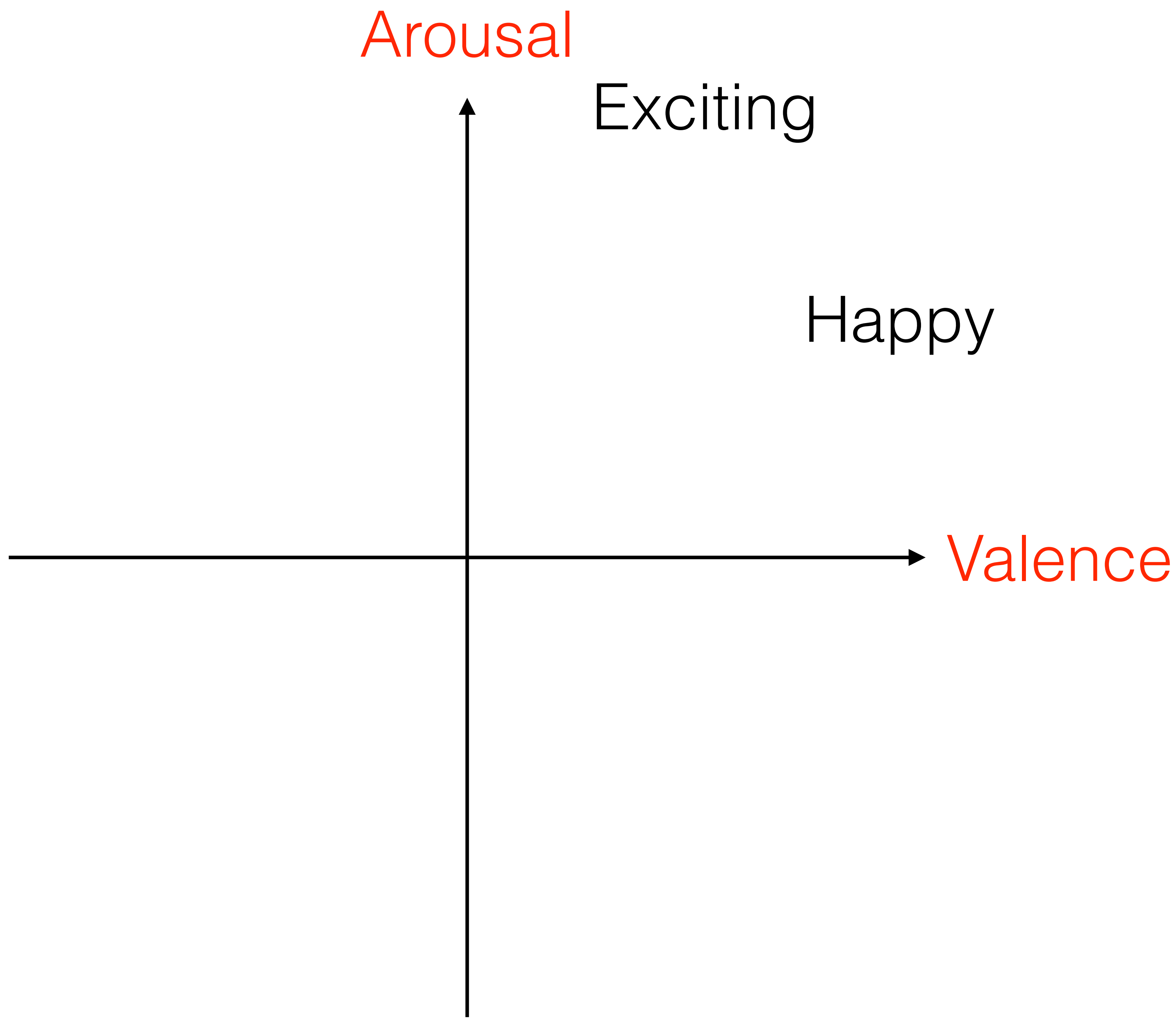
Two-dimensional model of emotion



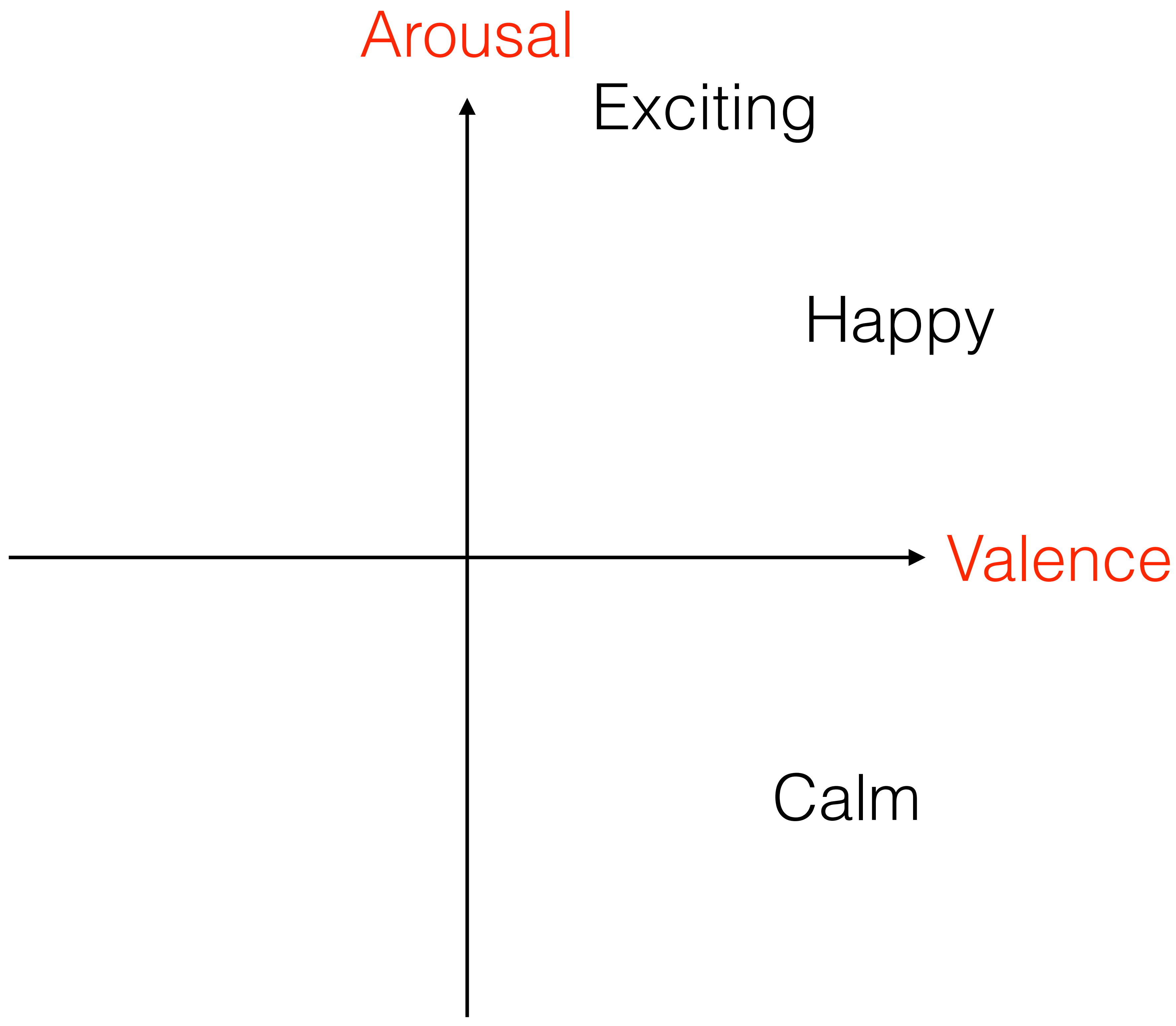
Two-dimensional model of emotion



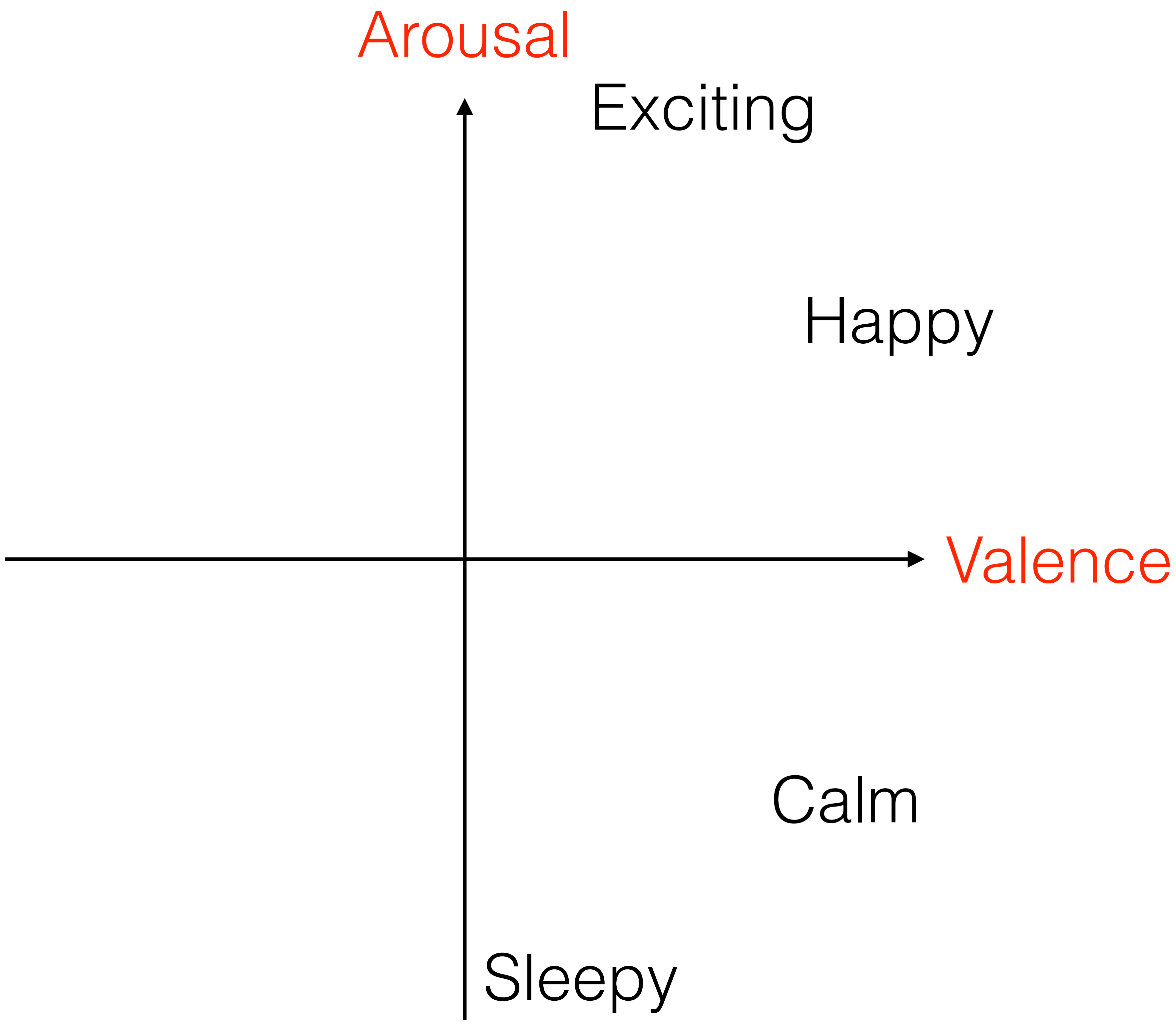
Two-dimensional model of emotion



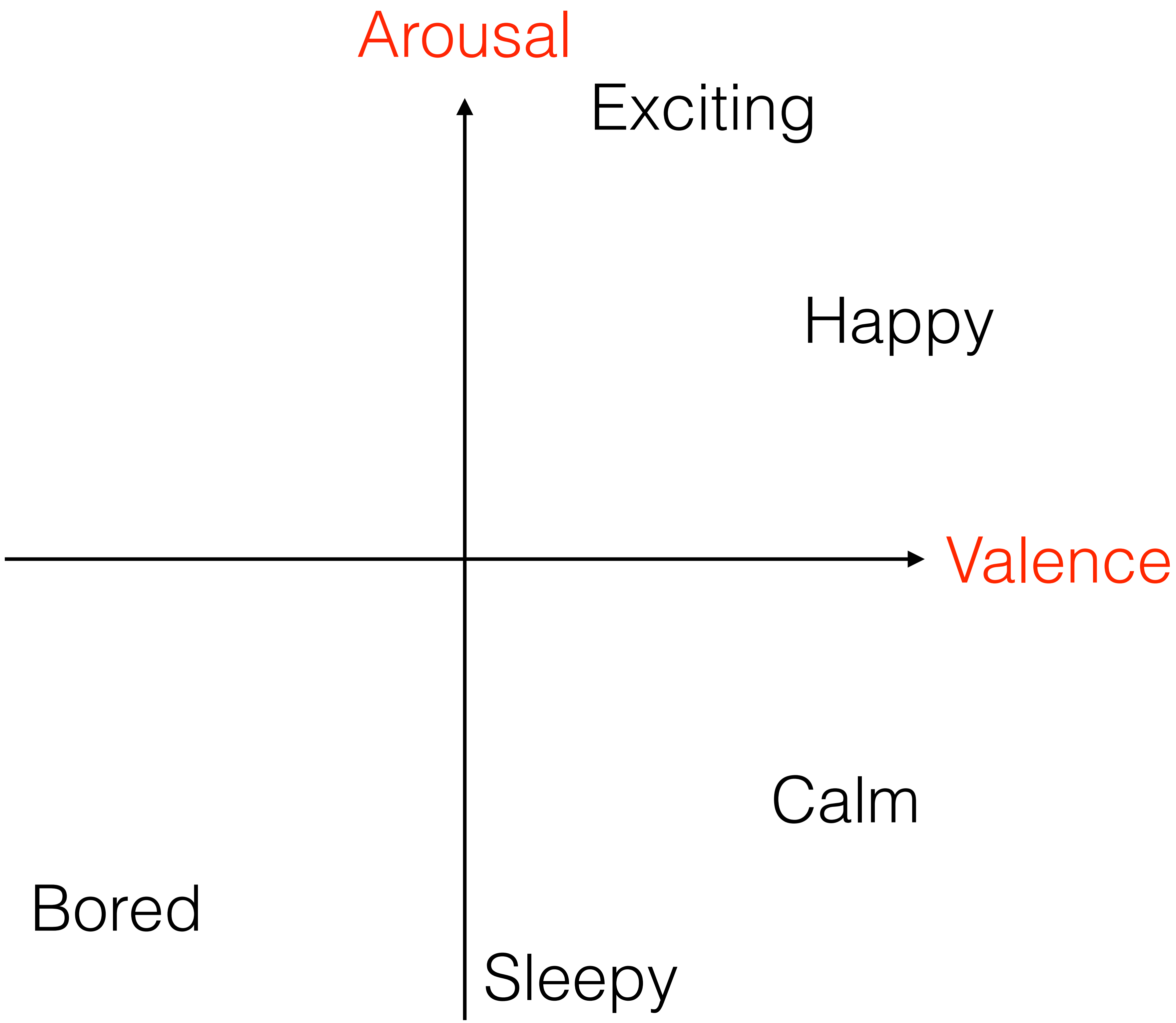
Two-dimensional model of emotion



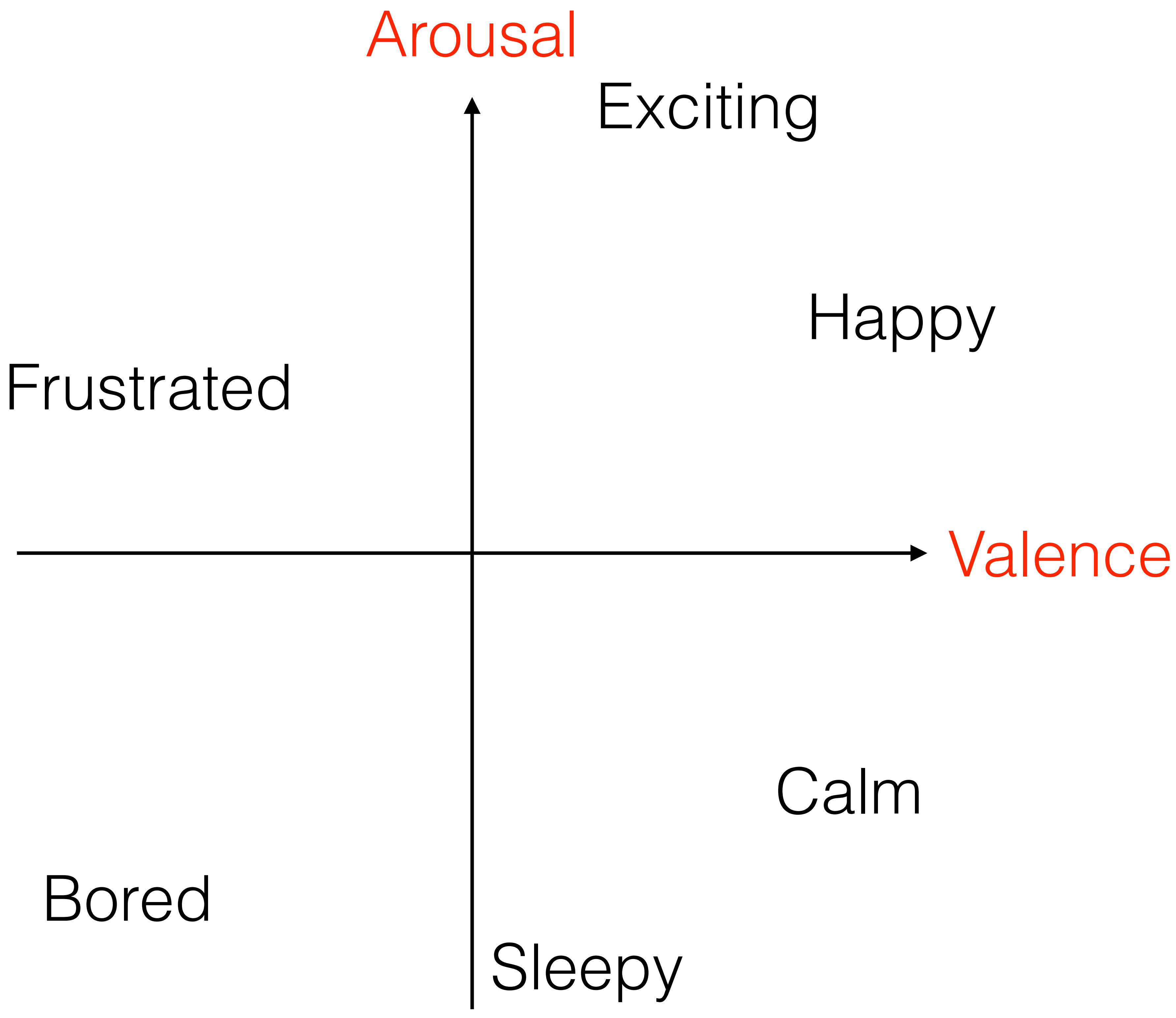
Two-dimensional model of emotion



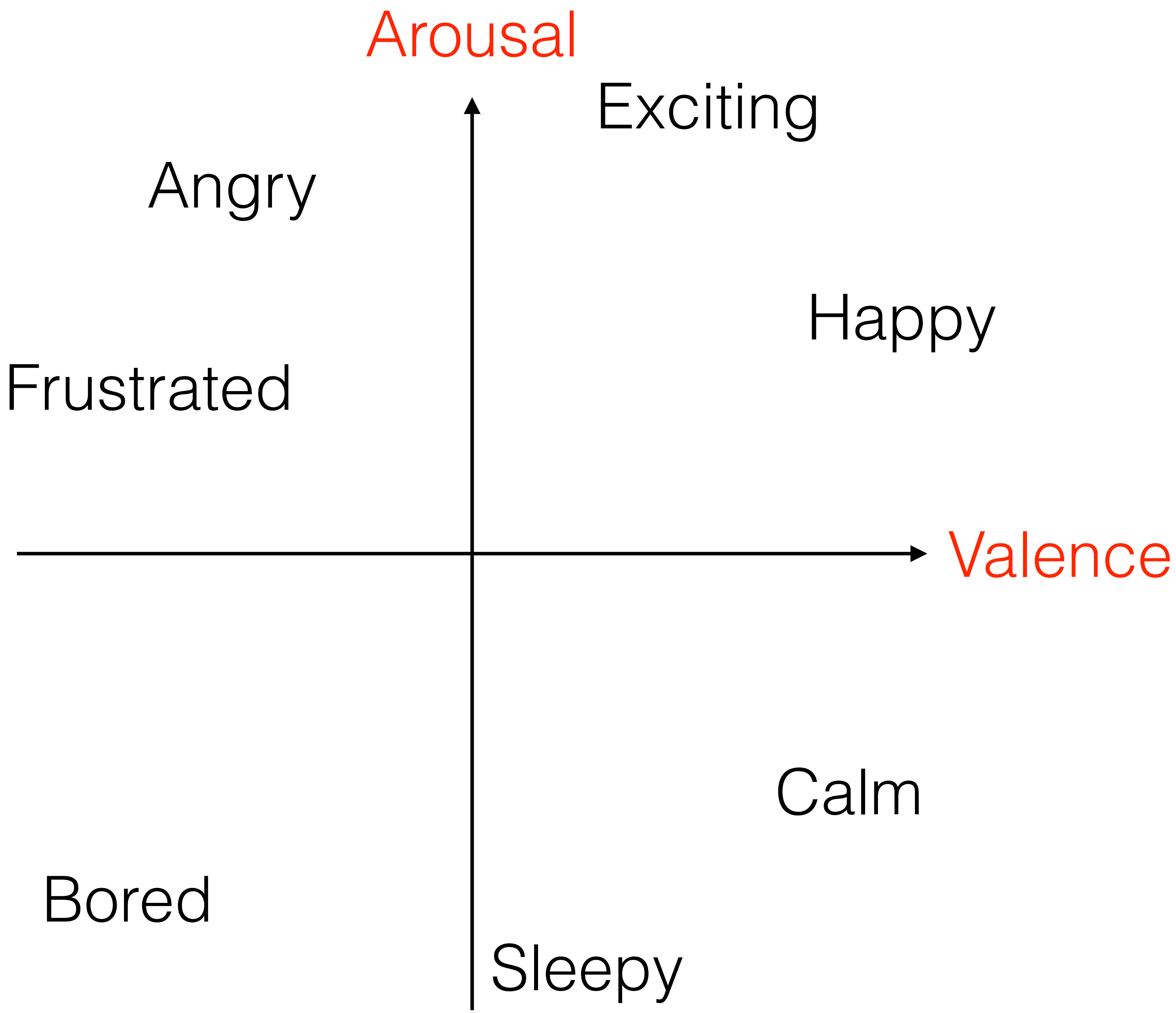
Two-dimensional model of emotion



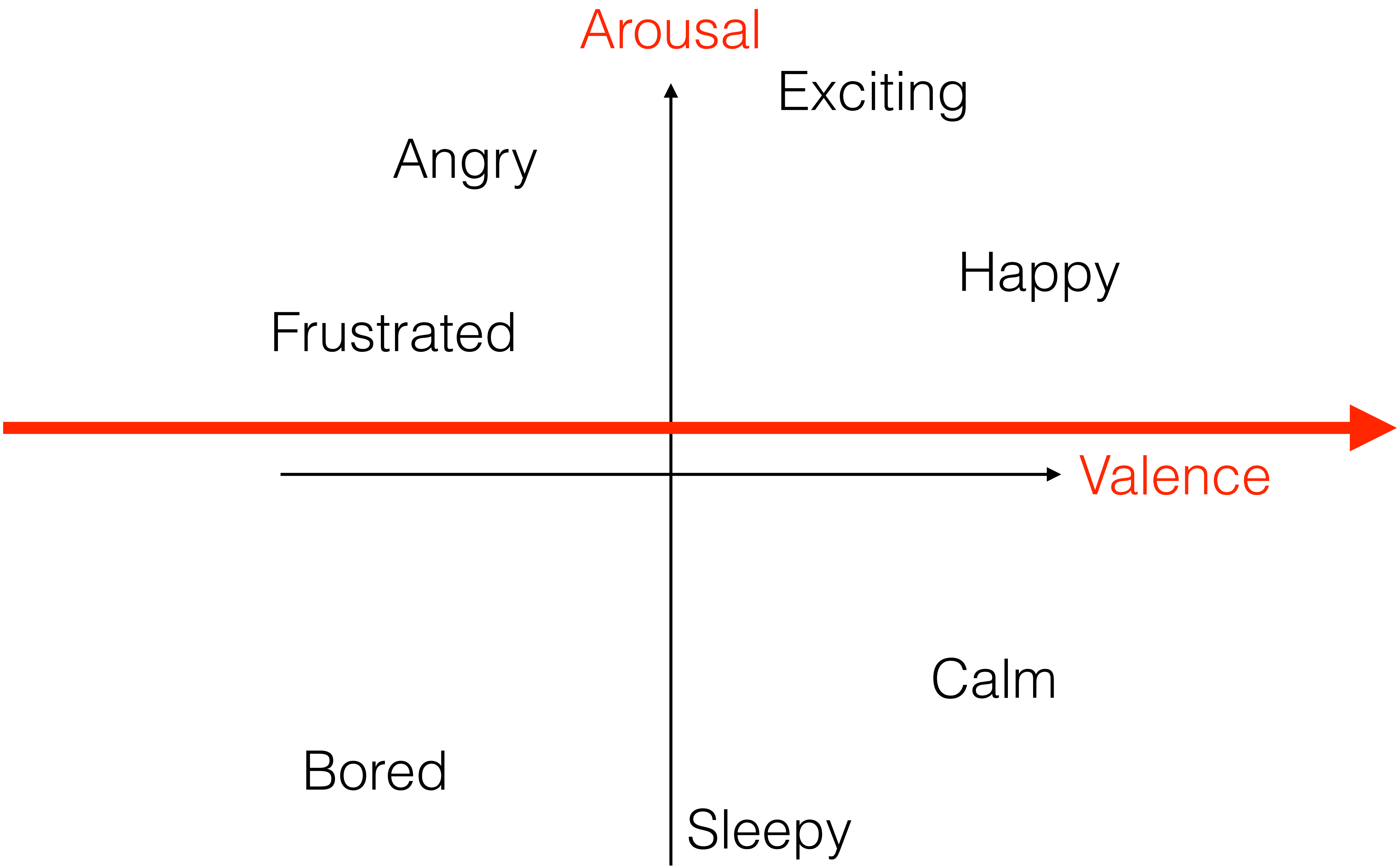
Two-dimensional model of emotion



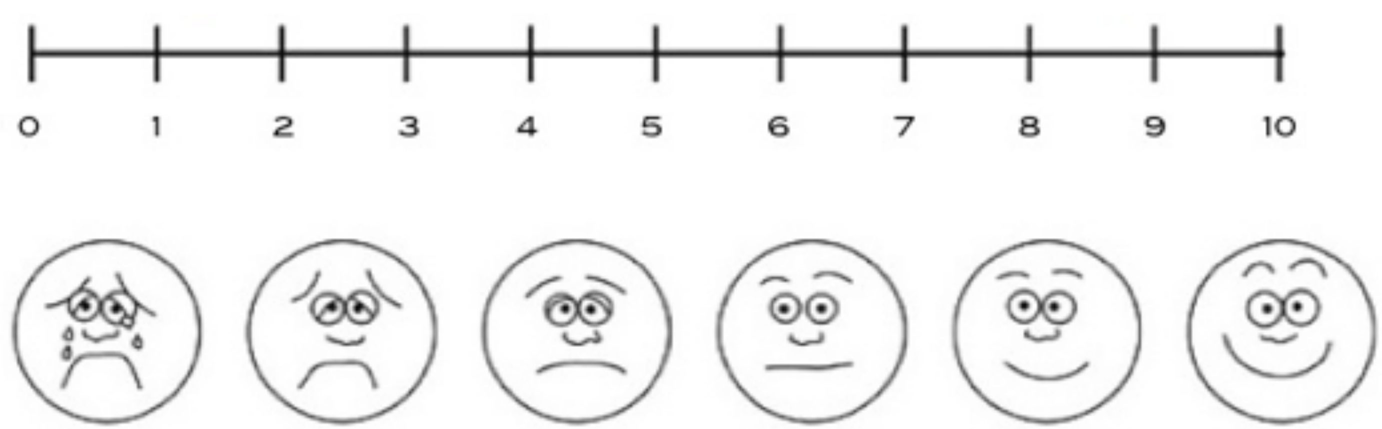
Two-dimensional model of emotion



Two-dimensional model of emotion

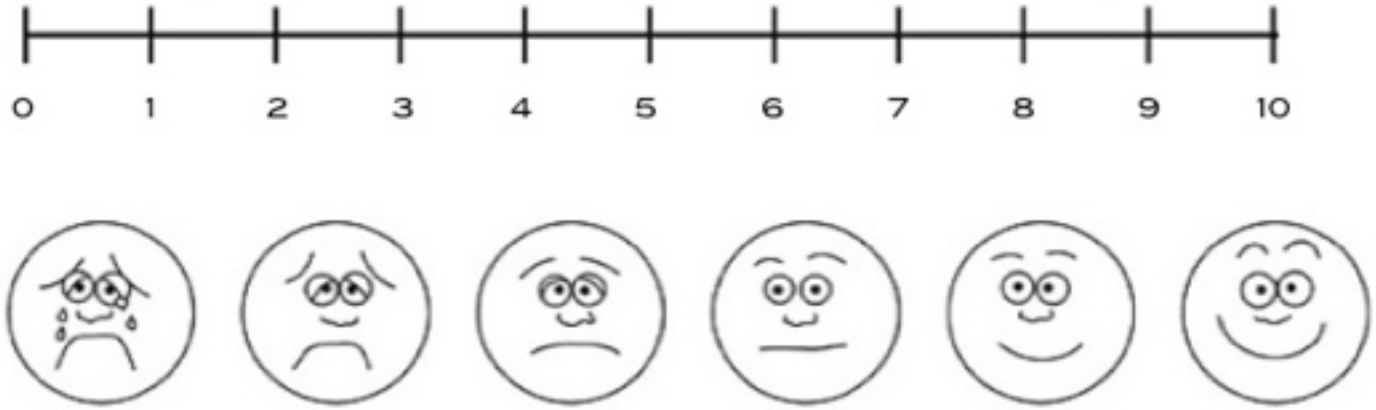


Affective Norms for English Words (ANEW)



Affective Norms for English Words (ANEW)

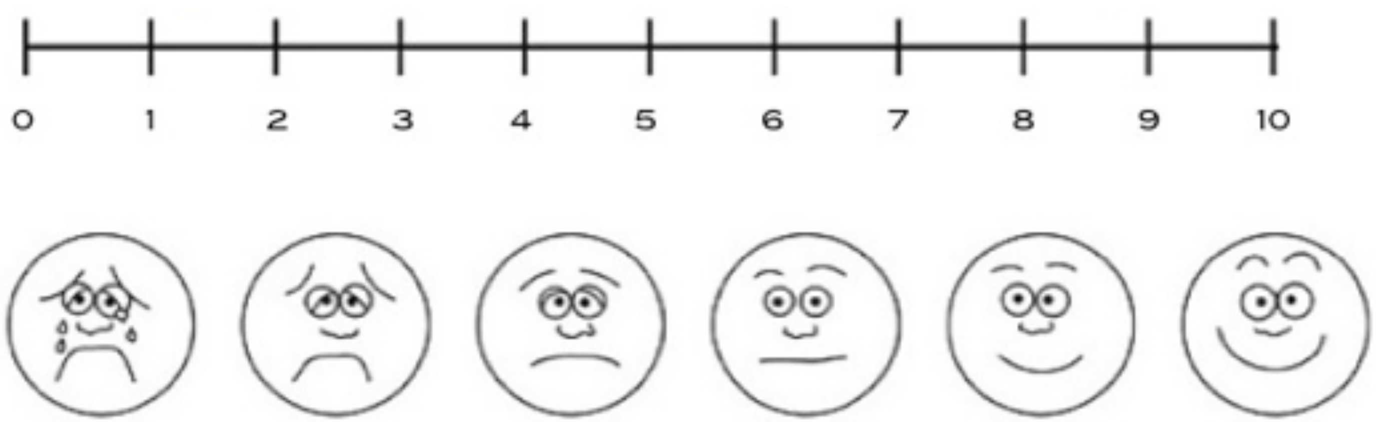
How happy is the word …



Affective Norms for English Words (ANEW)

"laughter"?

How happy is the word …

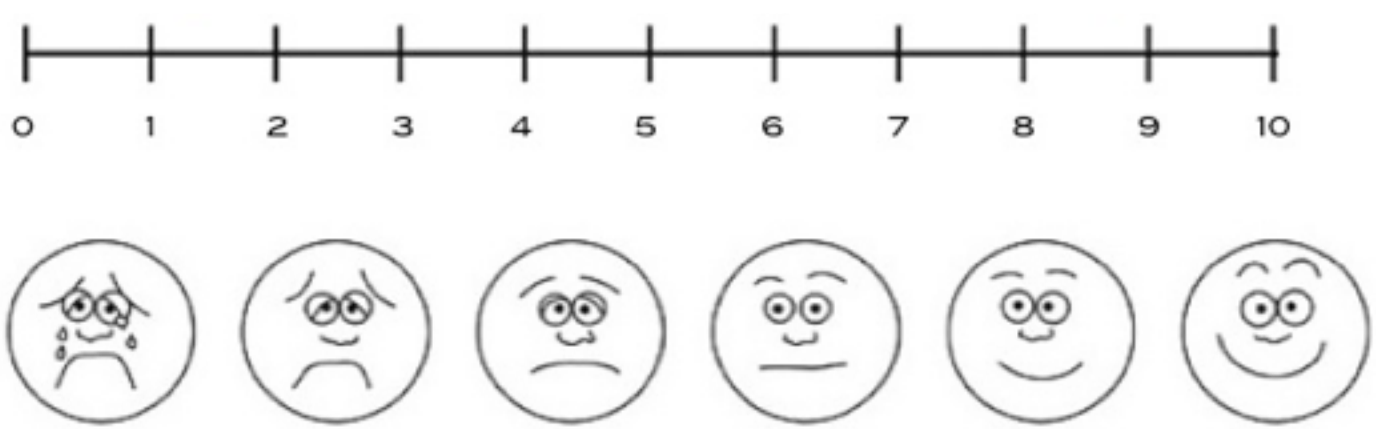


Affective Norms for English Words (ANEW)

"laughter"?

How happy is the word …

"war"?



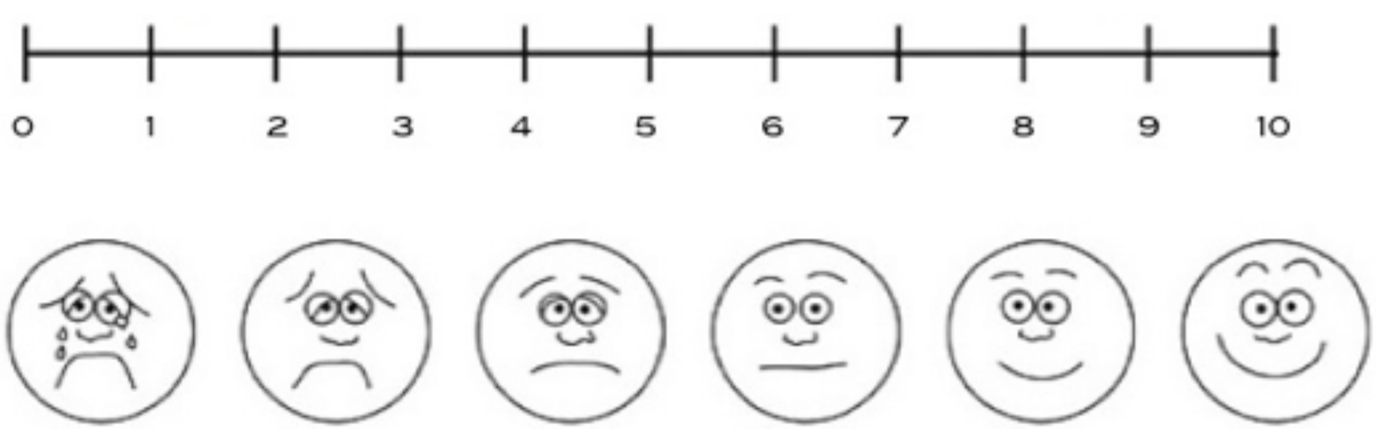
Affective Norms for English Words (ANEW)

“baby”?

“laughter”?

How happy is the word …

“war”?



Affective Norms for English Words (ANEW)

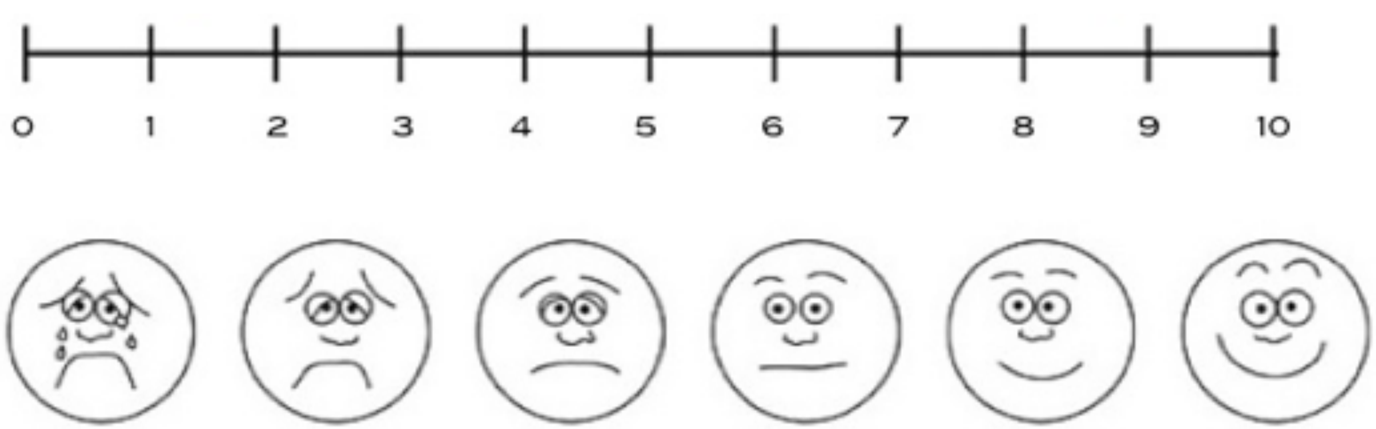
"baby"?

"laughter"?

How happy is the word …

"smile"?

"war"?



$$h_{\text{avg}}(\text{laughter}) = 8.50,$$

$$h_{\text{avg}}(\text{of}) = 4.94,$$

$$h_{\text{avg}}(\text{food}) = 7.44,$$

$$h_{\text{avg}}(\text{vanity}) = 4.30,$$

$$h_{\text{avg}}(\text{reunion}) = 6.96,$$

$$h_{\text{avg}}(\text{greed}) = 3.06,$$

$$h_{\text{avg}}(\text{truck}) = 5.48,$$

$$h_{\text{avg}}(\text{hate}) = 2.34,$$

$$h_{\text{avg}}(\text{the}) = 4.98,$$

$$h_{\text{avg}}(\text{funeral}) = 2.10,$$

and $h_{\text{avg}}(\text{terrorist}) = 1.30.$

$h_{\text{avg}}(\text{laughter}) = 8.50,$

$h_{\text{avg}}(\text{of}) = 4.94,$

$h_{\text{avg}}(\text{food}) = 7.44,$

$h_{\text{avg}}(\text{vanity}) = 4.30,$

$h_{\text{avg}}(\text{reunion}) = 6.96,$

$h_{\text{avg}}(\text{greed}) = 3.06,$

$h_{\text{avg}}(\text{truck}) = 5.48,$

$h_{\text{avg}}(\text{hate}) = 2.34,$

$h_{\text{avg}}(\text{the}) = 4.98,$

$h_{\text{avg}}(\text{funeral}) = 2.10,$

and $h_{\text{avg}}(\text{terrorist}) = 1.30.$

$$v_{\text{text}} = \frac{\sum_{i=1}^{n} v_i f_i}{\sum_{i=1}^{n} f_i}$$

$h_{\text{avg}}(\text{laughter}) = 8.50,$

$h_{\text{avg}}(\text{of}) = 4.94,$

$h_{\text{avg}}(\text{food}) = 7.44,$

$h_{\text{avg}}(\text{vanity}) = 4.30,$

$h_{\text{avg}}(\text{reunion}) = 6.96,$

$h_{\text{avg}}(\text{greed}) = 3.06,$

$h_{\text{avg}}(\text{truck}) = 5.48,$

$h_{\text{avg}}(\text{hate}) = 2.34,$

$h_{\text{avg}}(\text{the}) = 4.98,$

$h_{\text{avg}}(\text{funeral}) = 2.10,$

and $h_{\text{avg}}(\text{terrorist}) = 1.30.$

$$v_{\text{text}} = \frac{\sum_{i=1}^{n} v_i f_i}{\sum_{i=1}^{n} f_i}$$

Valence of a document (sentence)
~ the average of the word sentiment

Maybe not the most accurate nor sophisticated sentiment analysis method,

but

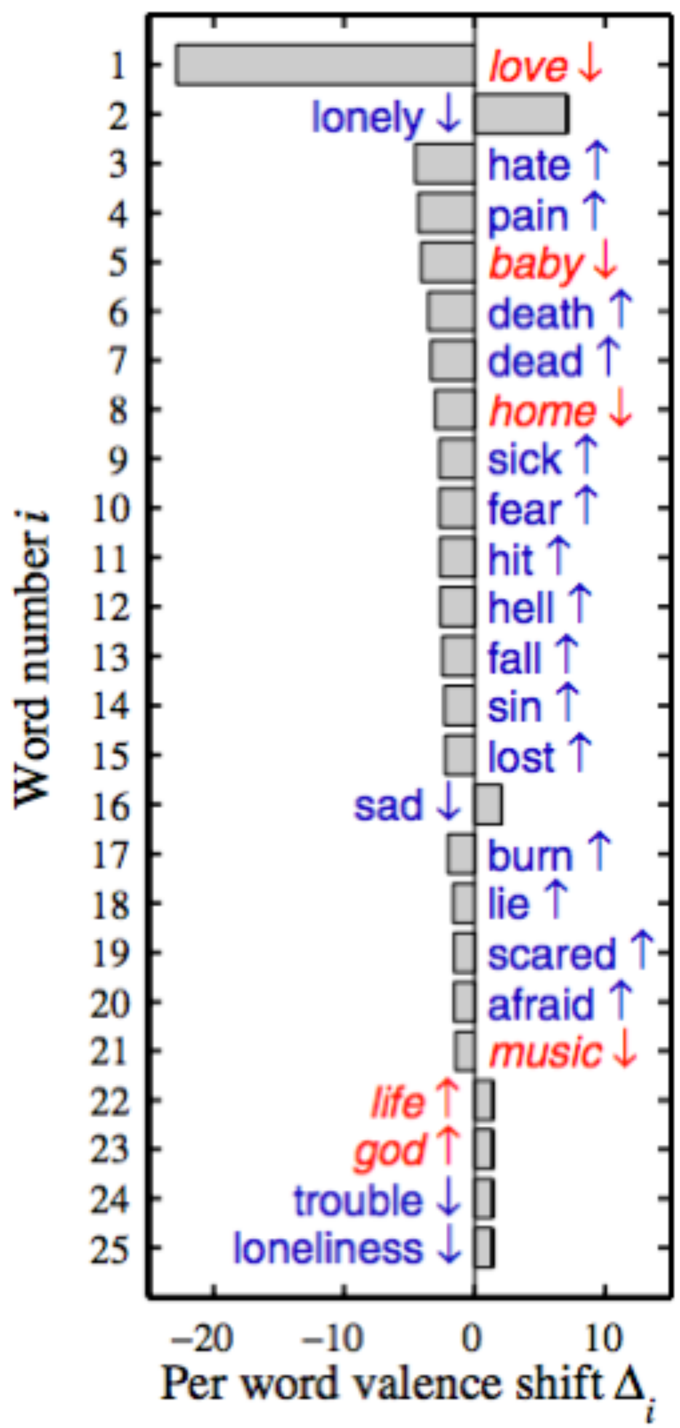
Super *easy*, *scalable*, and ***transparent***

Word-shift diagram

$$100 \cdot \frac{\overbrace{\left(h_i - h^{(\mathrm{ref})}\right)}^{+/-} \overbrace{\left(p_i - p_i^{(\mathrm{ref})}\right)}^{\uparrow/\downarrow}}{\left|h^{(\mathrm{comp})} - h^{(\mathrm{ref})}\right|}$$

PS Dodds, CM Danforth

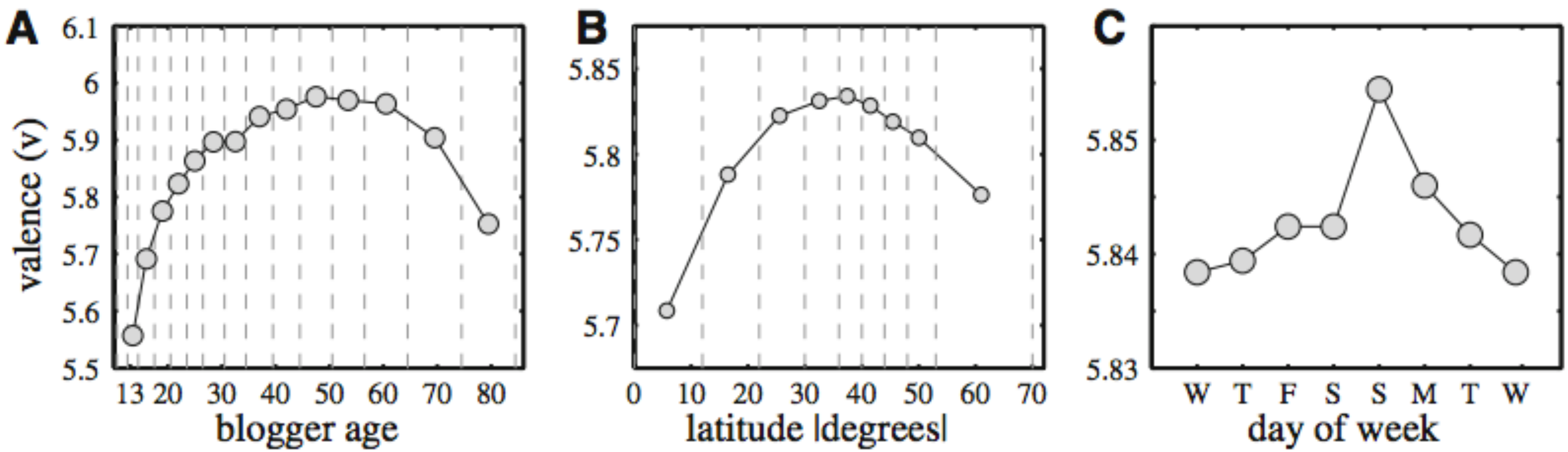
Word-shift diagram



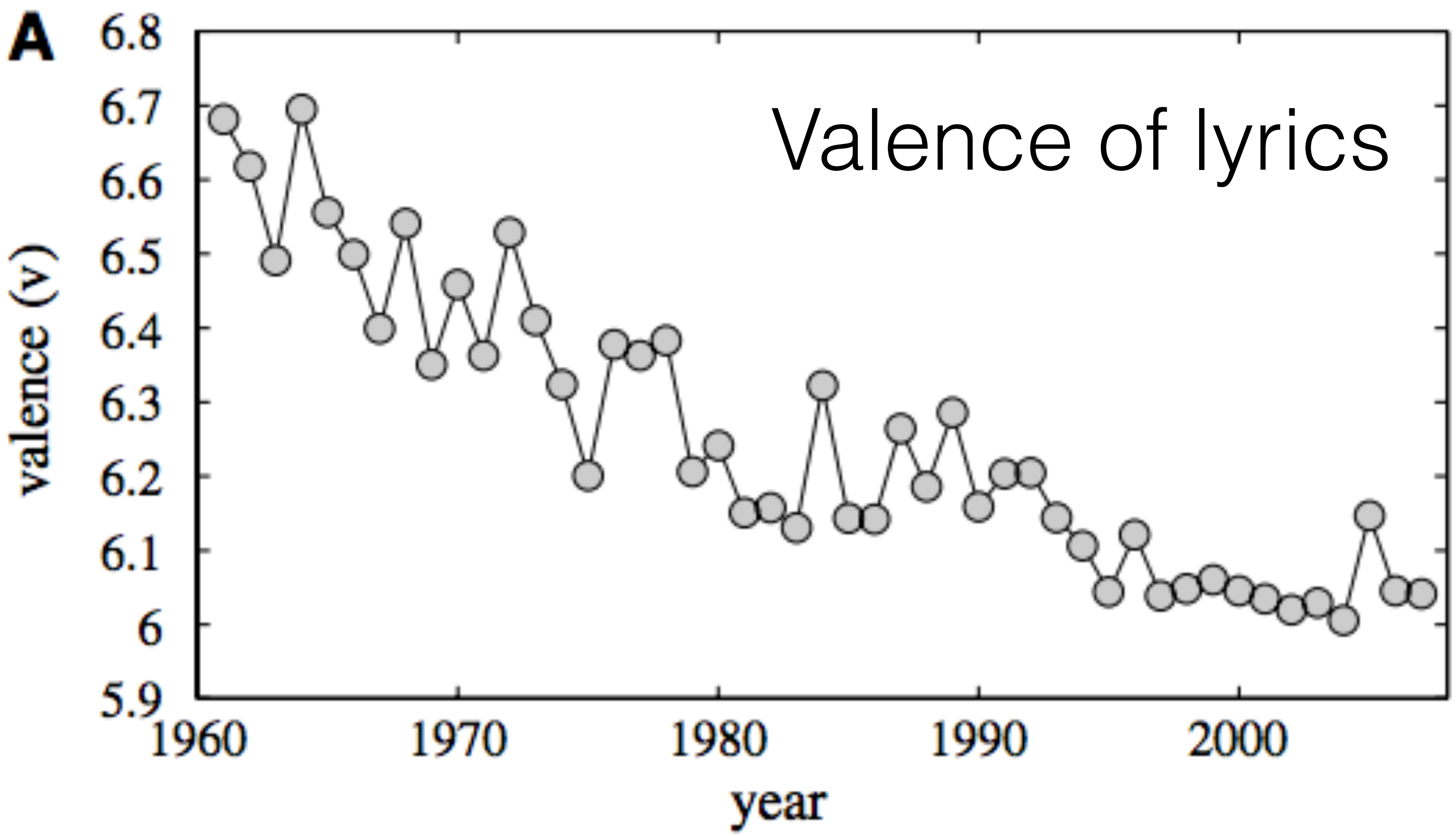
$$100 \cdot \frac{\overbrace{\left(h_i - h^{(\mathrm{ref})}\right)}^{+/-}\overbrace{\left(p_i - p_i^{(\mathrm{ref})}\right)}^{\uparrow/\downarrow}}{\left|h^{(\mathrm{comp})} - h^{(\mathrm{ref})}\right|}$$

PS Dodds, CM Danforth

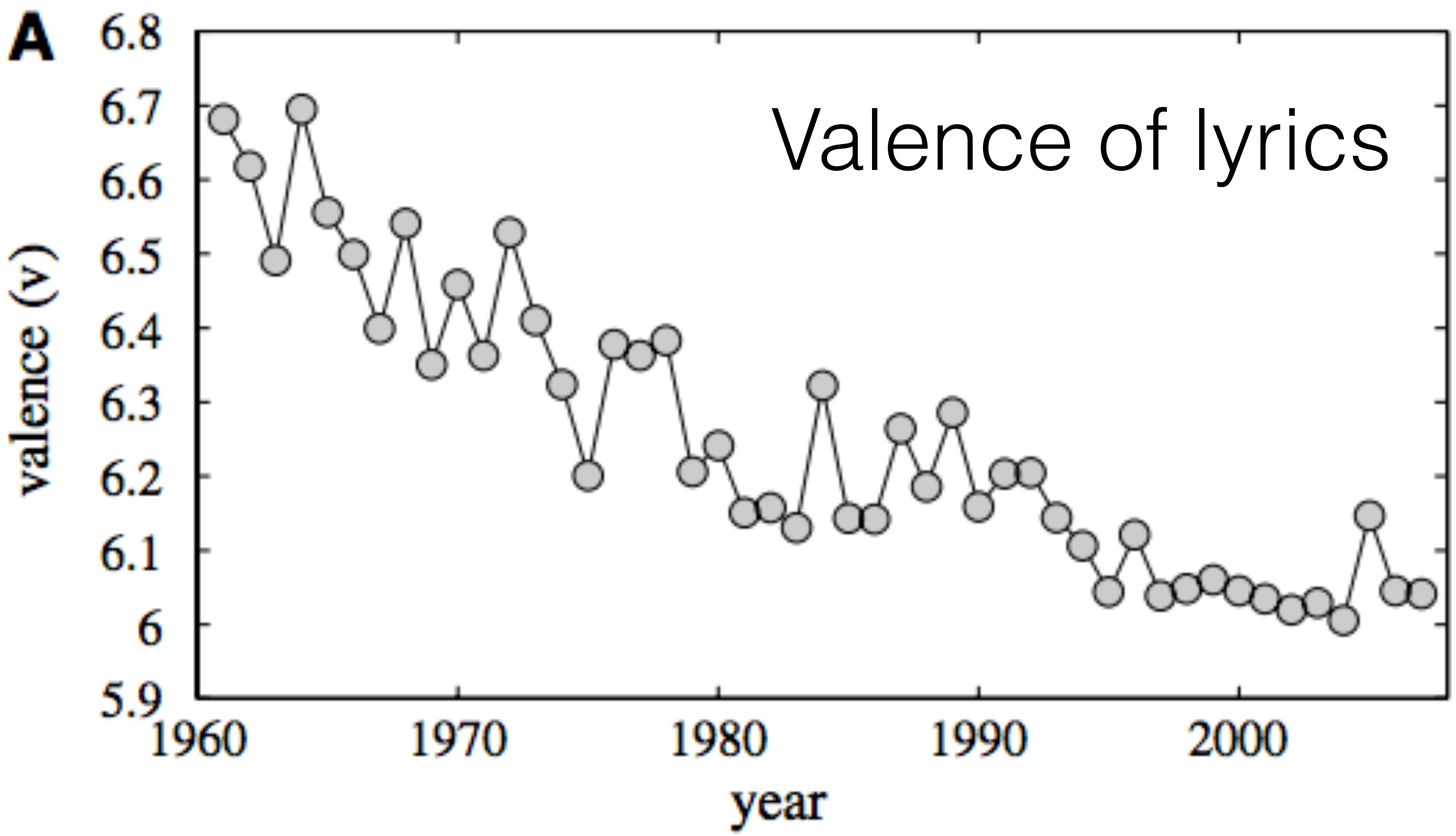
Valence of blog posts

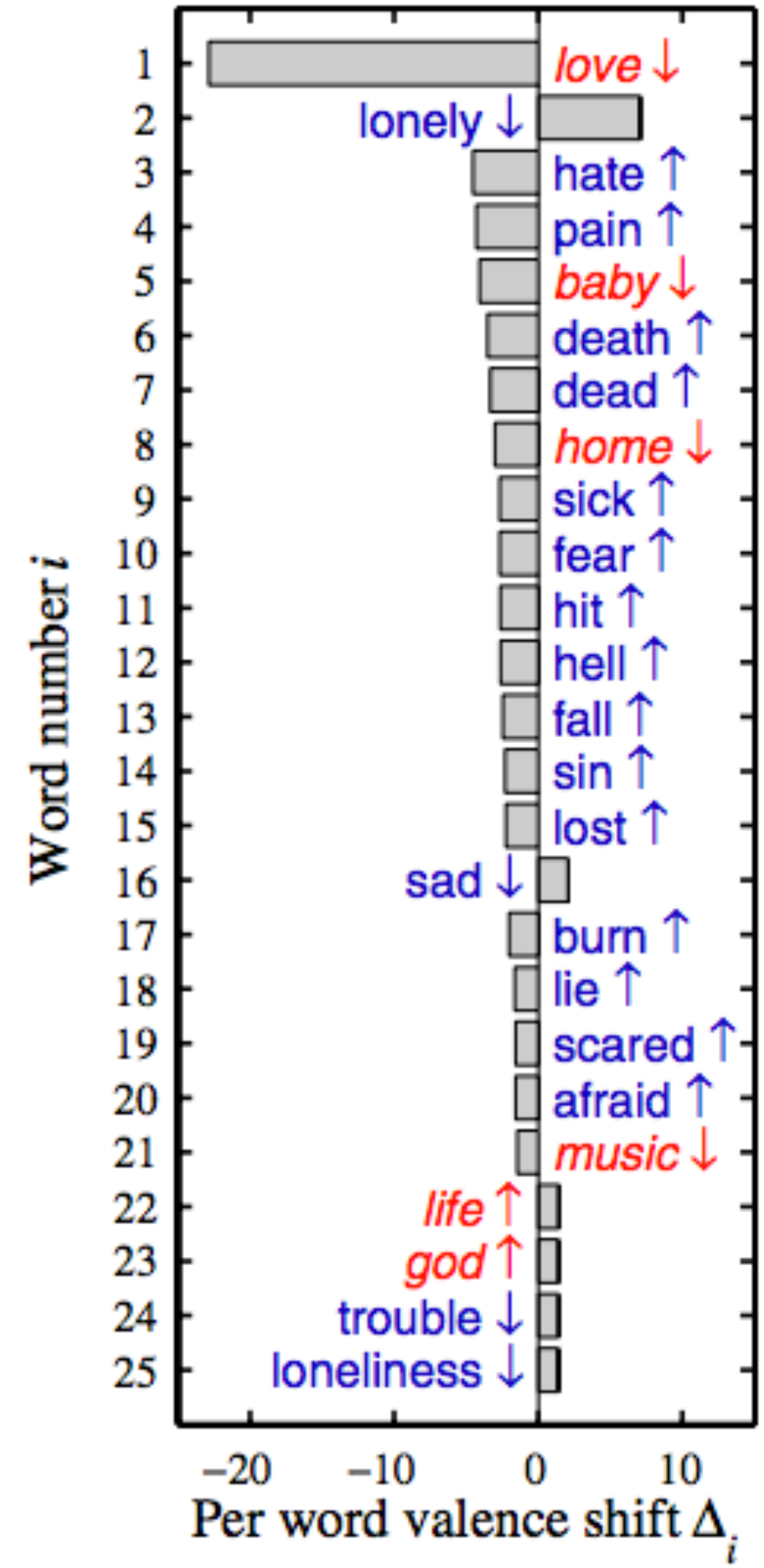


PS Dodds, CM Danforth

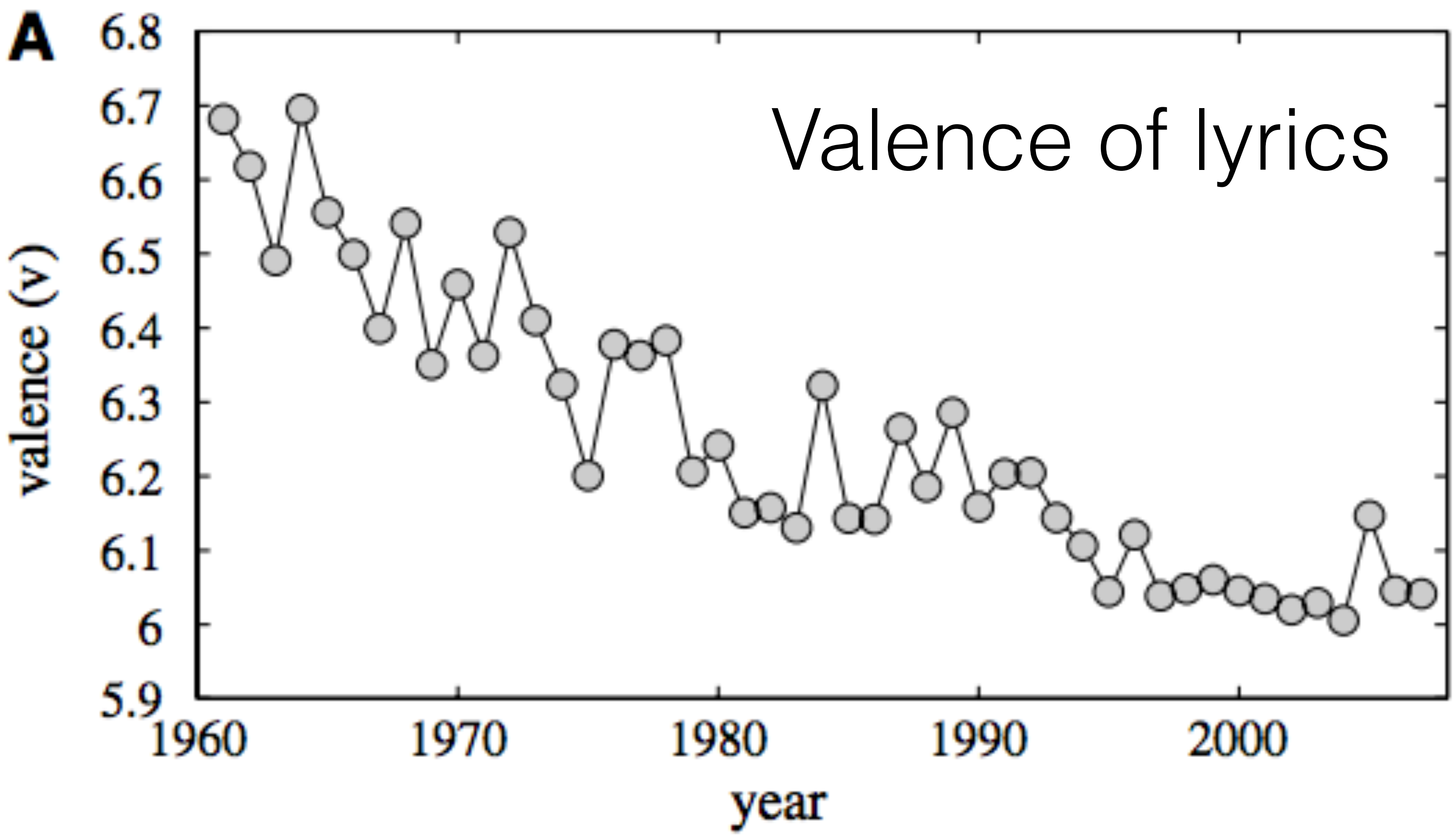


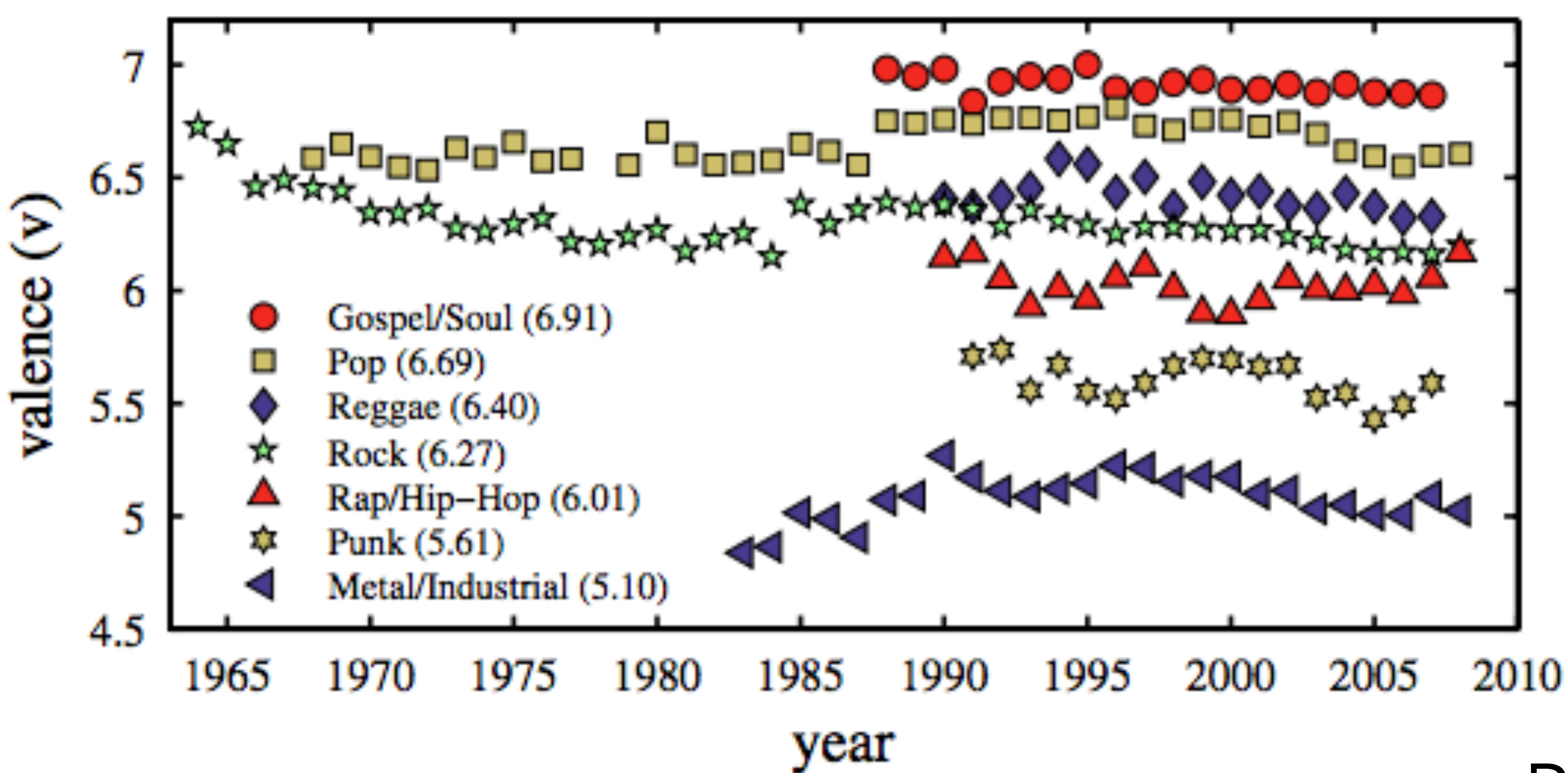
PS Dodds, CM Danforth

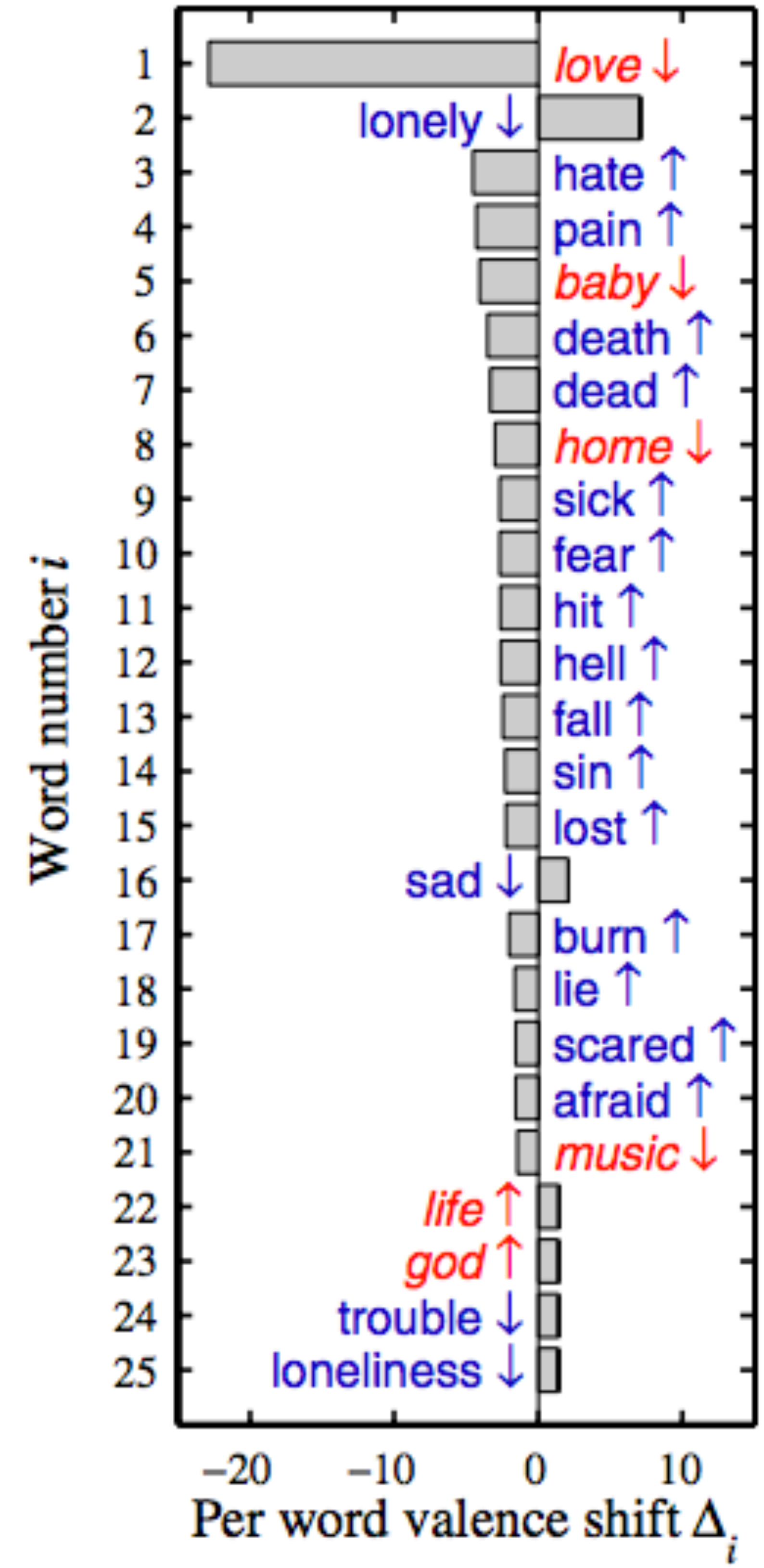




PS Dodds, CM Danforth



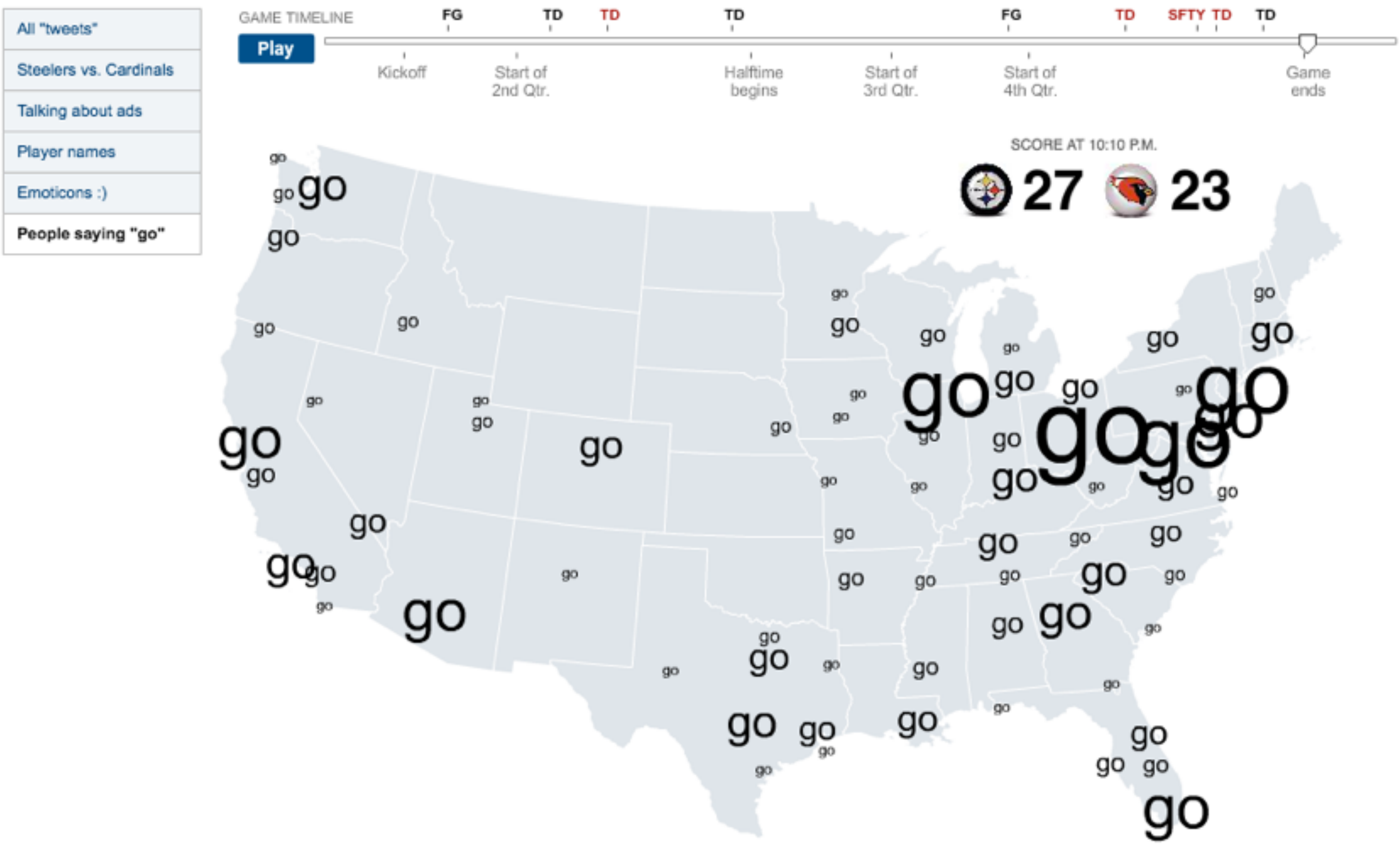




PS Dodds, CM Danforth

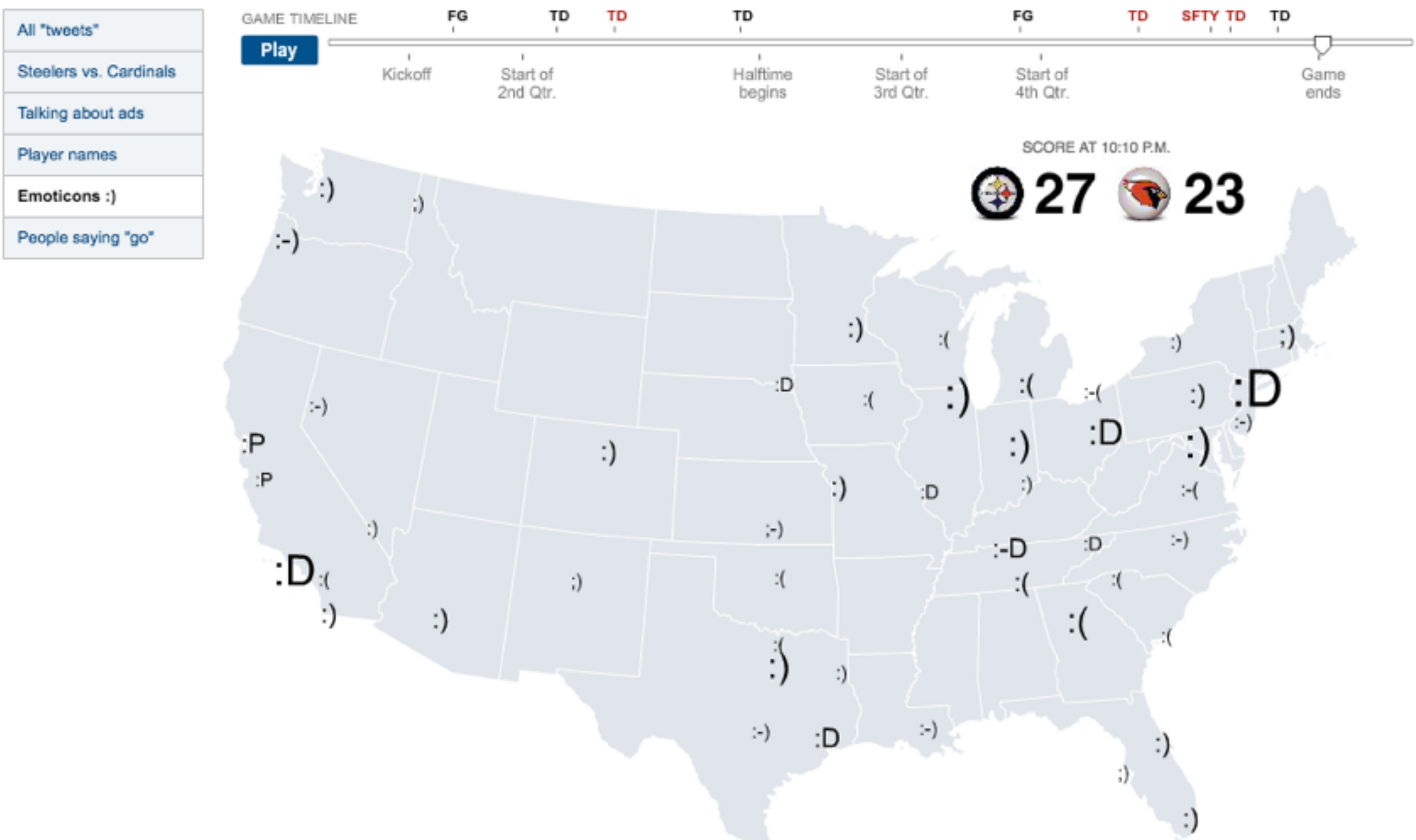
2.

Visualization of Super Bowl tweets (2009)



http://www.nytimes.com/interactive/2009/02/02/sports/20090202_superbowl_twitter.html?emc=eta3

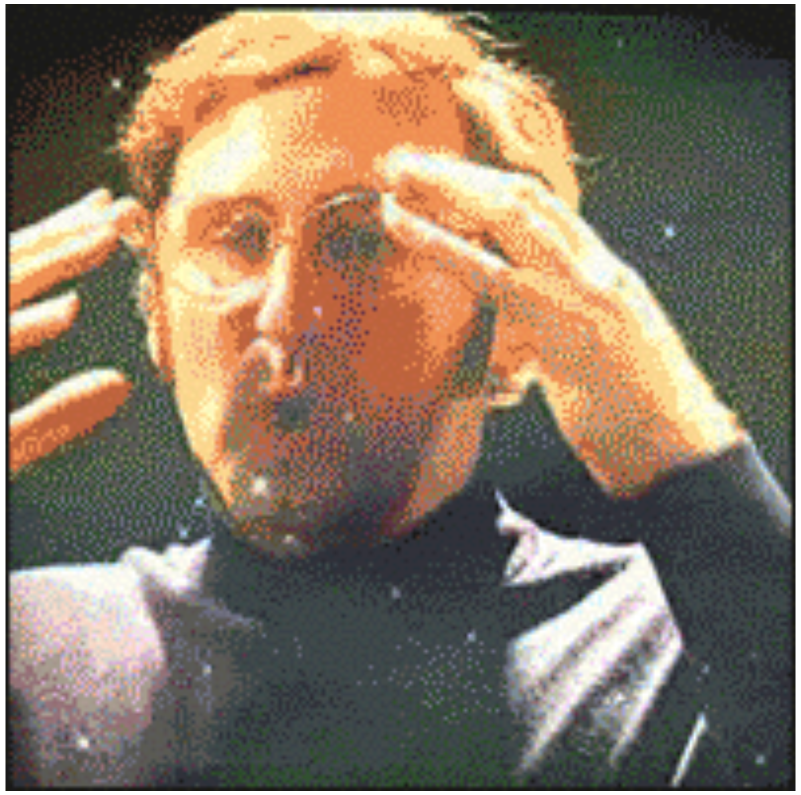
Visualization of Super Bowl tweets (2009)



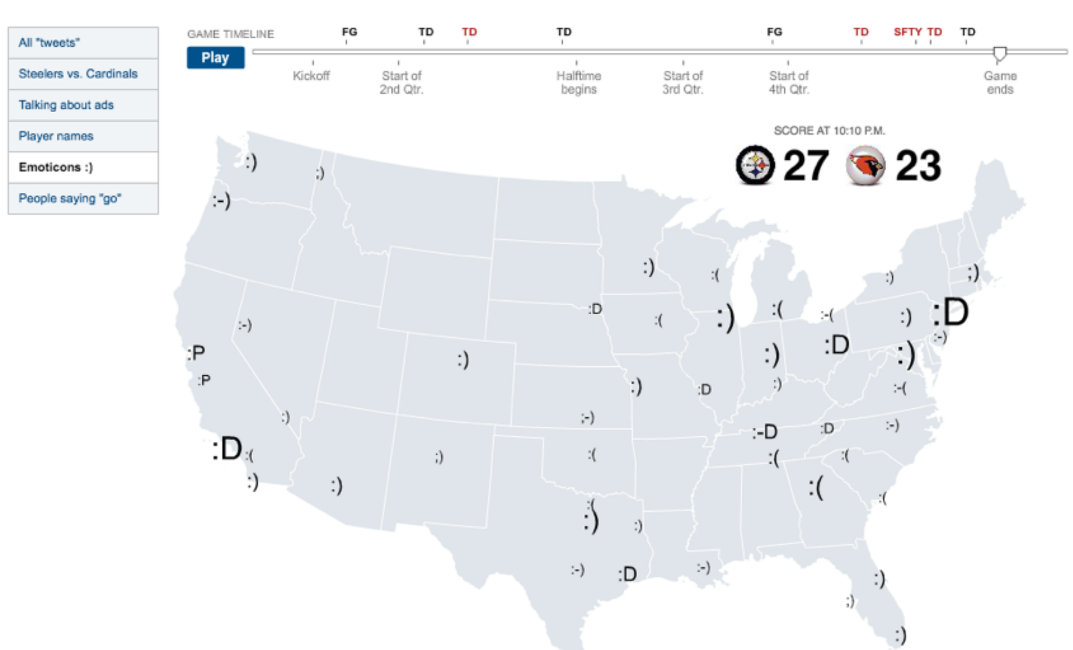
http://www.nytimes.com/interactive/2009/02/02/sports/20090202_superbowl_twitter.html?emc=eta3

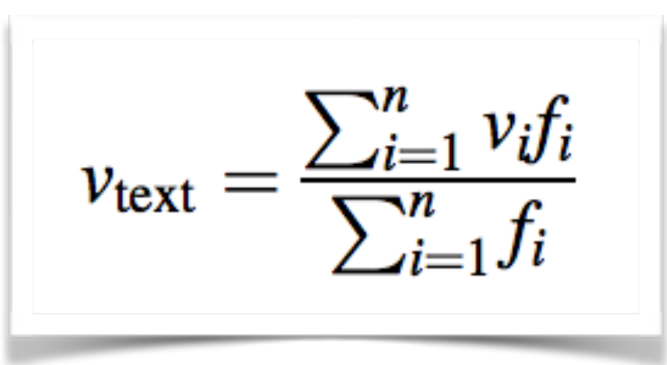
(simple) text analysis
+ rich text data

(simple) text analysis
+ rich text data



Can we map daily **mood** of the whole nation through Twitter?





$$v_{\text{text}} = \frac{\sum_{i=1}^{n} v_i f_i}{\sum_{i=1}^{n} f_i}$$

Alan Mislove

Sune Lehmann

+ James Bagrow, JP Onnela, J Niels Rosenquist

Can we map daily **mood** of the whole nation through Twitter?

300 million tweets (Sep 2006 - Aug 2009)

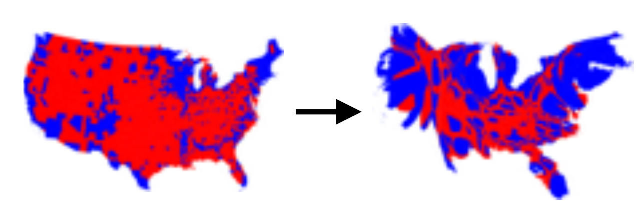
+

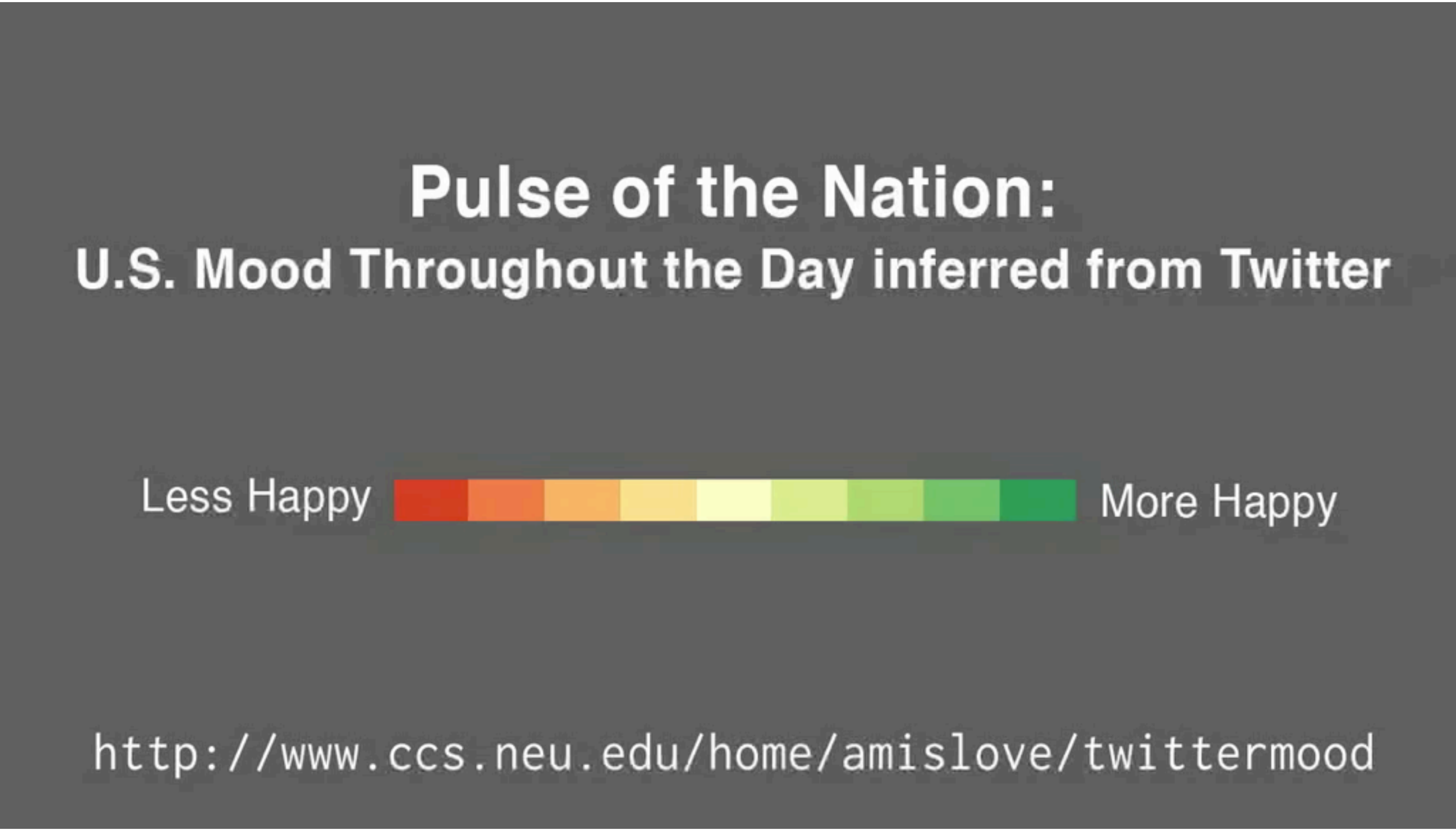
ANEW sentiment

+

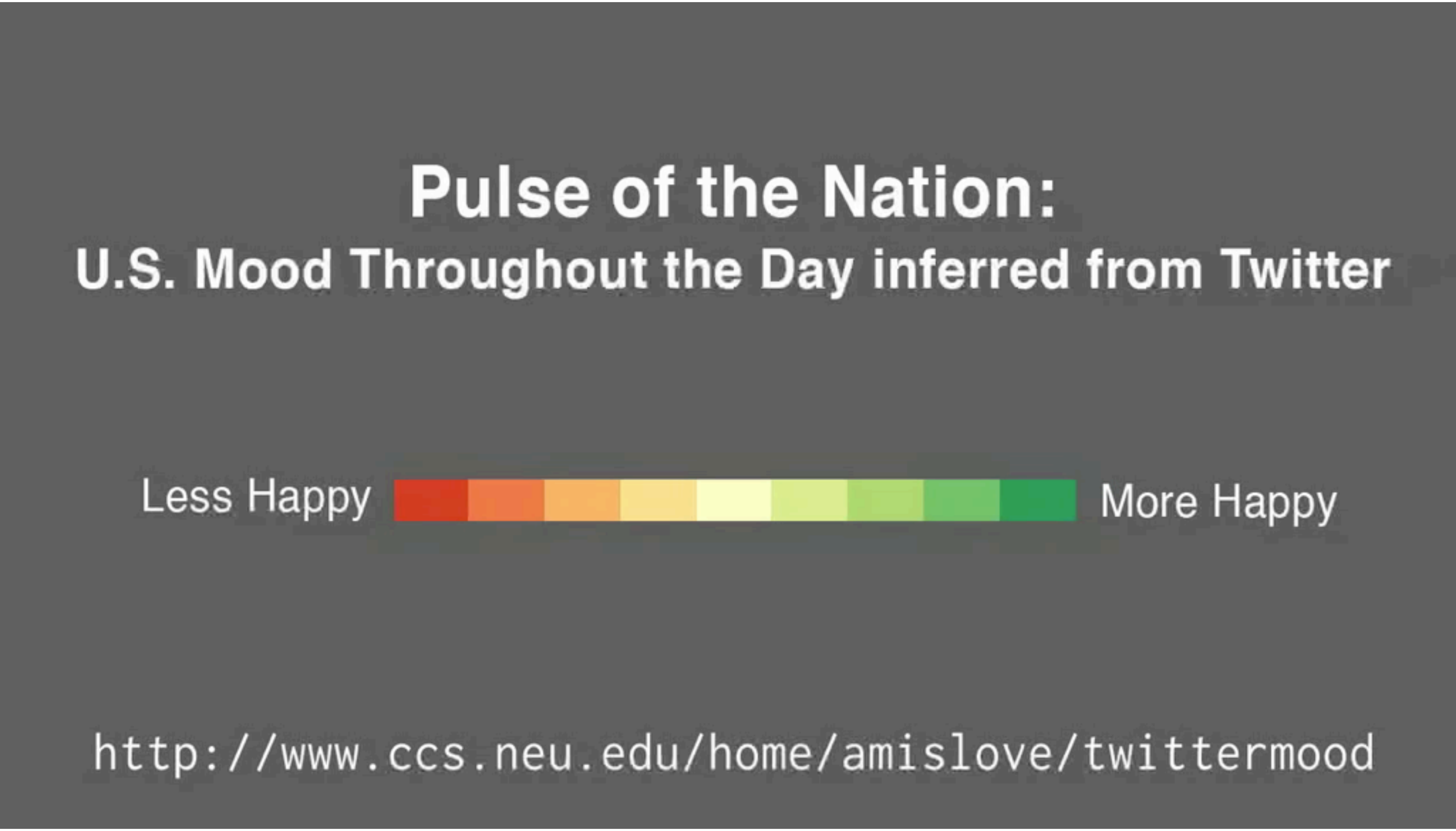
Simple geocoding to states

+

Cartogram 



http://www.ccs.neu.edu/home/amislove/twittermood/



http://www.ccs.neu.edu/home/amislove/twittermood/

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

Scott A. Golder* and Michael W. Macy

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

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

Experimental psychologists have repeatedly demonstrated that positive and negative affect are independent dimensions. Positive affect (PA) includes enthusiasm, delight, activeness, and alertness, whereas negative affect (NA) includes distress, fear, anger, guilt, and disgust (*6*). Thus, low PA indicates the absence of positive feelings, not the presence of negative feelings.

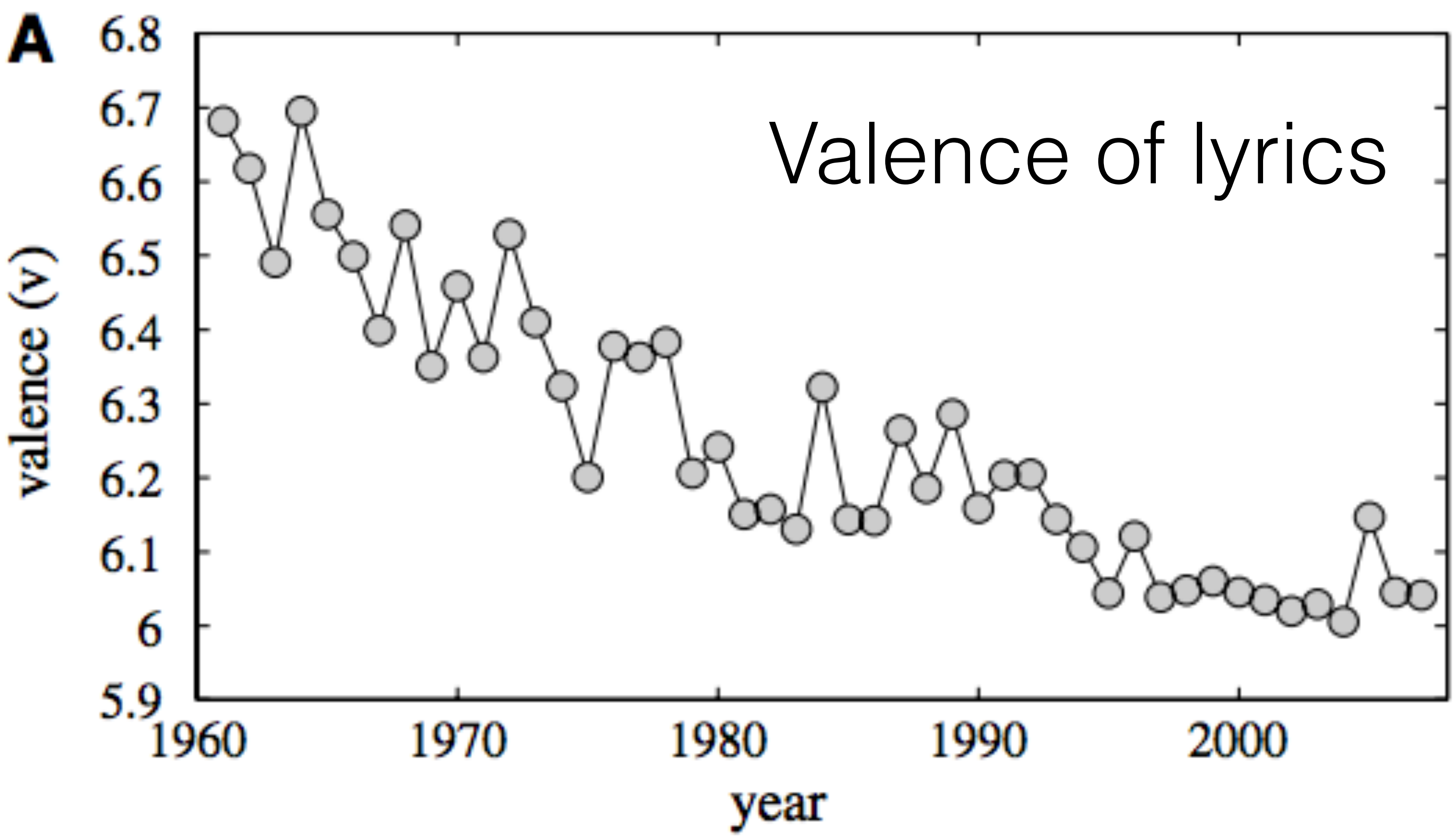
Laboratory studies have shown that diurnal mood swings reflect endogenous circadian rhythms interacting with the duration of prior wakefulness or sleep. The circadian component corresponds to change in core body temperature, which is lowest at the end of the night and peaks during late afternoon. The sleep-dependent component is elevated at waking and declines throughout the day (*7*). Other studies have variously observed a single PA peak 8 to 10 hours after waking (*8*), a

Department of Sociology, Cornell University, Ithaca, NY 14853, USA.

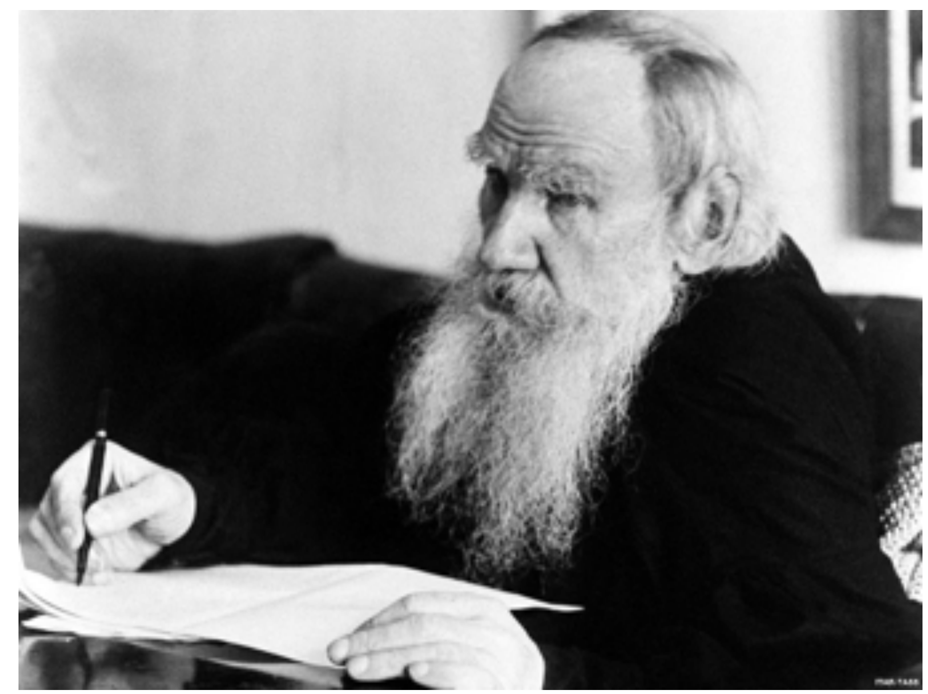
*To whom correspondence should be addressed. E-mail: sag262@cornell.edu

#1: Lyrics and Emotion

Coming back to music,



Interesting because music evokes strong emotion



"Music is the **shorthand** of emotion."

Music != lyrics

Music != lyrics

Melody

Music != lyrics

Melody

Harmony

Music != lyrics

Melody

Rhythm

Harmony

Music != lyrics

Melody

Timbre

Rhythm

Harmony

Music != lyrics

Melody

Timbre

Rhythm

Harmony

. . .

Music != lyrics

They are not "texts" :(

Melody
Timbre
Rhythm
Harmony
...

For instance, harmony has power.

C Major

C Minor

For instance, harmony has power.

C Major

C Minor

For instance, harmony has power.

C Major

C Minor

:)

For instance, harmony has power.

C Major

C Minor

:)

For instance, harmony has power.

C Major

C Minor

:)

For instance, harmony has power.

C Major

:)

C Minor

:(

https://www.youtube.com/watch?v=uNaQE9K3eO0

https://www.youtube.com/watch?v=uNaQE9K3eO0

Analyzing Lyrics & Chords together?

Nakul Dhande

"Maybe we can use sentiment analysis of **lyrics** in combination with **chords**…"

Difficult: what kinds of emotion (or ‘meaning’) does this chord convey?

Difficult: what kinds of emotion (or ‘meaning’) does this chord convey?

↓

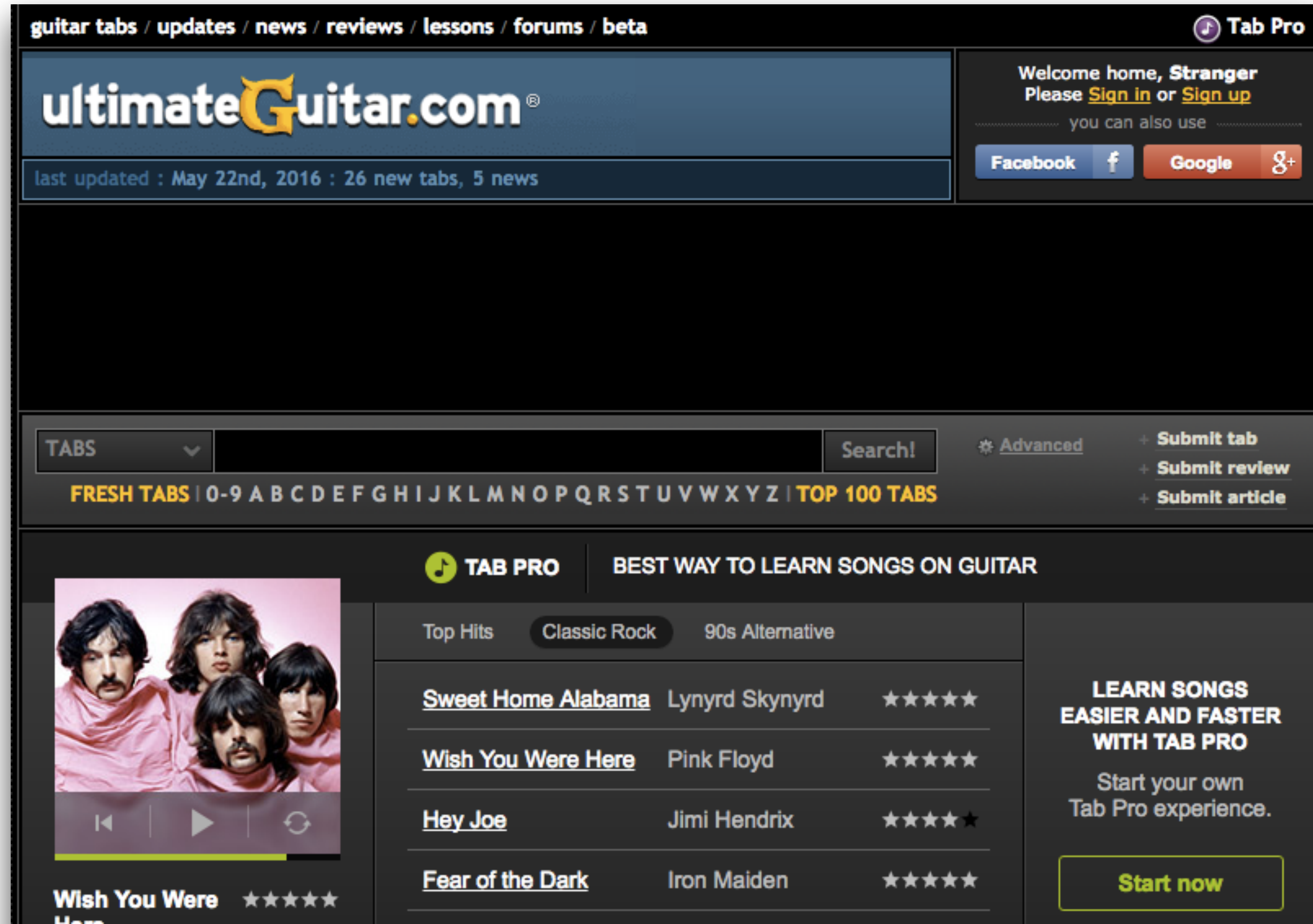
Somewhat feasible: what kinds of words are associated with this chord? How happy are they?

Difficult: what kinds of emotion (or 'meaning') does this chord convey?

Somewhat feasible: what kinds of words are associated with this chord? How happy are they?

"Did you lose the keys here?"
"No, but the light is much better here."

Data (~100k songs)



Hallelujah Chords version 9

by Leonard Cohen

x 121
excellent!

5 comments
read comments

53,351 views
381 this week

add to favorites
launch autoscroll
font size
transpose
display chords
guitar tuner

Contributor Pencom
Difficulty intermediate

```
#-------------------------------PLEASE NOTE-------------------------------------#
# This file is the author's own work and represents their interpretation of the #
# song. You may only use this file for private study, scholarship, or research. #
#-------------------------------------------------------------------------------#
Hallelujah chords
Leonard Cohen (live London version)

C Am  C Am  C Am  C Am

         C                 Am
Now I've heard there was a secret chord,
     C                    Am
that David played, and it pleased the Lord
    F                G               C       G
But you don't really care for music, do you?
   C                  F           G
It goes like this the fourth, the fifth,
    Am              F
the minor fall, the major lift
    G               Em          Am
The baffled king composing Hallelujah
     F
Hallelujah
     Am
Hallelujah
     F
Hallelujah
     C G C    G
```


```
         C                 Am
Now I've heard there was a secret chord,
     C                    Am
that David played, and it pleased the Lord
    F                G                C       G
But you don't really care for music, do you?
   C                  F           G
It goes like this the fourth, the fifth,
    Am              F
the minor fall, the major lift
    G               Em          Am
The baffled king composing Hallelujah
     F
Hallelujah
     Am
```

```
         C                 Am
Now I've heard there was a secret chord,
     C                   Am
that David played, and it pleased the Lord
    F               G              C       G
But you don't really care for music, do you?
   C                F          G
It goes like this the fourth, the fifth,
    Am             F
the minor fall, the major lift
    G              Em          Am
The baffled king composing Hallelujah
     F
Hallelujah
     Am
```

```
         C                 Am
Now I've heard there was a secret chord,
     C                    Am
that David played, and it pleased the Lord
    F                C               C       G
But you don't really care for music, do you?
   C                  F           G
It goes like this the fourth, the fifth,
    Am              F
the minor fall, the major lift
    G               Em          Am
The baffled king composing Hallelujah
     F
Hallelujah
     Am
```

```
         C                 Am
Now I've heard there was a secret chord,
     C                    Am
that David played, and it pleased the Lord
    F                C               G
But you don't really care for music, do you?
   C                  F           G
It goes like this the fourth, the fifth,
    Am              F
the minor fall, the major lift
    G               Em          Am
The baffled king composing Hallelujah
     F
Hallelujah
     Am
```

```
         C                Am
Now I've heard there was a secret chord,
     C                    Am
that David played, and it pleased the Lord
    F                G               C       G
But you don't really care for music, do you?
   C                  F           G
It goes like this the fourth, the fifth,
    Am              F
the minor fall, the major lift
    G               Em          Am
The baffled king composing Hallelujah
     F
Hallelujah
     Am
```

```
         C                 Am
Now I've heard there was a secret chord,
     C                    Am
that David played, and it pleased the Lord
    F                C               G
But you don't really care for music, do you?
   C                  F           G
It goes like this the fourth, the fifth,
    Am              F
the minor fall, the major lift
    G               Em          Am
The baffled king composing Hallelujah
     F
Hallelujah
     Am
```

Each chord (type): a “document”

Chords	Accompanied lyrics
C **Major**	Now I’ve heard that there was a
A **Minor**	Secret chord, that
C **Major**	David Played, and it
A **Minor**	Pleased the Lord but
…	…

Sentiment analysis using

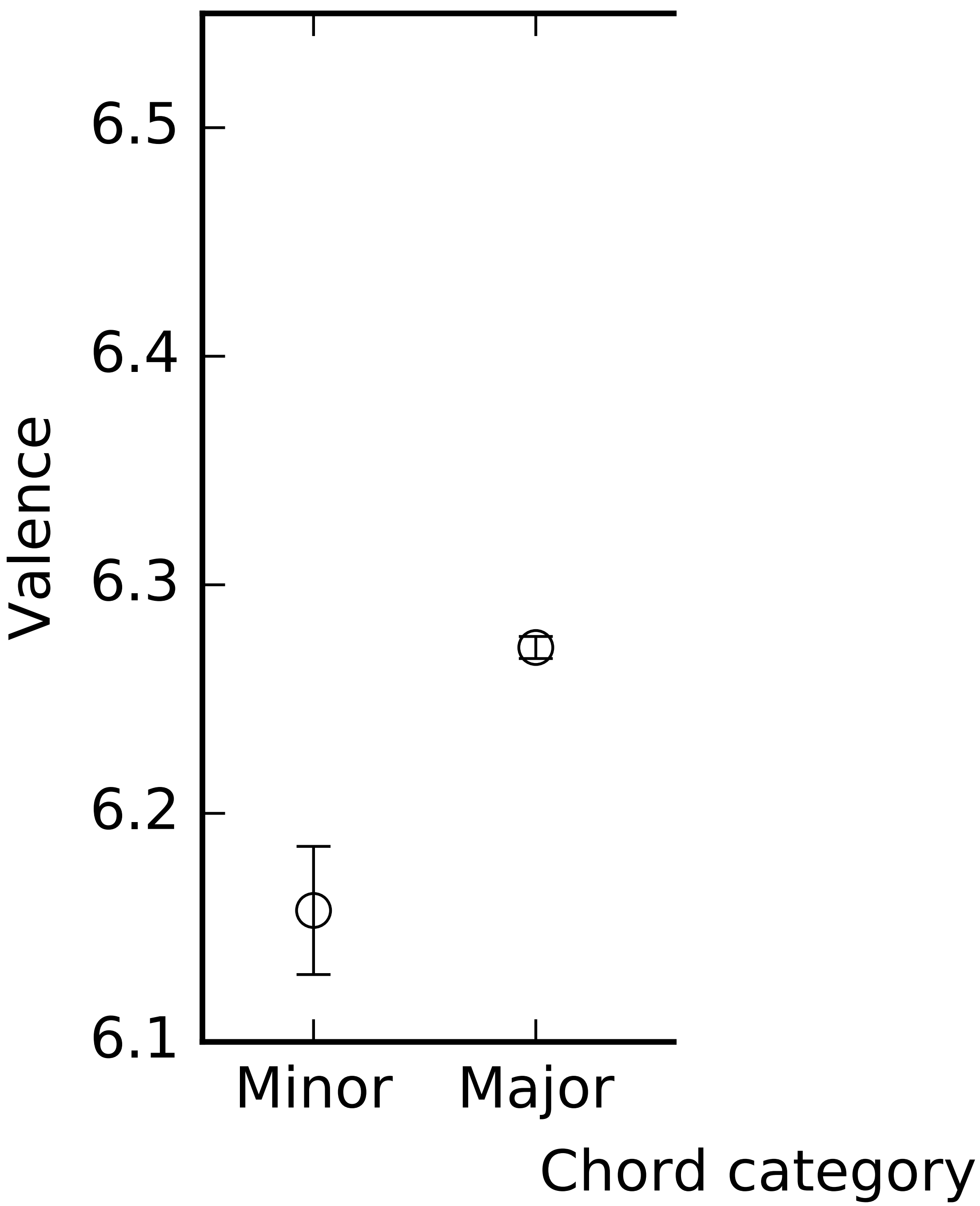
“**labMT**”
(language assessment by
Mechanical Turk)

Danforth, Dodds, et al.

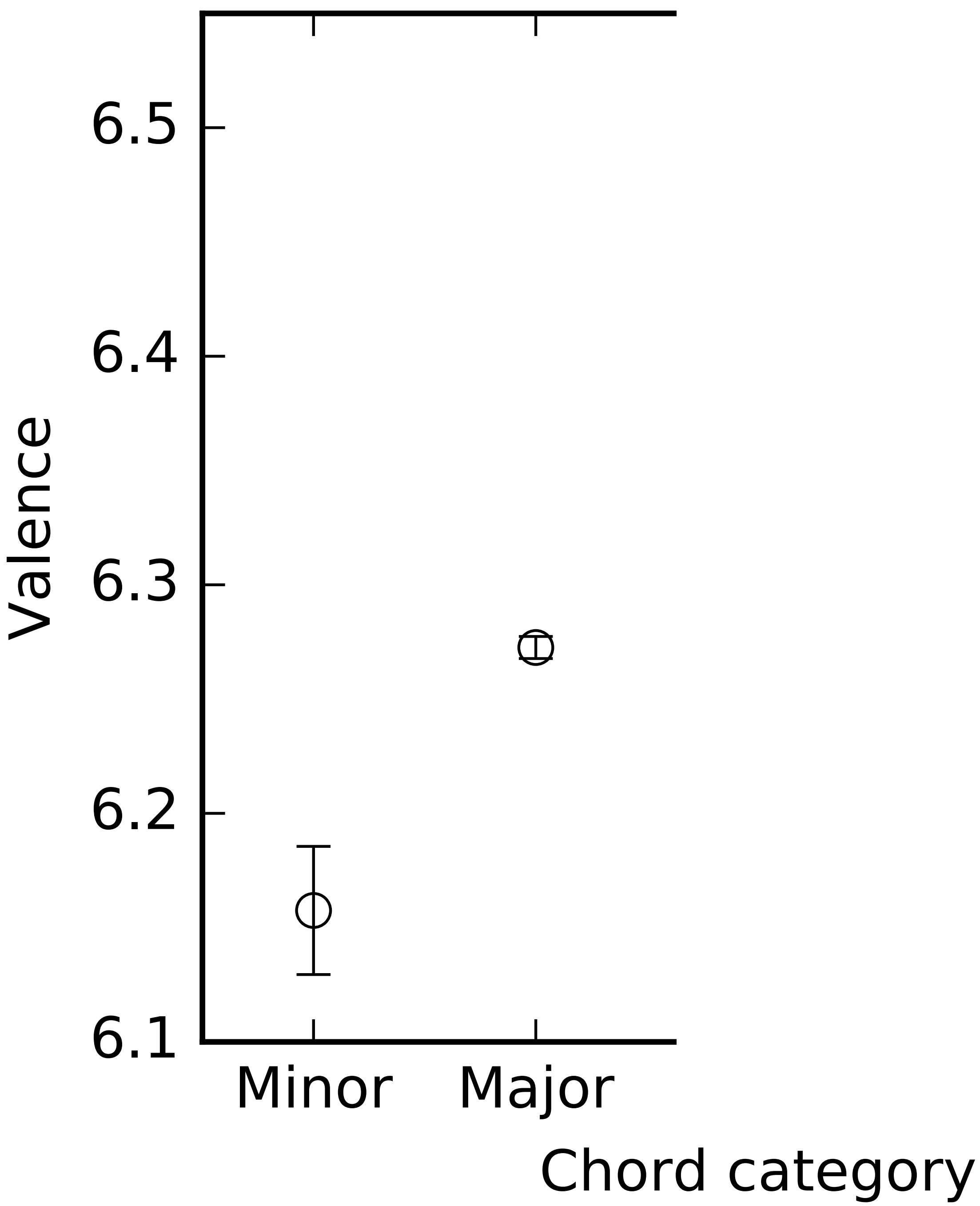
Are Major chords happier than Minor chords?

Are Major chords associated with happier words than Minor chords?

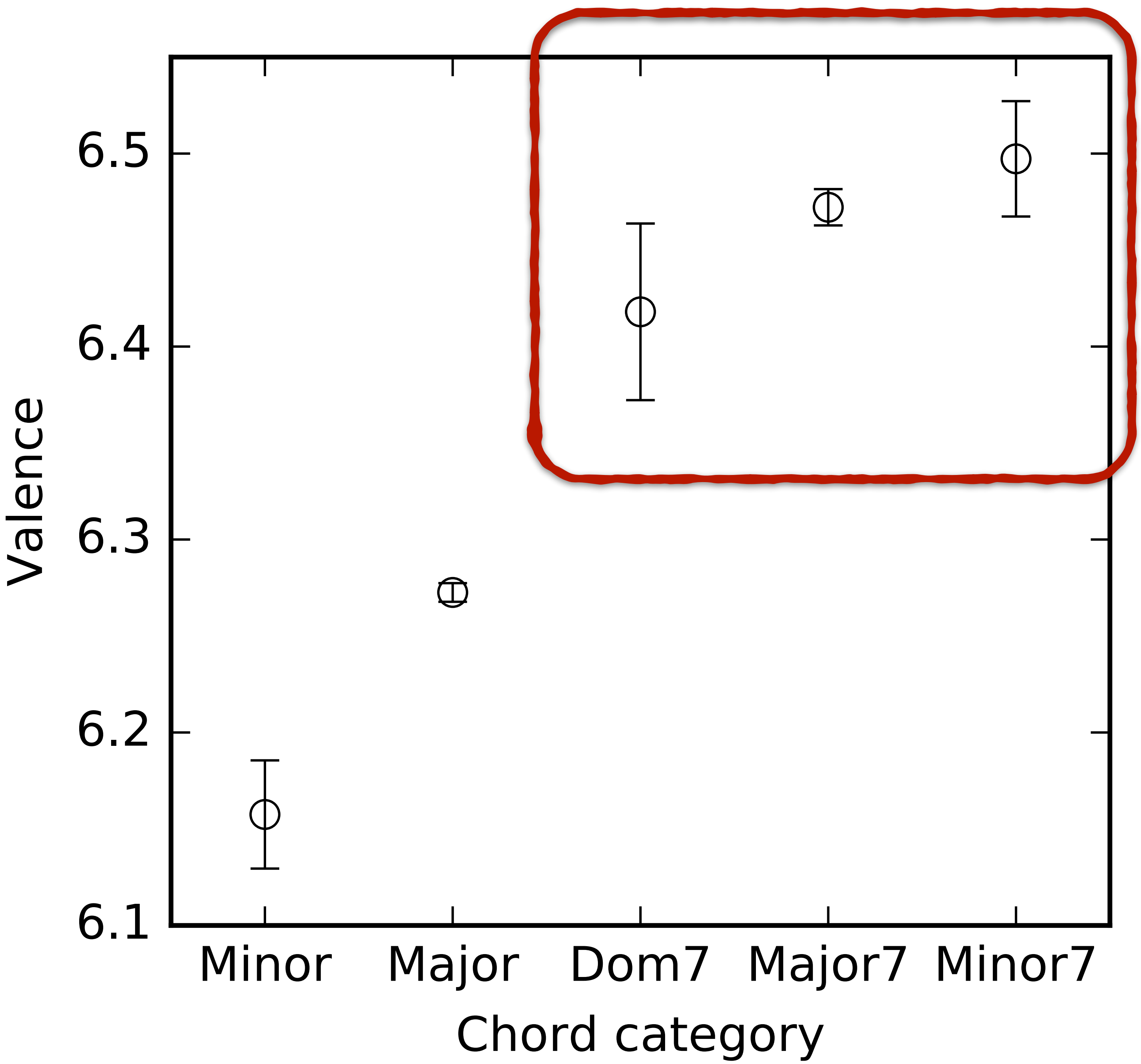
Are Major chords associated with happier words than Minor chords?



Are Major chords associated with happier words than Minor chords?

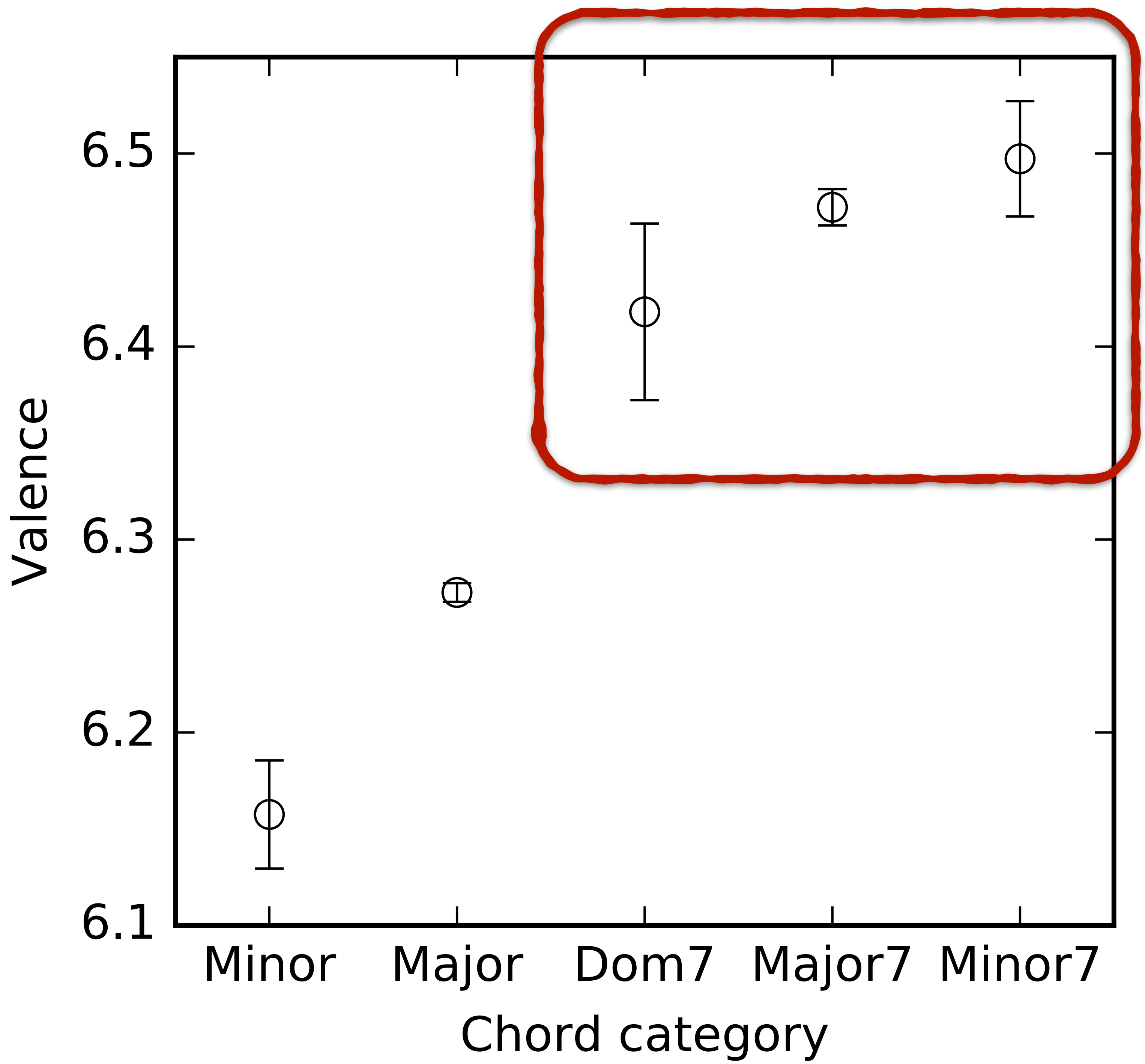


Yes! But,


6.5
6.4
6.3
6.2
6.1
Valence
Minor
Major
Dom7
Major7
Minor7
Chord category

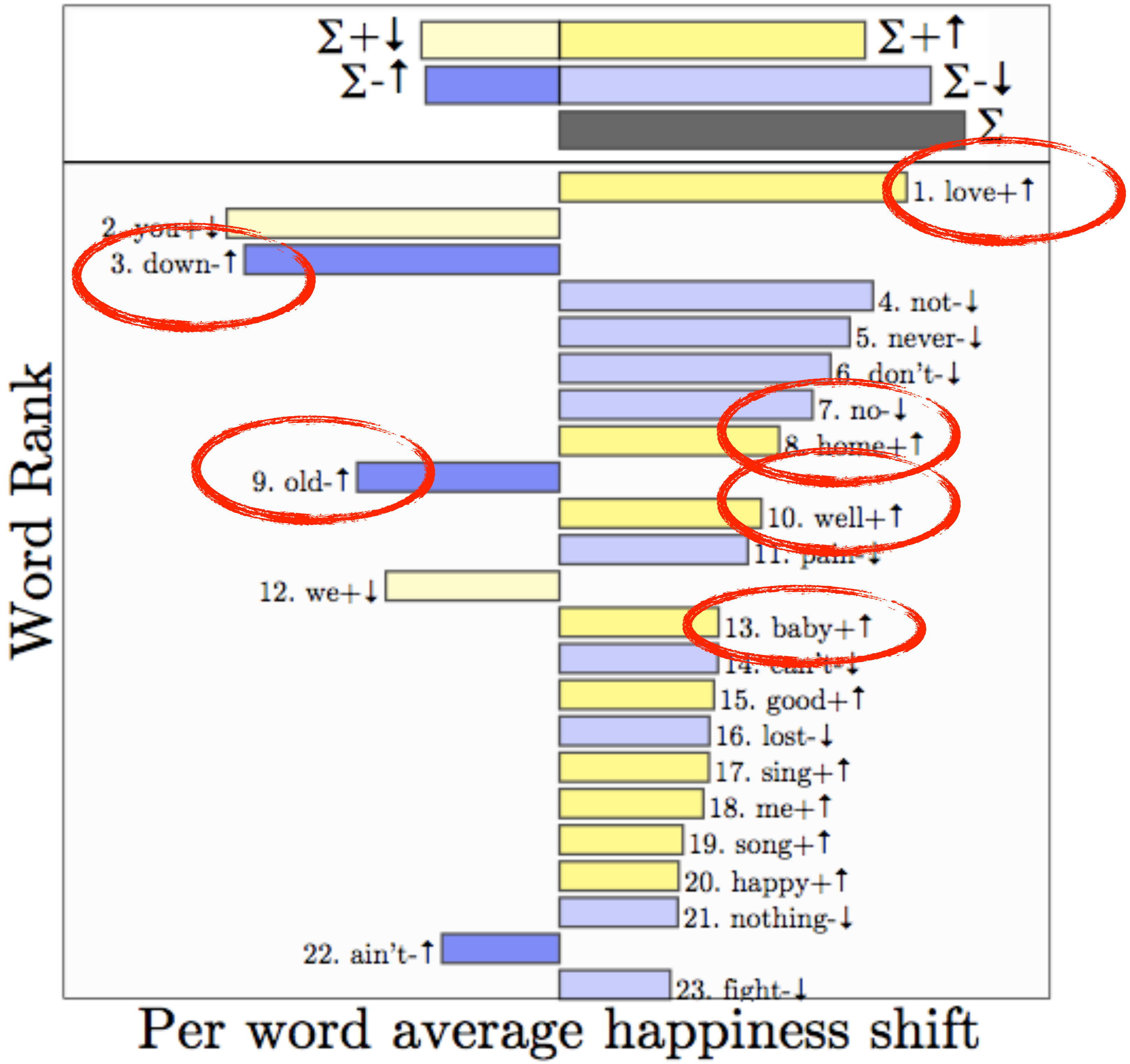
Yes! But,

7th chords are associated with happier words than Major chords



Yes! But,

Major, compared with Minor

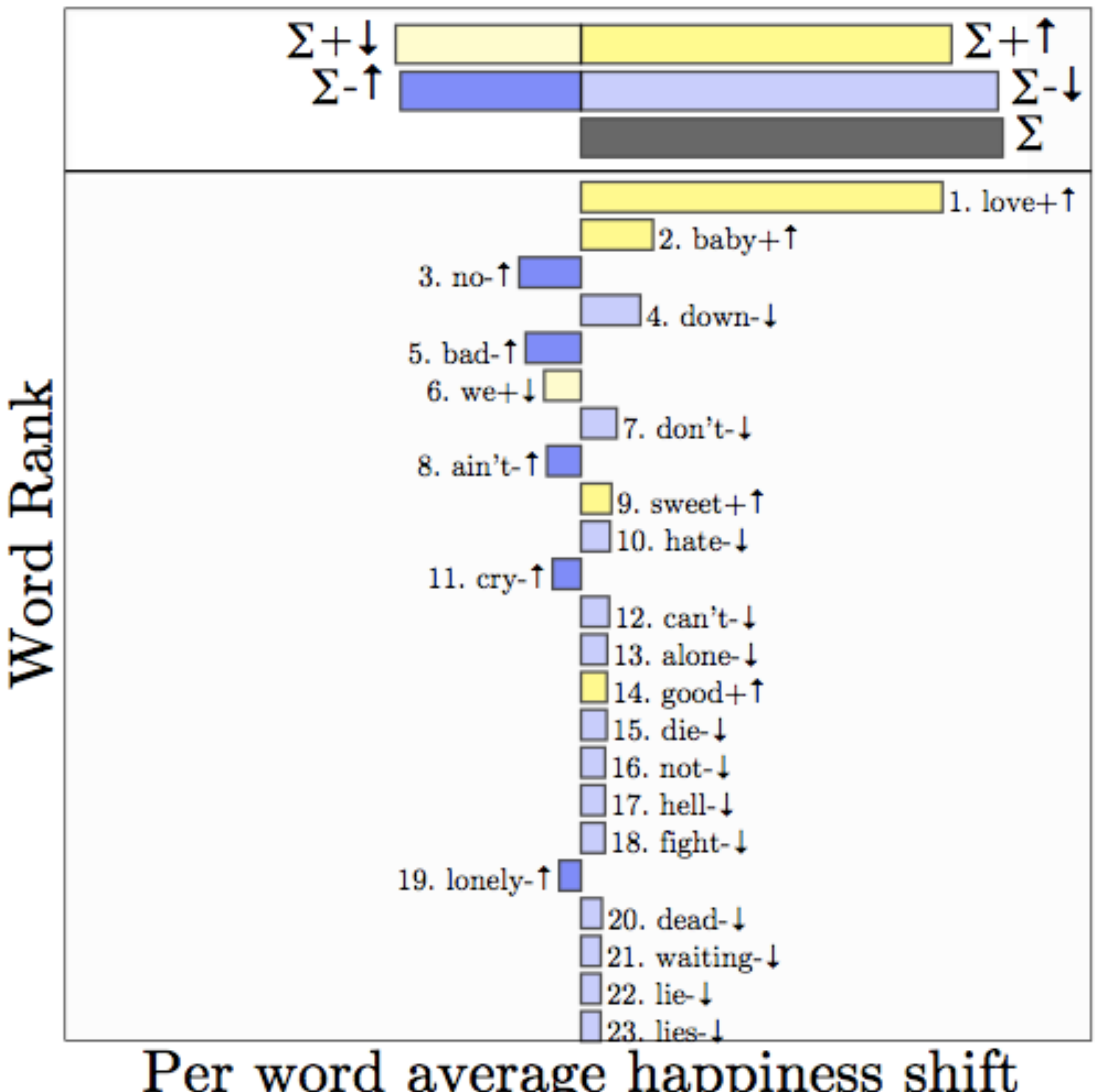


love, down, **home**, old, **well**, **baby**, **good**, **sing**, **song**, **happy**

vs.

you, **not**, **never**, **don't**, **no**, pain, we, **can't**, lost, **nothing**

7ths vs. Major



love, baby, no, bad, ain't, sweet, cry, good, lonely

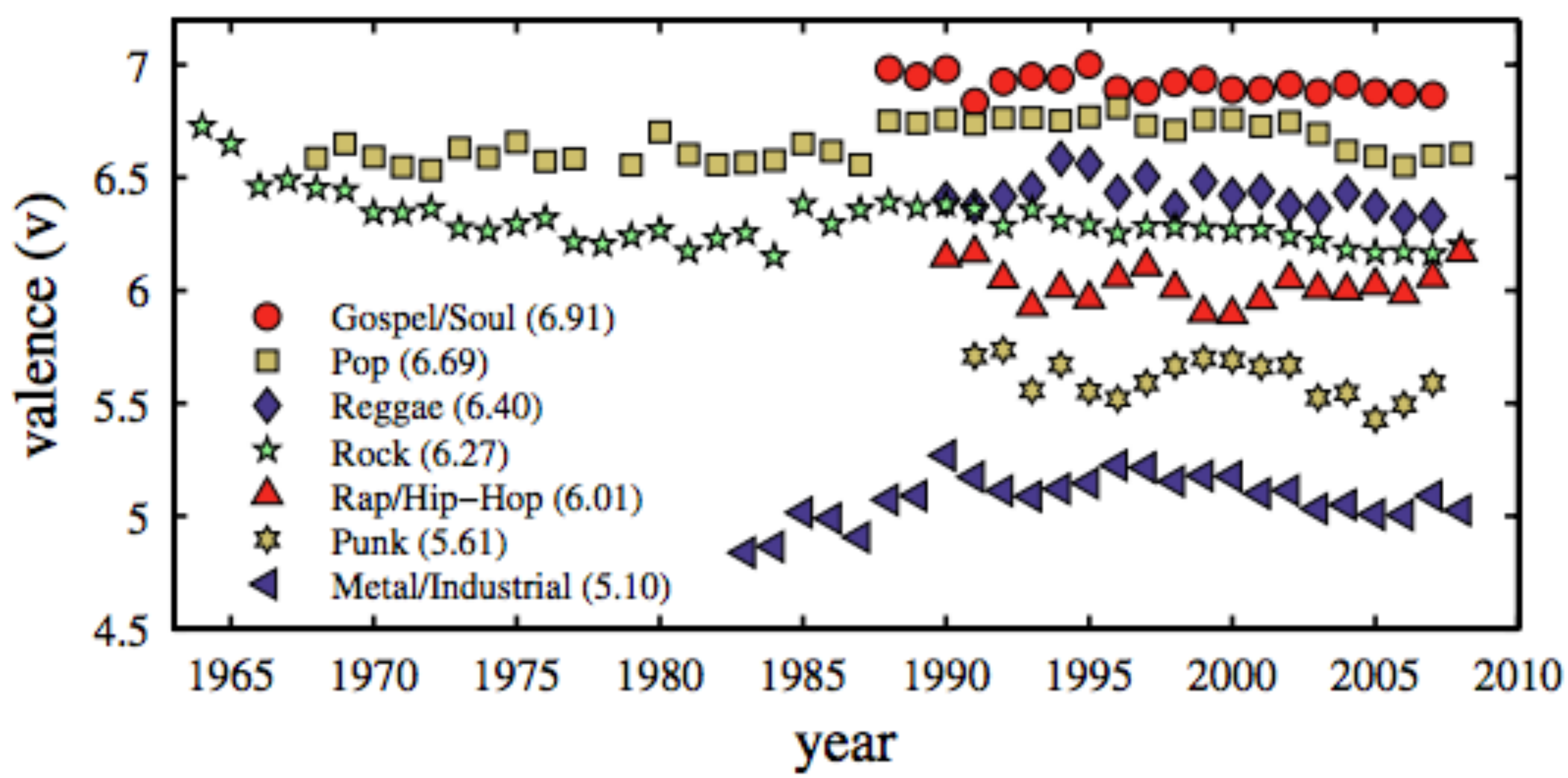
vs.

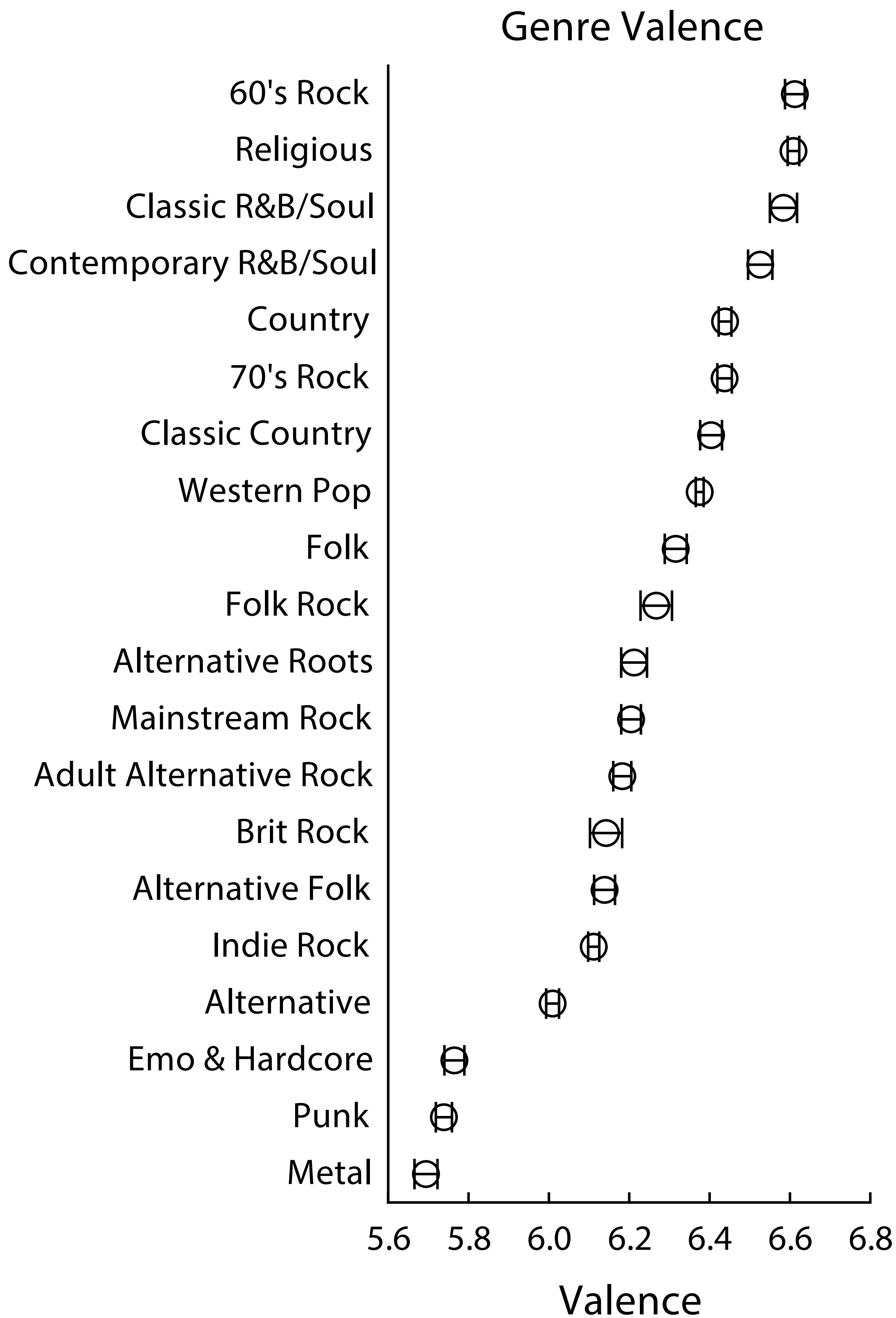
down, we, don't, hate, can't, alone, die, not, hell, fight, dead, waiting, lie,

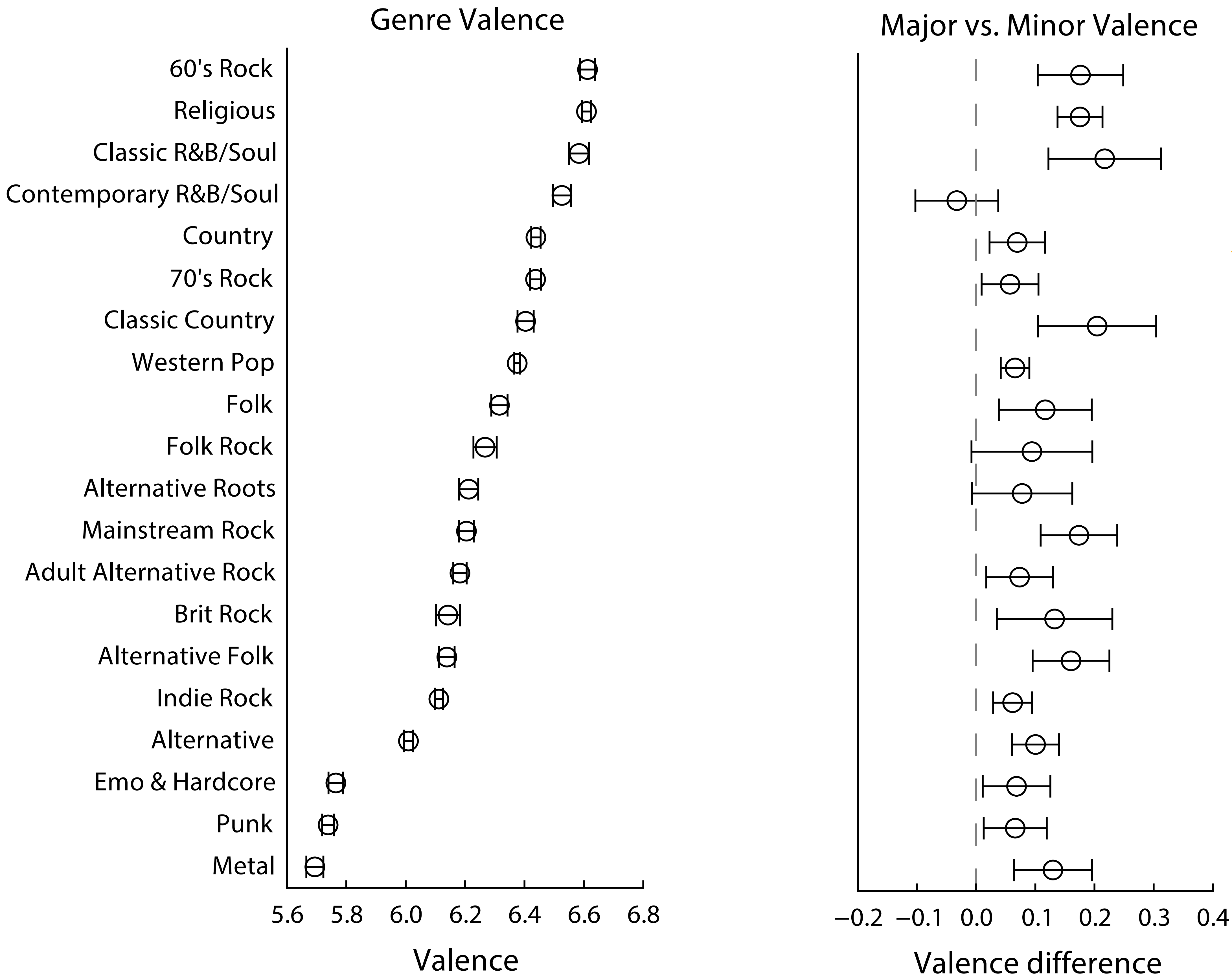
…

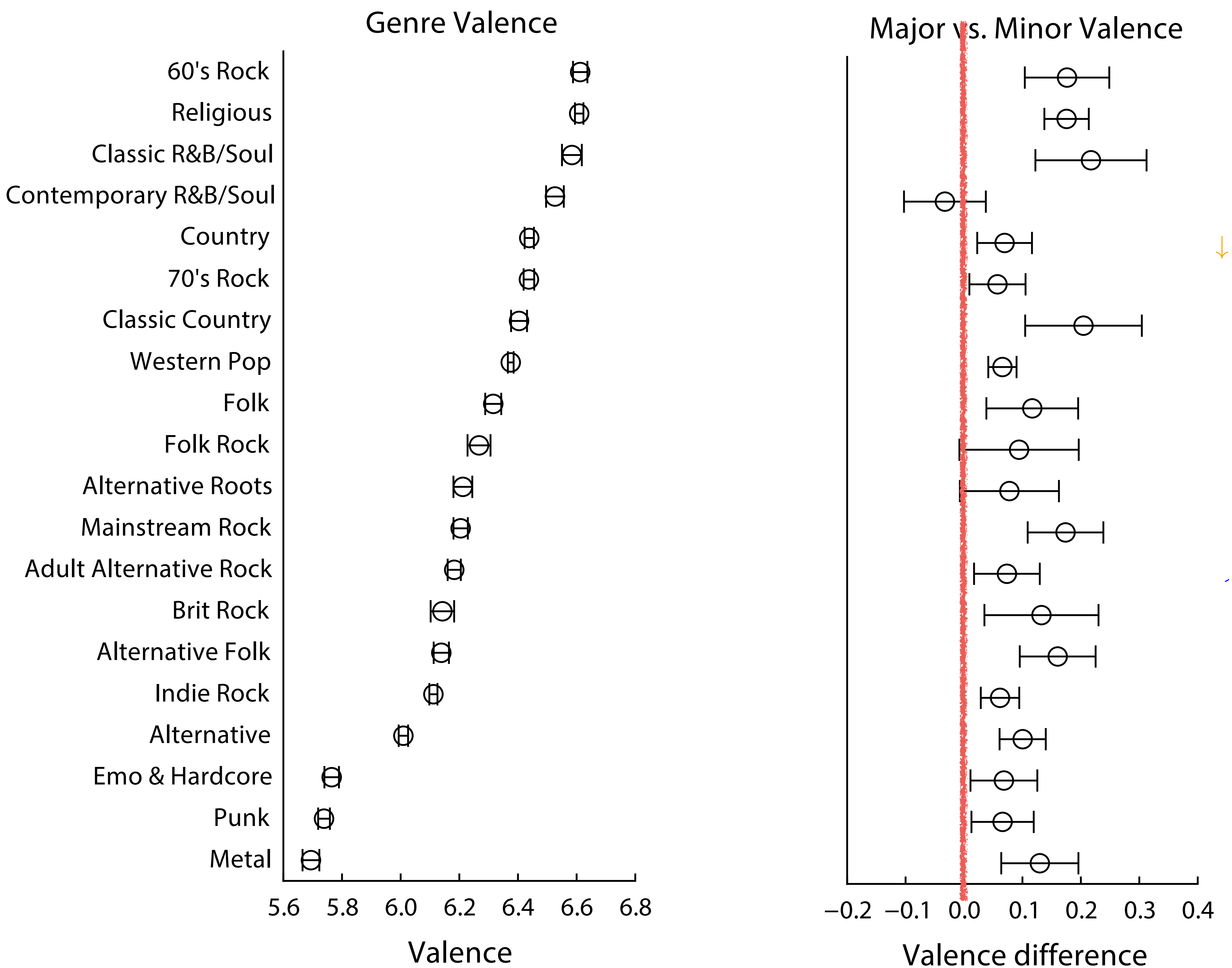
But how about genres?

But how about genres?

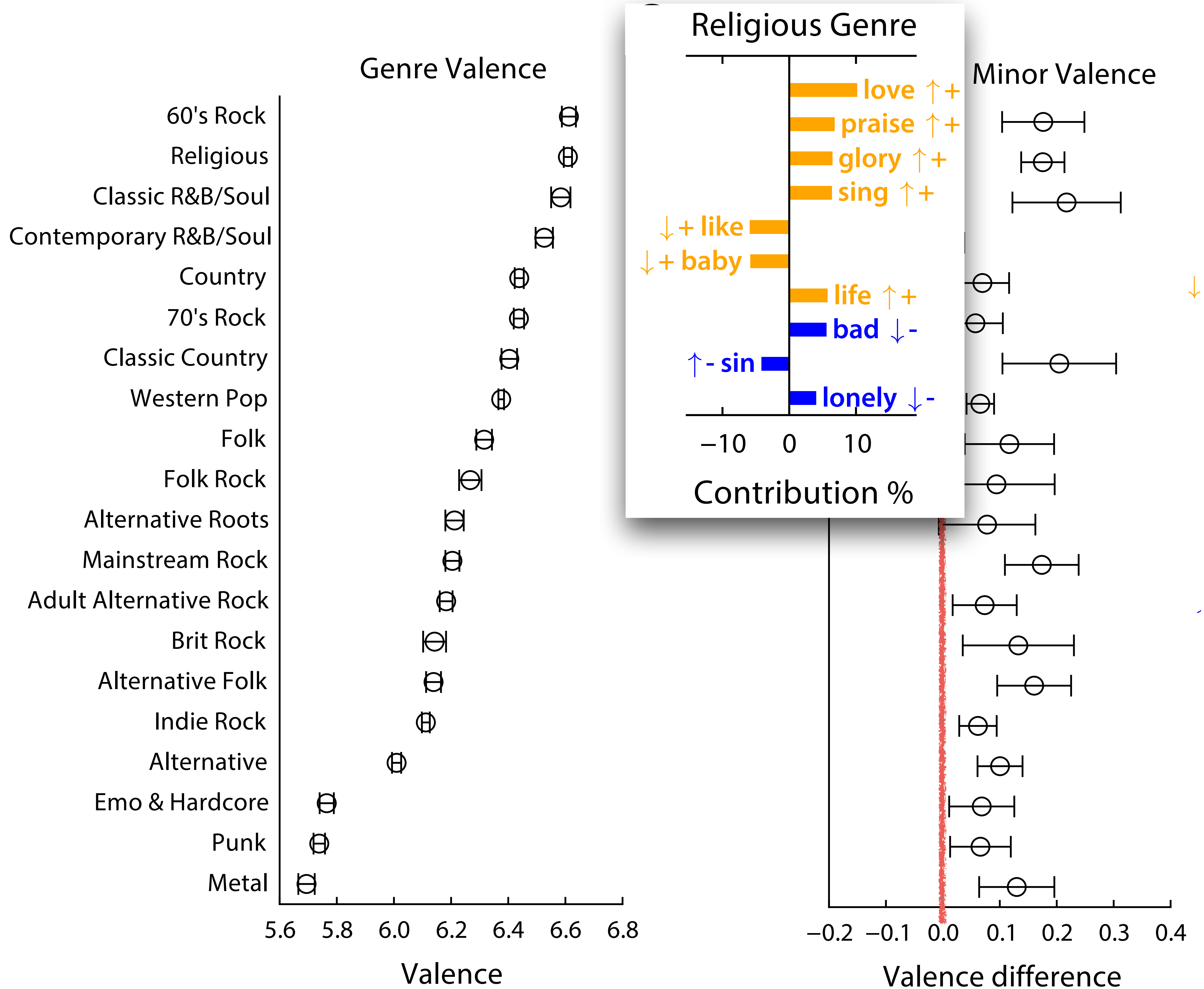



Genre Valence
60's Rock
Religious
Classic R&B/Soul
Contemporary R&B/Soul
Country
70's Rock
Classic Country
Western Pop
Folk
Folk Rock
Alternative Roots
Mainstream Rock
Adult Alternative Rock
Brit Rock
Alternative Folk
Indie Rock
Alternative
Emo & Hardcore
Punk
Metal
5.6
5.8
6.0
6.2
6.4
6.6
6.8
Valence

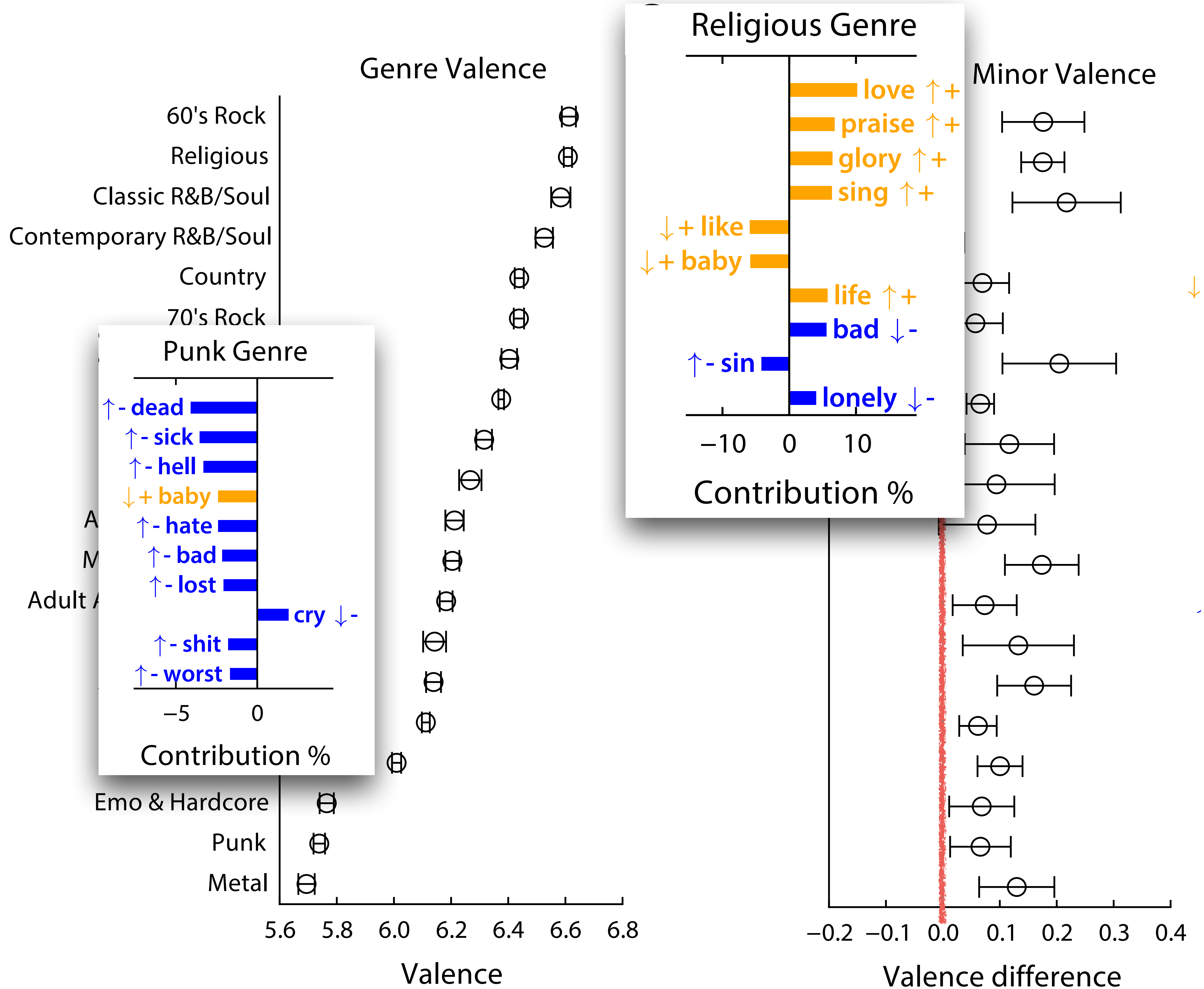

Genre Valence
60's Rock
Religious
Classic R&B/Soul
Contemporary R&B/Soul
Country
70's Rock
Classic Country
Western Pop
Folk
Folk Rock
Alternative Roots
Mainstream Rock
Adult Alternative Rock
Brit Rock
Alternative Folk
Indie Rock
Alternative
Emo & Hardcore
Punk
Metal
5.6 5.8 6.0 6.2 6.4 6.6 6.8
Valence
Major vs. Minor Valence
−0.2 −0.1 0.0 0.1 0.2 0.3 0.4
Valence difference

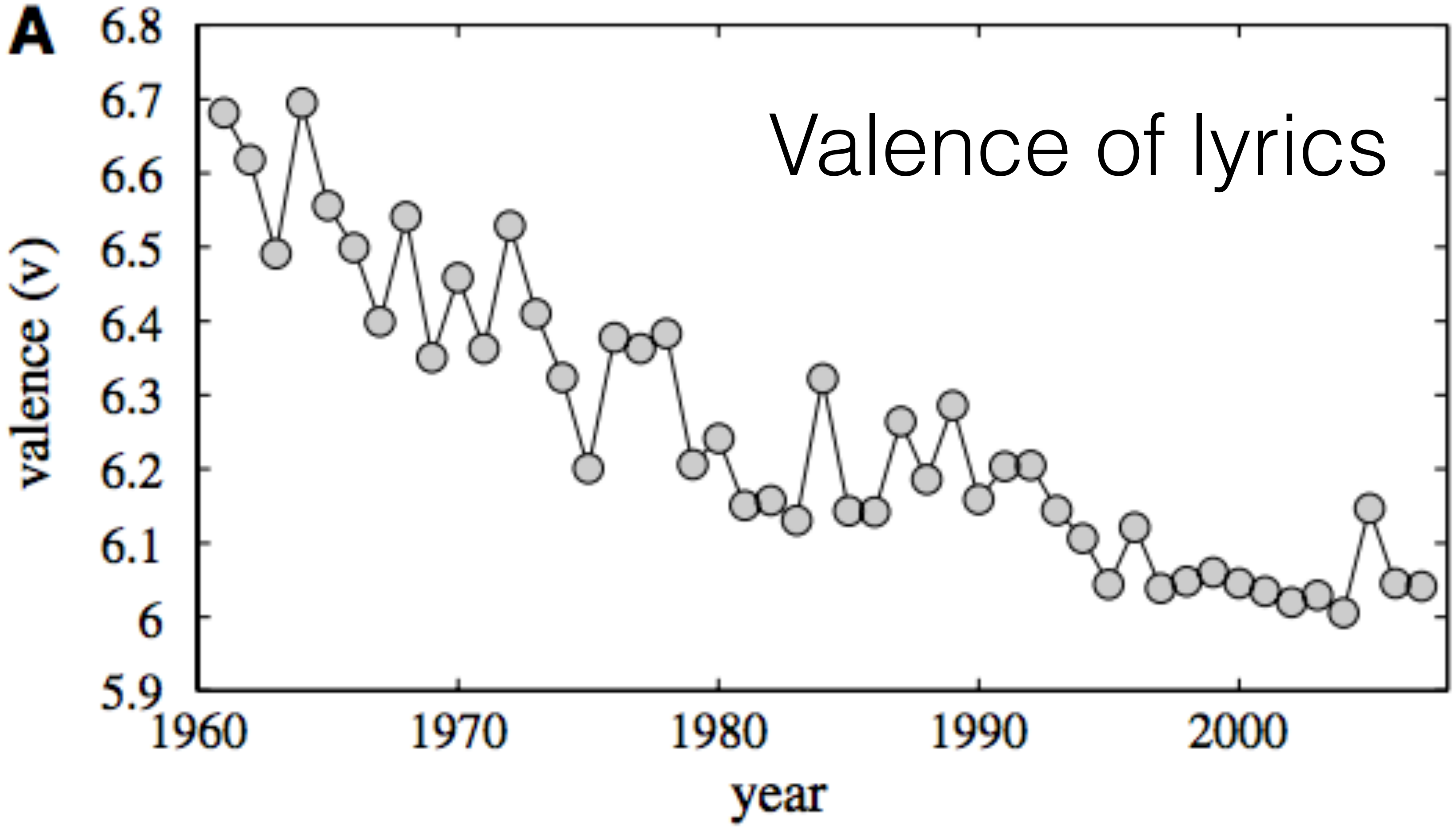


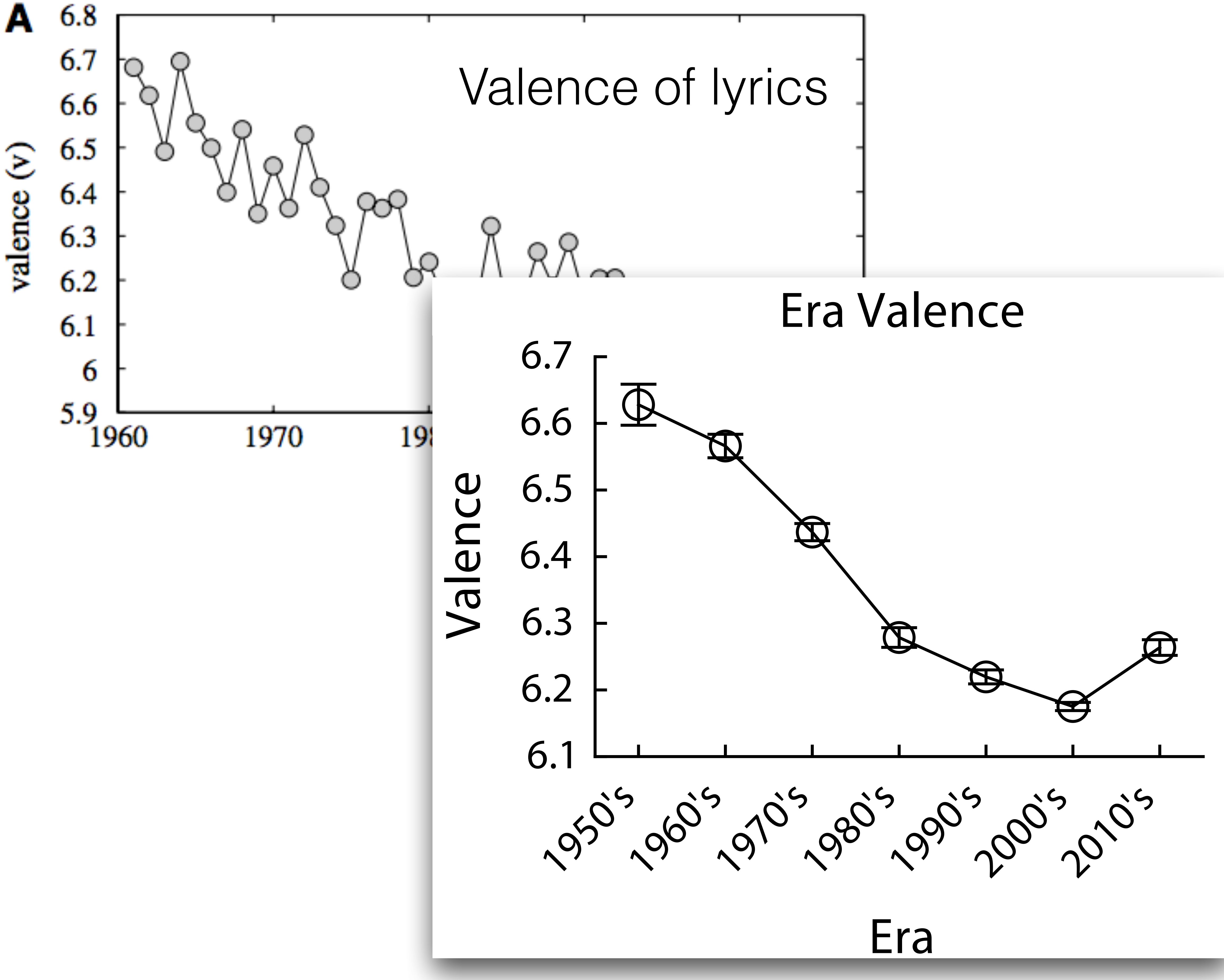
Genre Valence
60's Rock
Religious
Classic R&B/Soul
Contemporary R&B/Soul
Country
70's Rock
Classic Country
Western Pop
Folk
Folk Rock
Alternative Roots
Mainstream Rock
Adult Alternative Rock
Brit Rock
Alternative Folk
Indie Rock
Alternative
Emo & Hardcore
Punk
Metal
5.6
5.8
6.0
6.2
6.4
6.6
6.8
Valence
Major vs. Minor Valence
−0.2
−0.1
0.0
0.1
0.2
0.3
0.4
Valence difference

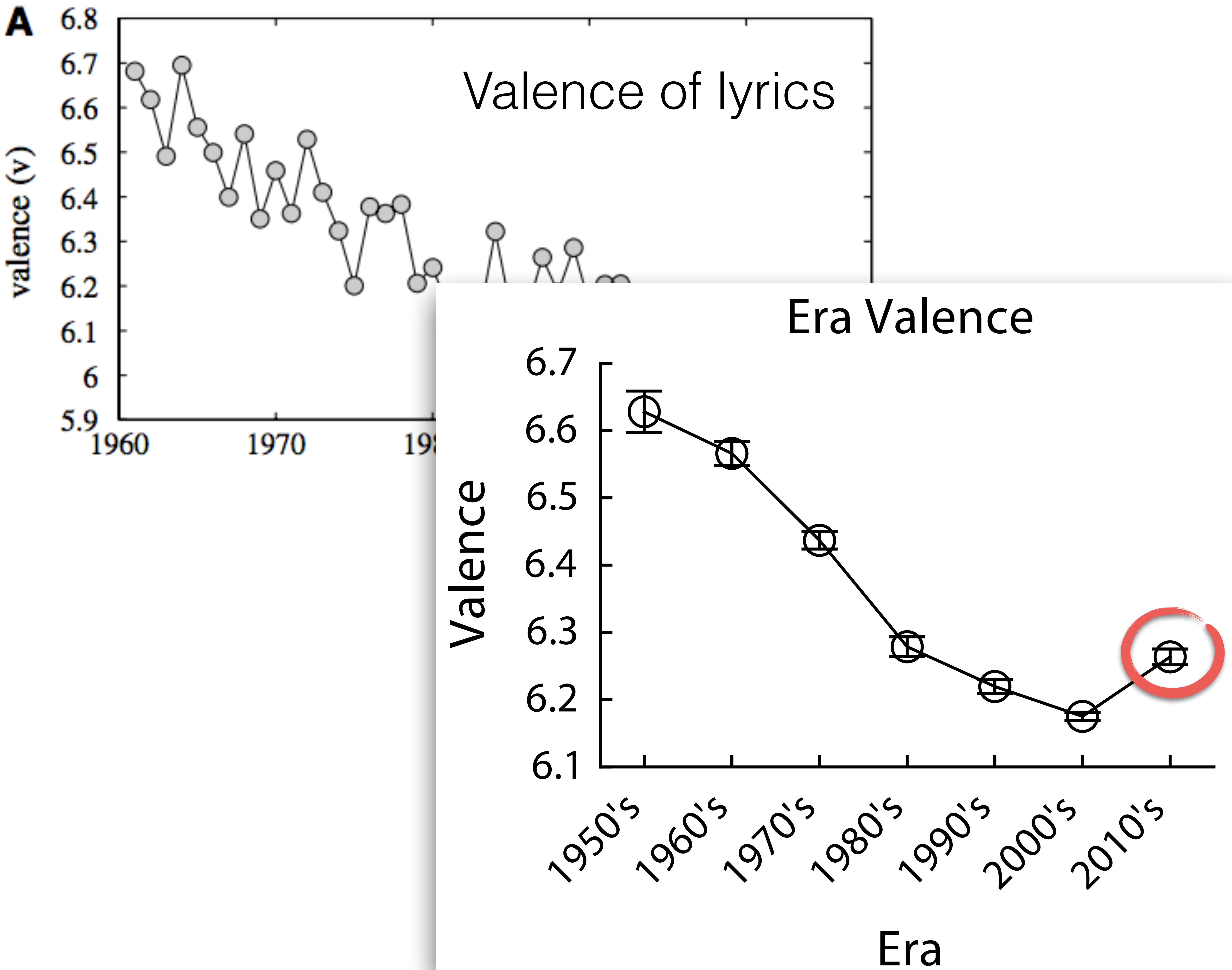


Genre Valence
60's Rock
Religious
Classic R&B/Soul
Contemporary R&B/Soul
Country
70's Rock
Classic Country
Western Pop
Folk
Folk Rock
Alternative Roots
Mainstream Rock
Adult Alternative Rock
Brit Rock
Alternative Folk
Indie Rock
Alternative
Emo & Hardcore
Punk
Metal
5.6 5.8 6.0 6.2 6.4 6.6 6.8
Valence
Religious Genre
love ↑+
praise ↑+
glory ↑+
sing ↑+
↓+ like
↓+ baby
life ↑+
bad ↓-
↑- sin
lonely ↓-
−10 0 10
Contribution %
Minor Valence
−0.2 −0.1 0.0 0.1 0.2 0.3 0.4
Valence difference


Genre Valence
60's Rock
Religious
Classic R&B/Soul
Contemporary R&B/Soul
Country
70's Rock
Emo & Hardcore
Punk
Metal
5.6
5.8
6.0
6.2
6.4
6.6
6.8
Valence
Punk Genre
↑- dead
↑- sick
↑- hell
↓+ baby
↑- hate
↑- bad
↑- lost
cry ↓-
↑- shit
↑- worst
−5
0
Contribution %
Religious Genre
love ↑+
praise ↑+
glory ↑+
sing ↑+
↓+ like
↓+ baby
life ↑+
bad ↓-
↑- sin
lonely ↓-
−10
0
10
Contribution %
Minor Valence
−0.2
−0.1
0.0
0.1
0.2
0.3
0.4
Valence difference

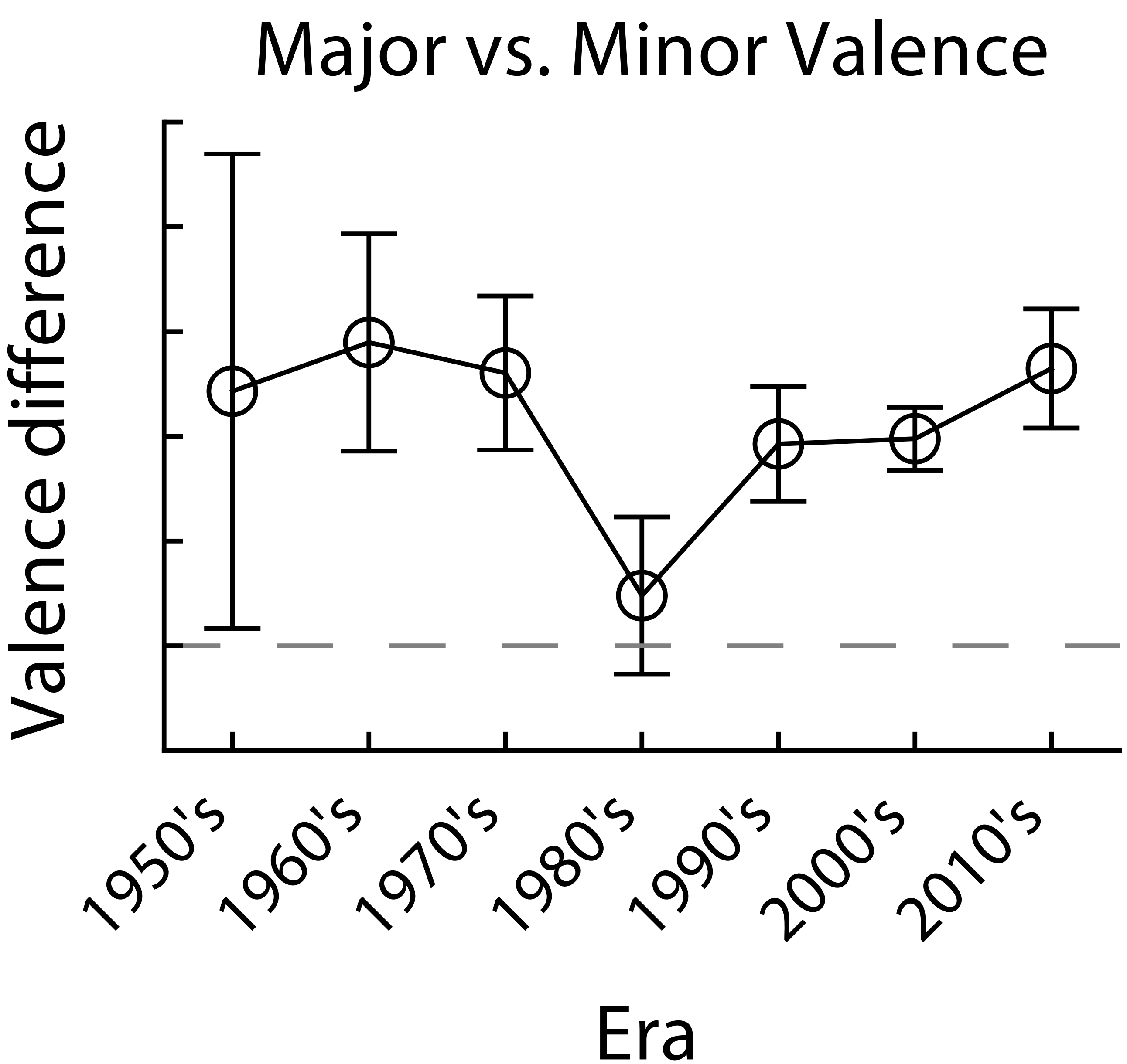

A
Valence of lyrics
valence (v)
6.8
6.7
6.6
6.5
6.4
6.3
6.2
6.1
6
5.9
1960
1970
1980
1990
2000
year


A
Valence of lyrics
valence (v)
6.8
6.7
6.6
6.5
6.4
6.3
6.2
6.1
6
5.9
1960
1970
Era Valence
Valence
6.7
6.6
6.5
6.4
6.3
6.2
6.1
1950's
1960's
1970's
1980's
1990's
2000's
2010's
Era

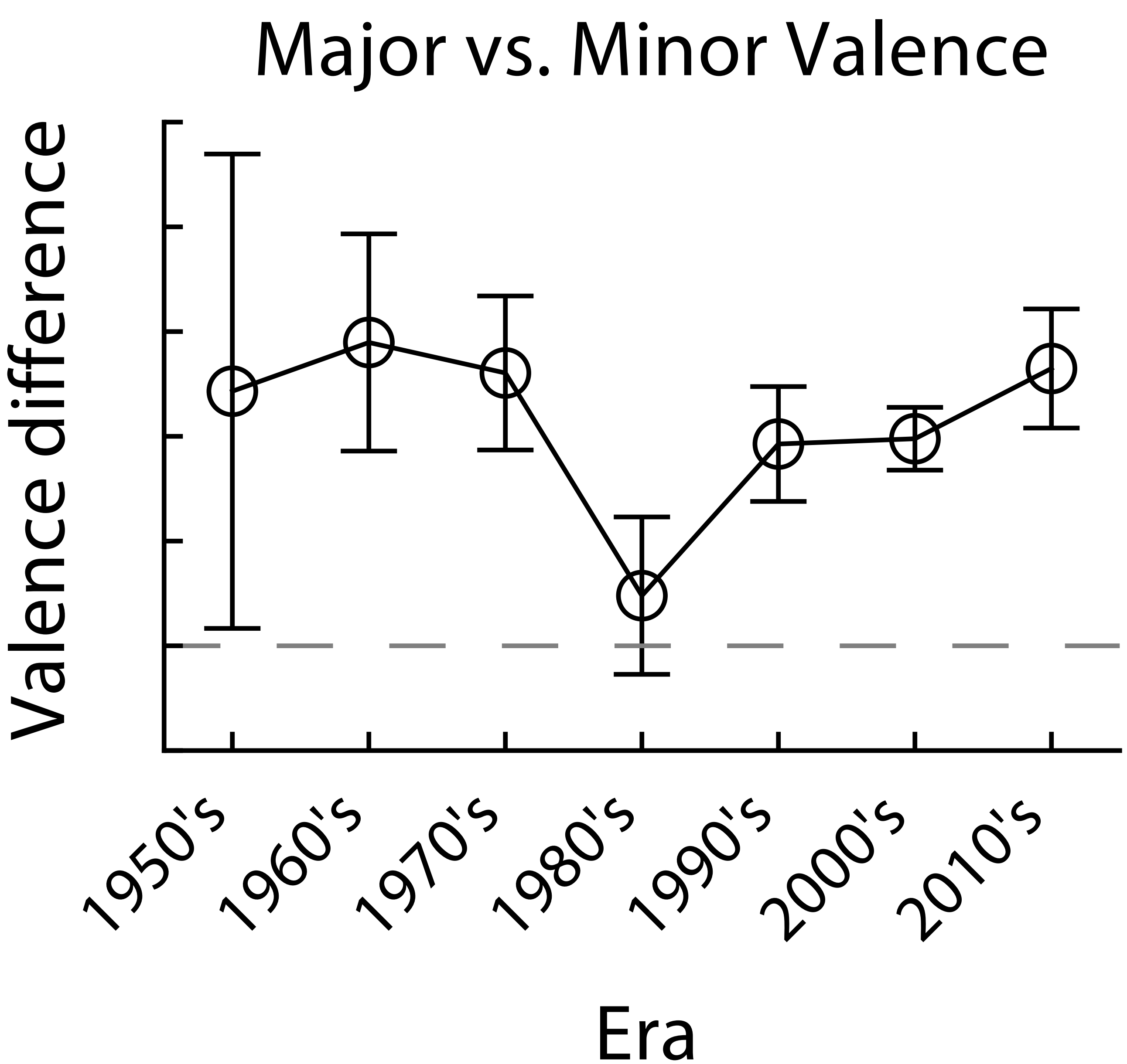


A
Valence of lyrics
valence (v)
6.8
6.7
6.6
6.5
6.4
6.3
6.2
6.1
6
5.9
1960
1970
Era Valence
Valence
6.7
6.6
6.5
6.4
6.3
6.2
6.1
1950's
1960's
1970's
1980's
1990's
2000's
2010's
Era

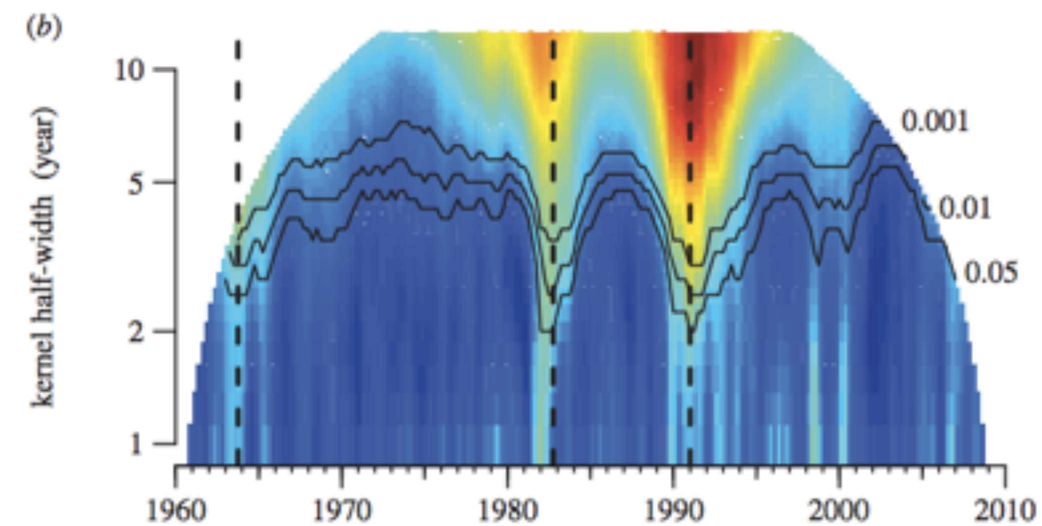
Was 80’s weird?



Was 80's weird?

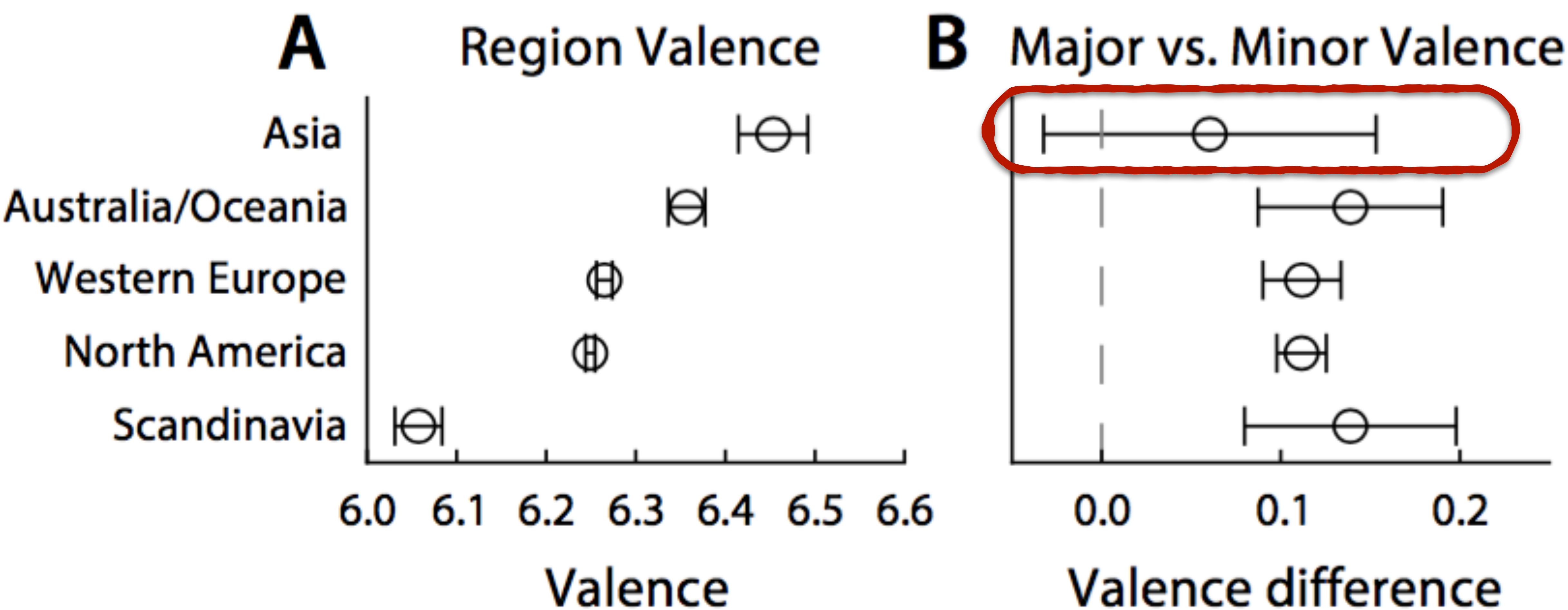


Topic analysis of sound features of Billboard 100



M. Mauch et al., Roy. Sci. Open Science (2016).

Regional difference



Summary

Summary

- Lyrics as a proxy to understand ‘meaning’ of other musical elements?

Summary

- Lyrics as a proxy to understand 'meaning' of other musical elements?

- Curious associations: 7th > Major > Minor; Minor ~ negations; 7th ~ Love, …

Summary

- Lyrics as a proxy to understand ‘meaning’ of other musical elements?

- Curious associations: 7th > Major > Minor; Minor ~ negations; 7th ~ Love, …

- Major - Minor difference is fairly consistent, but not as robust as we may assume.

Summary

- Lyrics as a proxy to understand ‘meaning’ of other musical elements?

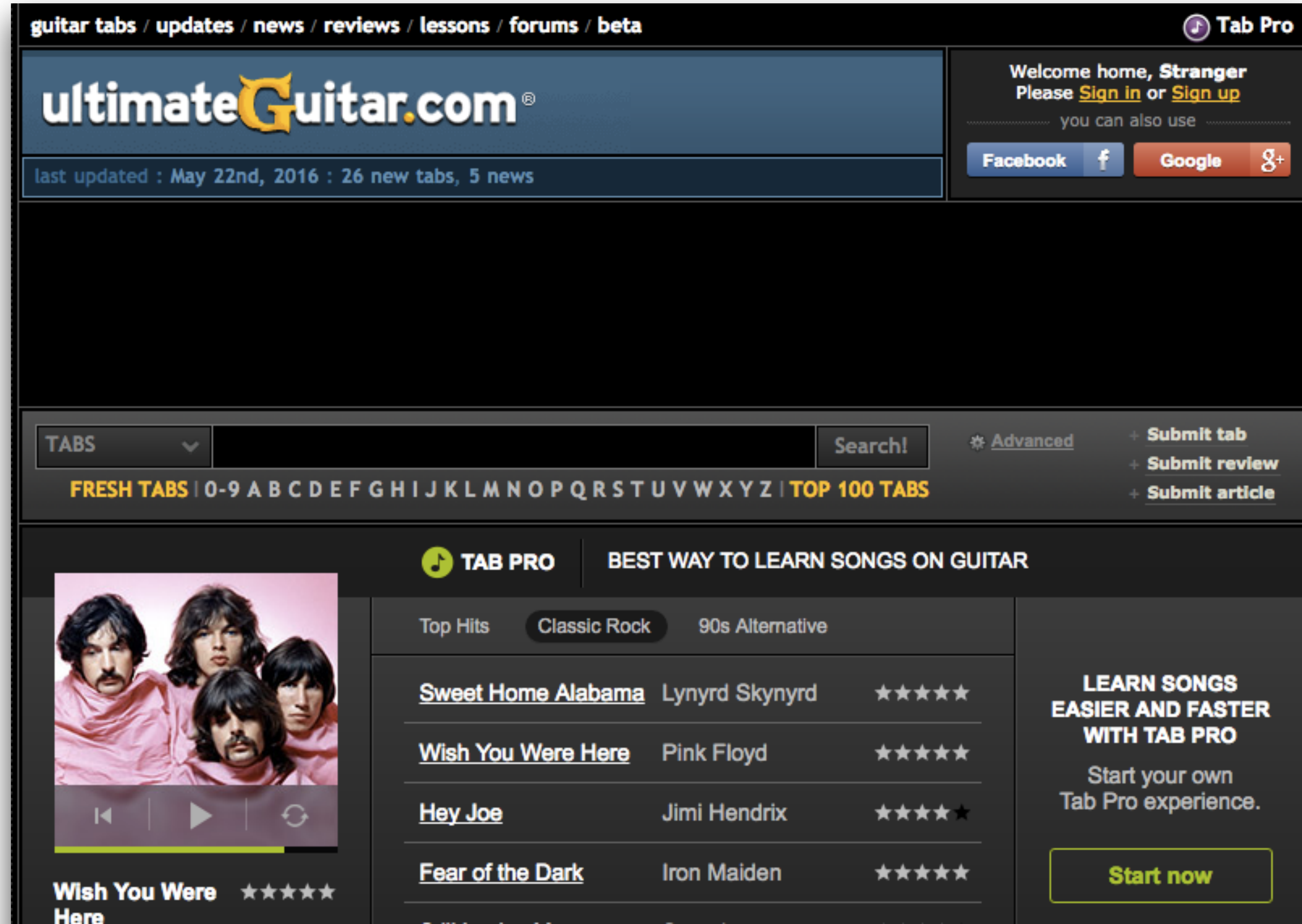
- Curious associations: 7th > Major > Minor; Minor ~ negations; 7th ~ Love, …

- Major - Minor difference is fairly consistent, but not as robust as we may assume.

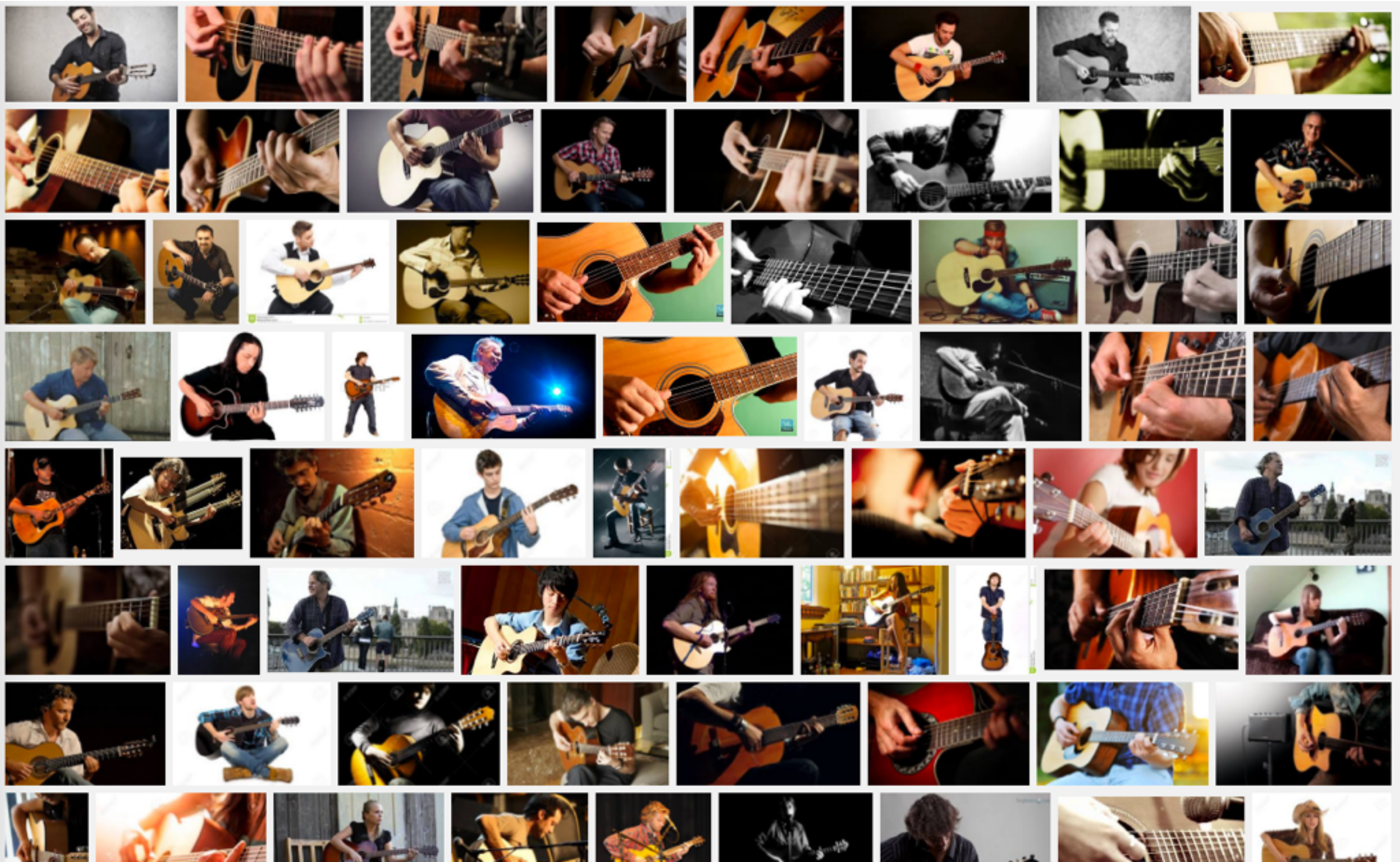
- Still lots of caveats!

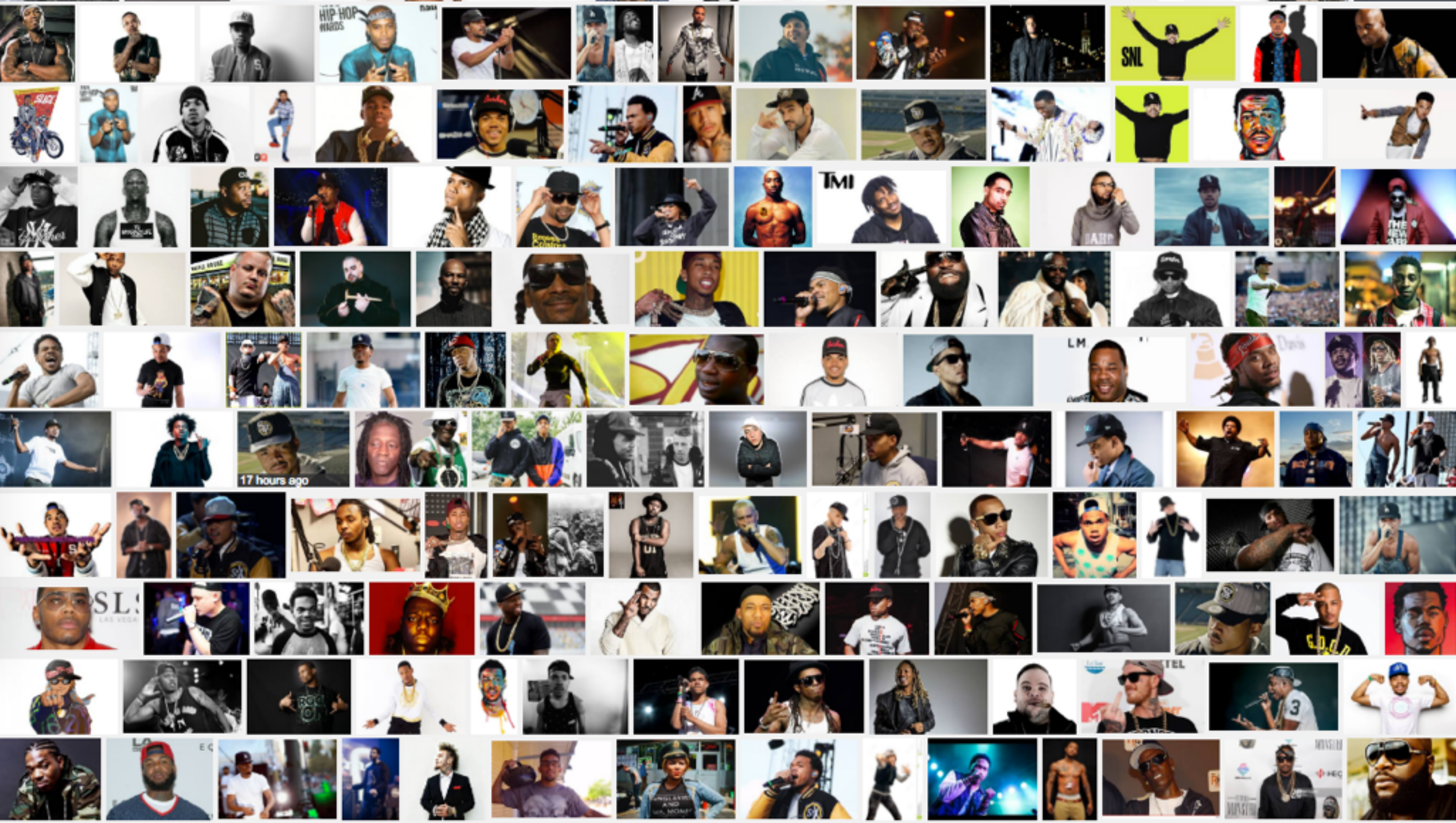
https://www.youtube.com/watch?v=oOlDewpCfZQ https://www.youtube.com/watch?v=oOlDewpCfZQ

#2 Lyrics and Society


guitar tabs / updates / news / reviews / lessons / forums / beta
Tab Pro
ultimateGuitar.com®
last updated : May 22nd, 2016 : 26 new tabs, 5 news
Welcome home, Stranger
Please Sign in or Sign up
you can also use
Facebook
Google
TABS
Search!
Advanced
Submit tab
Submit review
Submit article
FRESH TABS | 0-9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z | TOP 100 TABS
TAB PRO
BEST WAY TO LEARN SONGS ON GUITAR
Top Hits
Classic Rock
90s Alternative
Sweet Home Alabama
Lynyrd Skynyrd
Wish You Were Here
Pink Floyd
Hey Joe
Jimi Hendrix
Fear of the Dark
Iron Maiden
LEARN SONGS EASIER AND FASTER WITH TAB PRO
Start your own Tab Pro experience.
Start now
Wish You Were Here

“Acoustic guitar player”





“Rapper”

Jaehyuk Park

Can we study the relationship between **hip hop lyrics** and **social movements**?

Wu-Tang Clan Use Ferguson/Eric Garner Protest Footage For New Video

BY **BEN YAKAS** IN **ARTS & ENTERTAINMENT** ON DEC 6, 2014 11:40 AM

"F*** tha police" by N.W.A. (1988)

F*** the police coming straight from the underground
A young n**** got it bad cause I'm brown
And not the other color so police think
They have the authority to kill a minority
…
Searching my car, looking for the product
Thinking every n**** is selling narcotics
…
Punk police are afraid of me, huh
A young n**** on the warpath
And when I'm finished, it's gonna be a bloodbath
Of cops, dying in L.A

https://www.youtube.com/watch?v=-fVsA_Gm0No

https://www.youtube.com/watch?v=-fVsA_Gm0No

Los Angeles riots: Gangsta rap foretold them and grew after them

Toddy Tee and N.W.A were a ready-made soundtrack in April 1992. Ice Cube's and Dr. Dre's albums that year explained the feelings in South L.A. neighborhoods. Snoop Dogg, Tupac Shakur and more followed.

May 02, 2012 | By Ernest Hardy and August Brown, Los Angeles Times

Email Share G+1 7 Tweet Recommend 0

In 1985, Los Angeles rapper Toddy Tee released wh
salvo against police brutality in black neighborhoo
battering ram that then-LAPD Chief Daryl F. Gates
was a hit on local radio station KDAY-AM.

The track went on to become a protest anthem in t
device was often deployed against homes that were
rockhouse / Well, I know to you we all look the sa
five and ain't a damn thing changed ..." rapped To

The L.A. riots of 1992 arrived with its soundtrack i
gang life, a decimated school system, the toll of cr
being documented by West Coast rappers long be
Department officers was documented on tape. Inr
out of the Bronx in the last '70s — with hard-core,
neighborhoods around them.

"Even before the riots ... voices in L.A. hip-hop w
Singleton, whose 1991 film "Boyz n the Hood" wa
many Americans. "So many people who didn't gr
happened. You can live in a different part of L.A.
to 'F— tha Police,' you hear where they're coming

The riots gave marginalized music from the hood
music born of the very conditions that precipitat
labels began signing and promoting West Coast
for worse, the Southland style that became know

HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | U.S. Edition

The New York Times

News

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

Rappers Say the Riots Were No Surprise To Their Listeners

By SHEILA RULE
Published: May 26, 1992

Correction Appended

Long before the verdict in the Rodney G. King beating and the fires that engulfed South-Central Los Angeles, some black rap artists had been illuminating America's racial tensions. "Police think they have the authority to kill a minority," railed the rapper Ice Cube on "Straight Outta Compton," the 1988 debut album of N.W.A. (Niggas With Attitude).

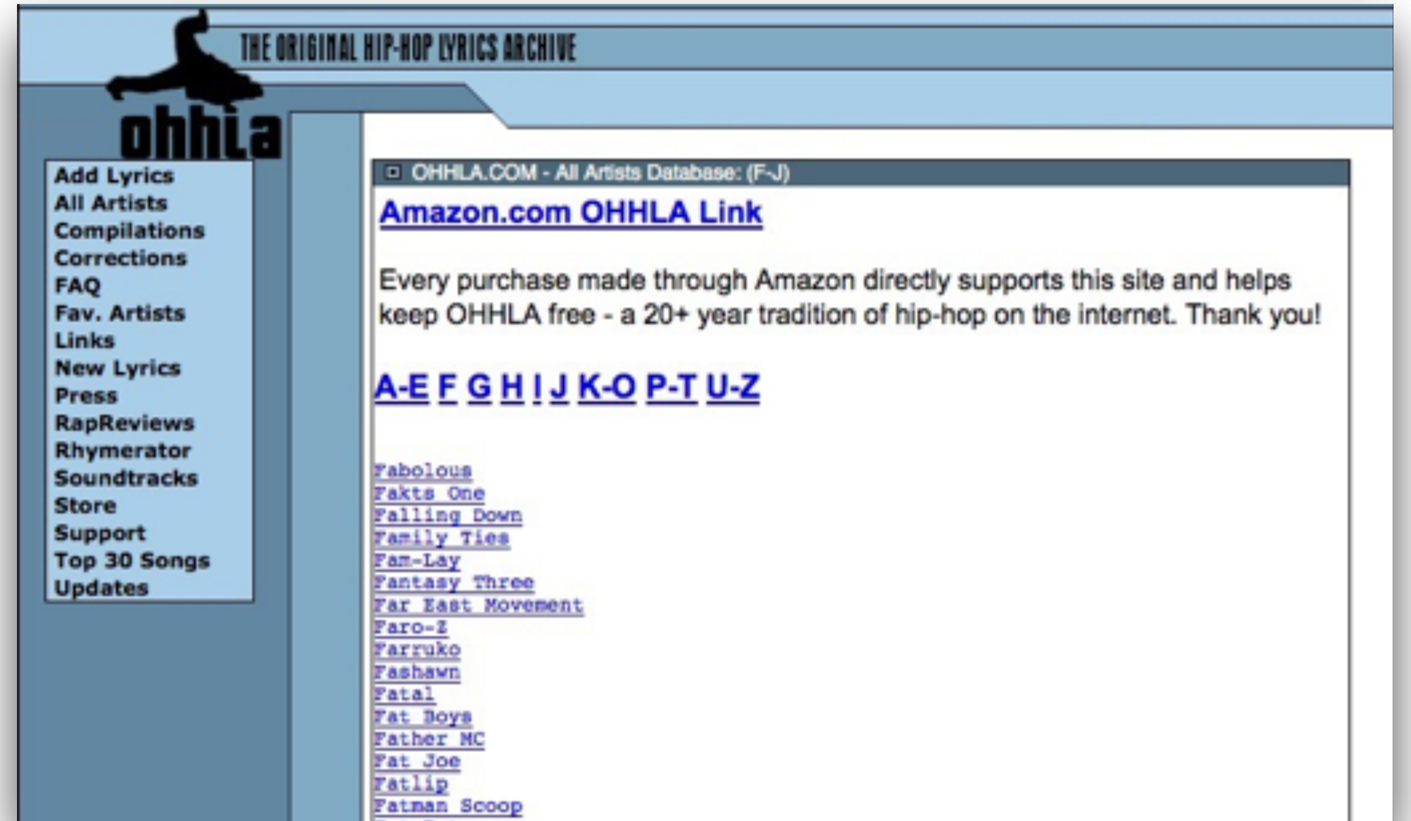
FACEBOOK
TWITTER
GOOGLE+
EMAIL
SHARE
PRINT
REPRINTS

These rappers distilled blacks' anger and prophesied its eruption to anyone who would listen. Millions of young blacks did.

Now, as South-Central residents rebuild, rappers differ on where their message will go

"Music … heralds, for it is prophetic. It has always been in its essence a herald of times to come."

—Jacques Attali,
Noise: the Political Economy of Music

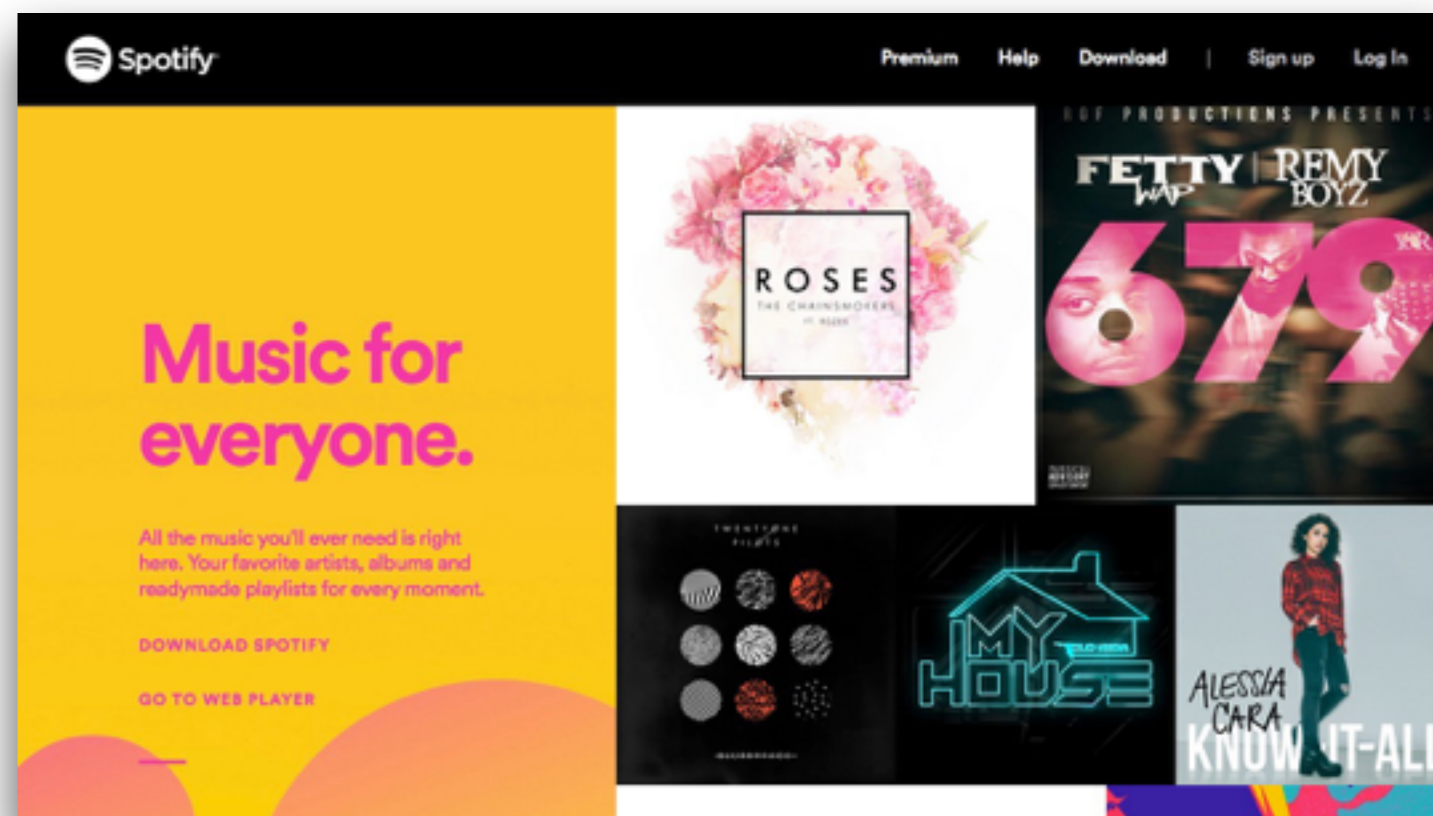


OHHLA.com

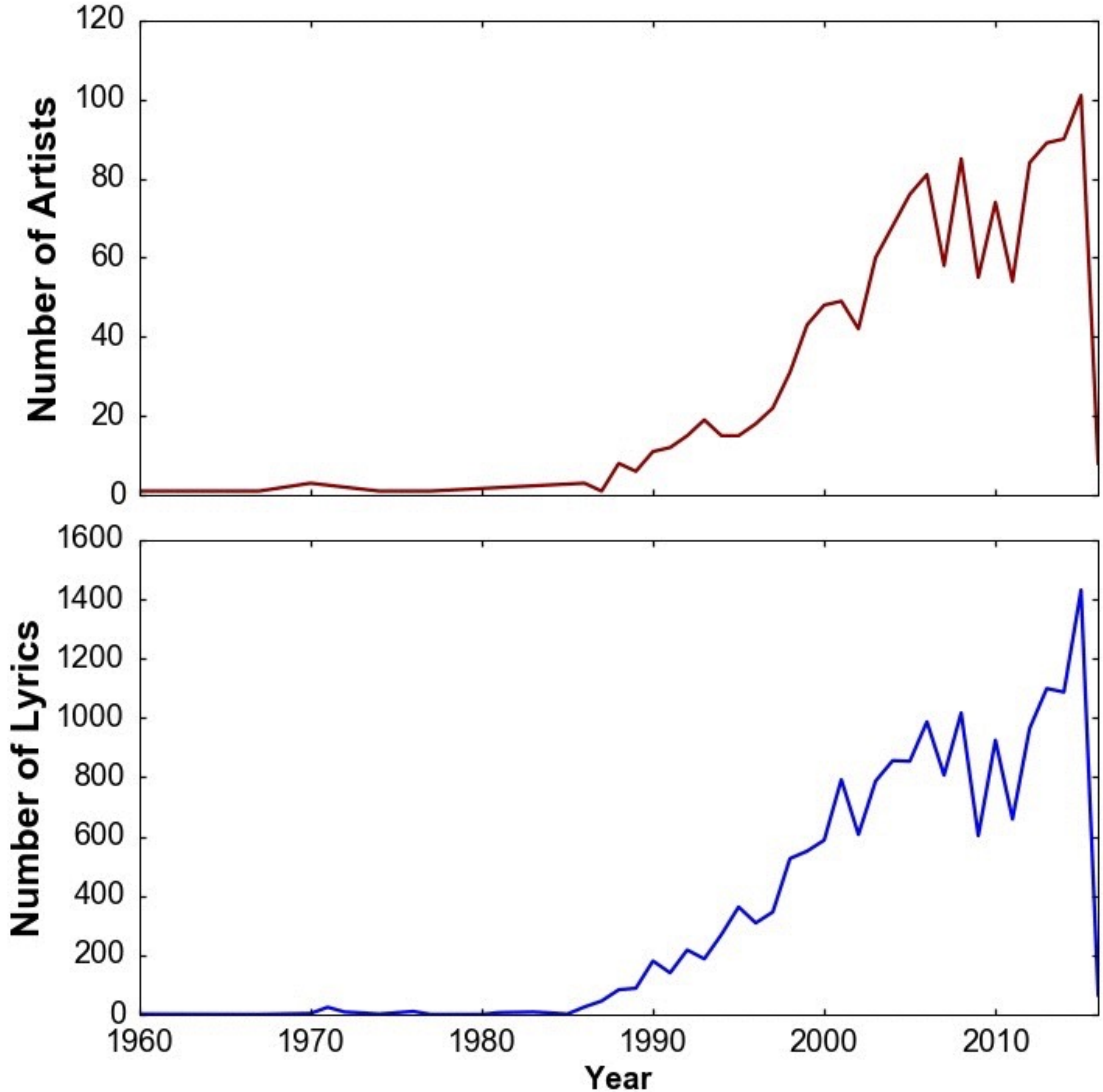
(lyrics)

spotify

(metadata)



Total 18,126 lyrics
from 3,340 albums
by 1,350 artists
during 1960 ~ 2016



How to analyze the
the corpus?

Counting + word2vec

A brief intro about word2vec

Word embedding

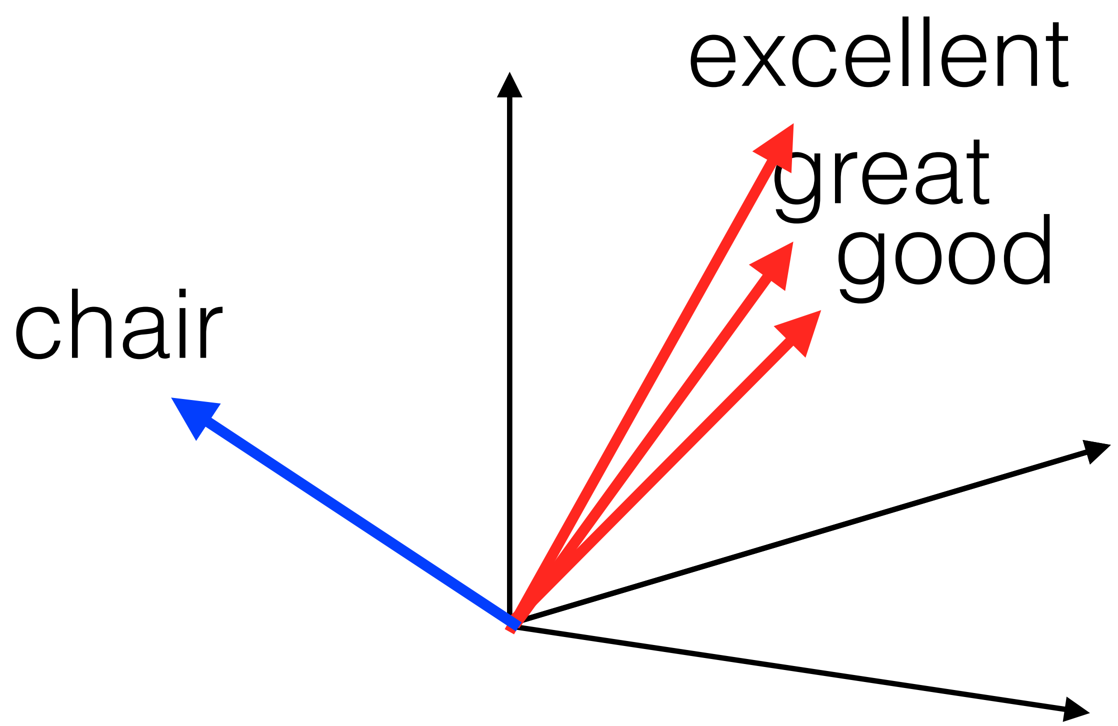
"One-hot" representation

```
the   = (1, 0, 0, 0, …, 0)
quick = (0, 1, 0, 0, …, 0)
…
```

```
Similarity(v(w1),  v(w2)) = 0
```

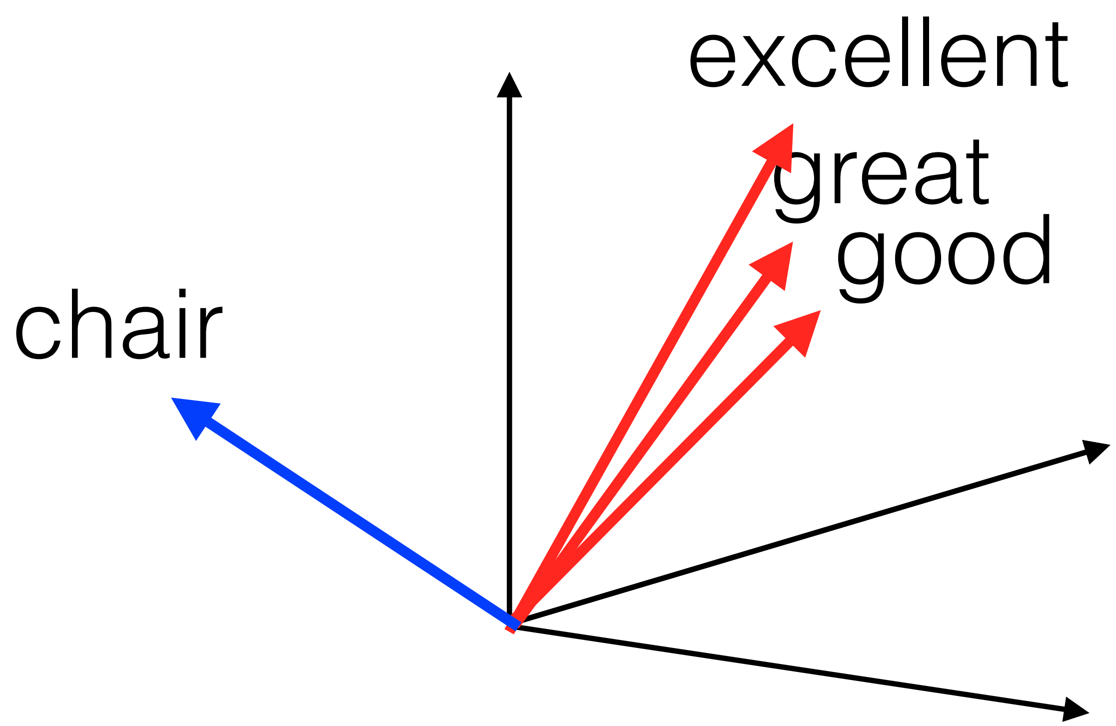
Can we find **dense, continuous, meaningful** representations of words?

Embedding ~ Extracting **features**



How can we put similar words into similar place?

Embedding ~ Extracting **features**



How can we put similar words into similar place?

Hand-crafting

Factorization of term-document matrix

Information theory (PMI)

.....

word2vec: Idea

Language model: how well can we **predict** the **next word** based on the **context**?

$$P(w_1, \ldots, w_m) = \prod_{i=1}^{m} P(w_i \mid w_1, \ldots, w_{i-1})$$

word2vec: Idea

Language model: how well can we **predict** the **next word** based on the **context**?

$$P(w_1, \ldots, w_m) = \prod_{i=1}^{m} P(w_i \mid w_1, \ldots, w_{i-1})$$

Why not "deep-learning" it?

word2vec: Idea

Language model: how well can we **predict** the **next word** based on the **context**?

$$P(w_1, \ldots, w_m) = \prod_{i=1}^{m} P(w_i \mid w_1, \ldots, w_{i-1})$$

Why not "deep-learning" it?

word2vec: Idea

Language model: how well can we **predict** the **next word** based on the **context**?

$$P(w_1, \ldots, w_m) = \prod_{i=1}^{m} P(w_i \mid w_1, \ldots, w_{i-1})$$

Why not "deep-learning" it?

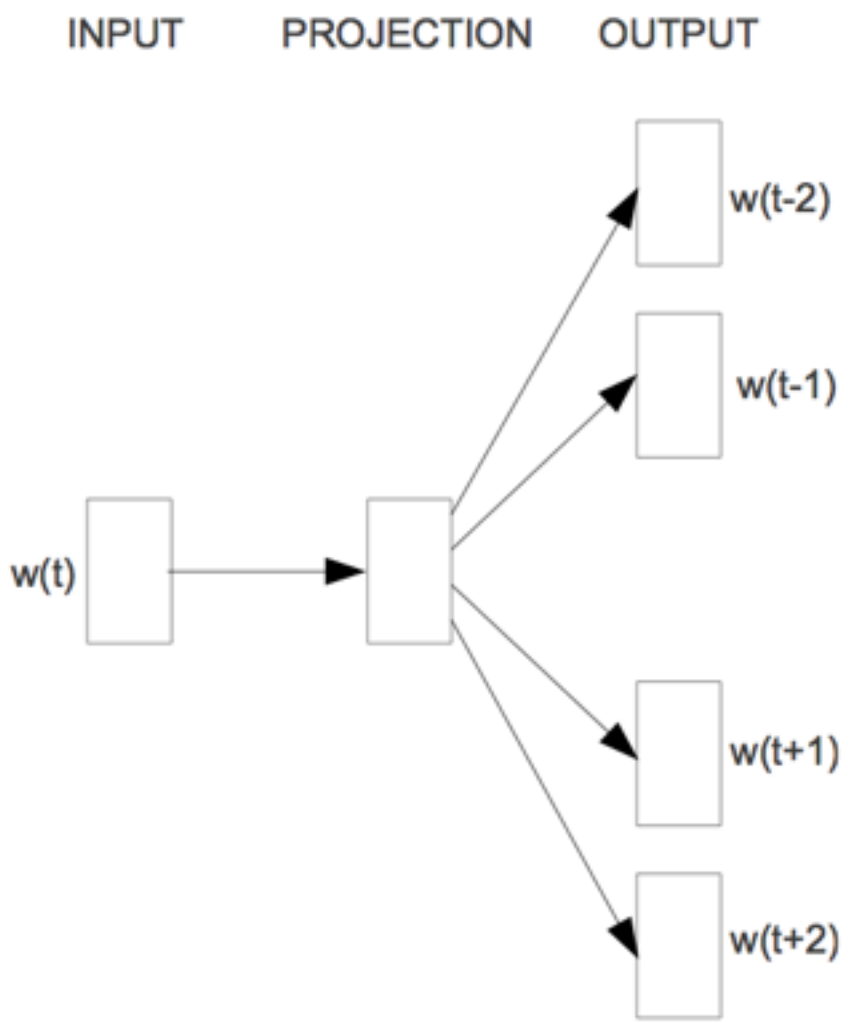
(i.e. why not trying to find the vector representations that can predict target words based on the context best?)

Each word has two vector representations ('in' and 'out') and you learn both.

Credit: Tomas Mikolov, Chris Moody

Each word has two vector representations ('in' and 'out') and you learn both.

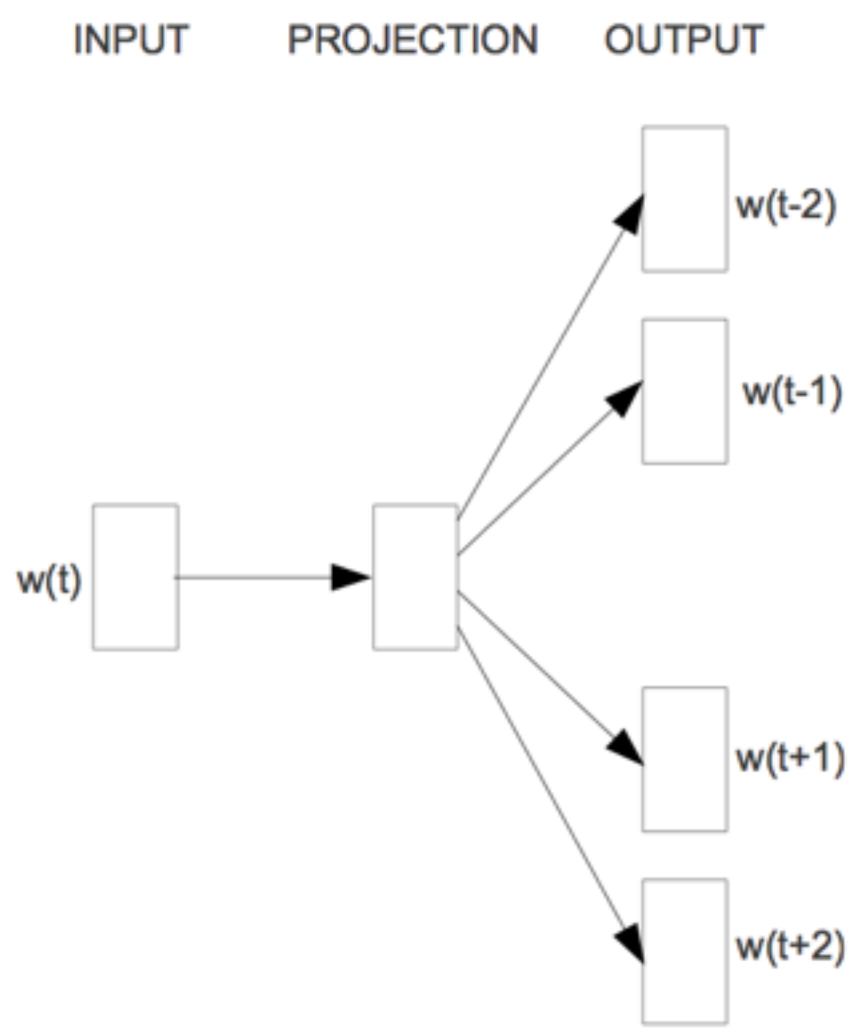
"Skip-gram model"



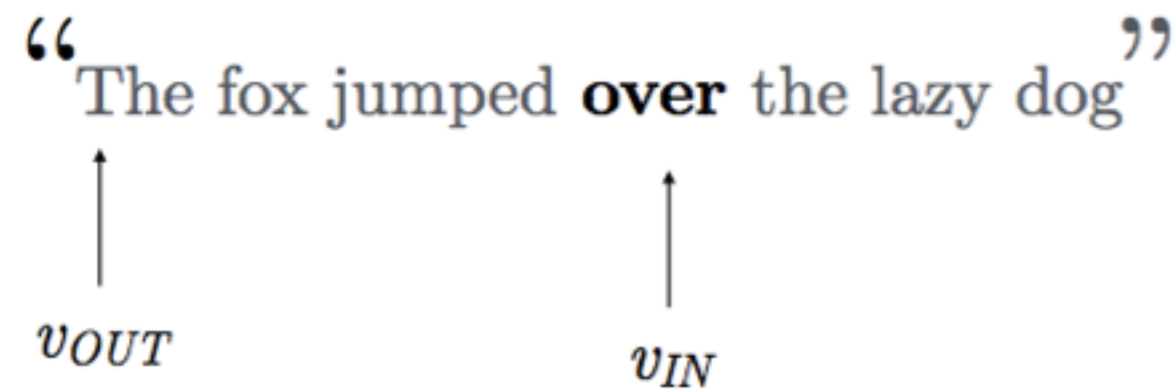
Credit: Tomas Mikolov, Chris Moody

Each word has two vector representations ('in' and 'out') and you learn both.

"Skip-gram model"



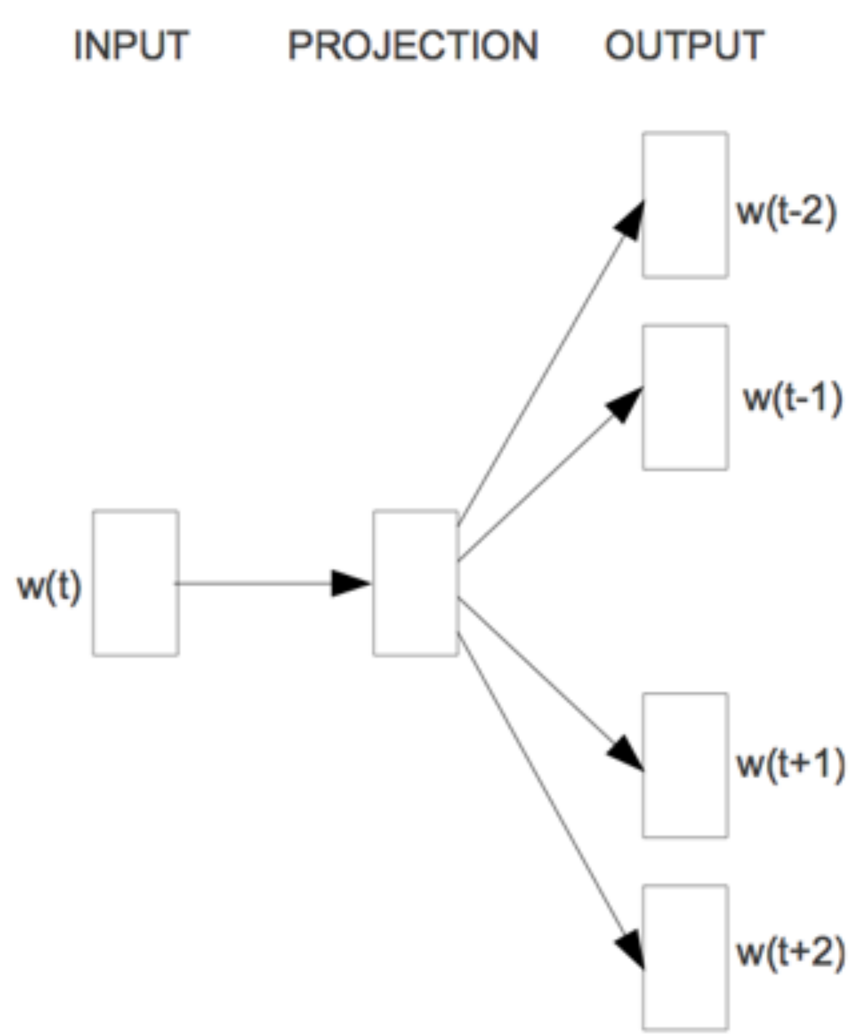
$P(v_{OUT}|v_{IN})$



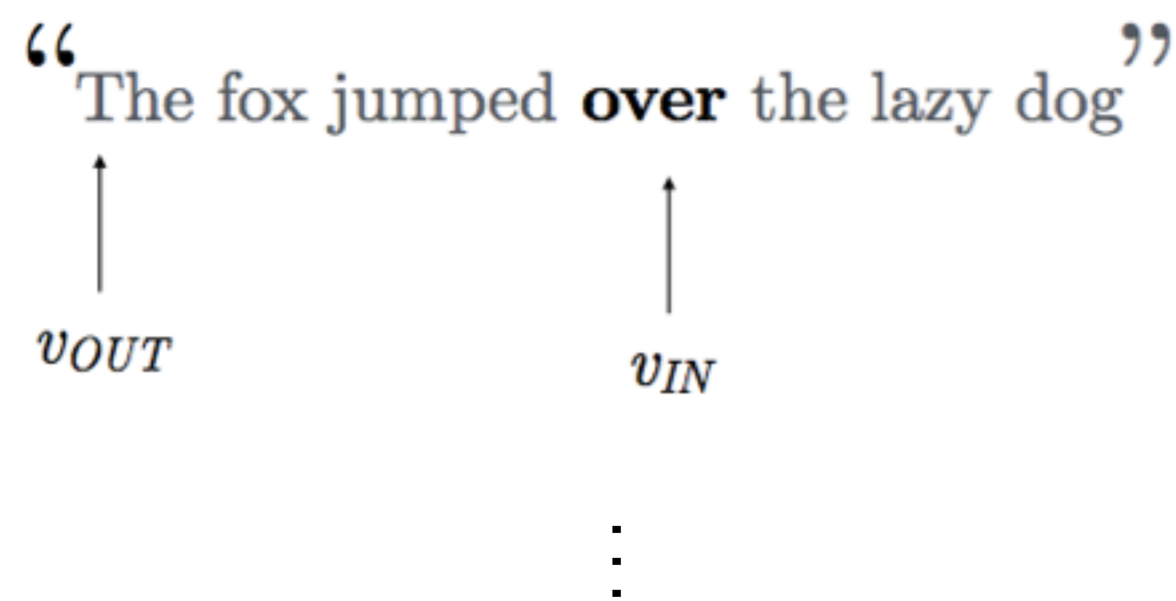
Credit: Tomas Mikolov, Chris Moody

Each word has two vector representations ('in' and 'out') and you learn both.

"Skip-gram model"



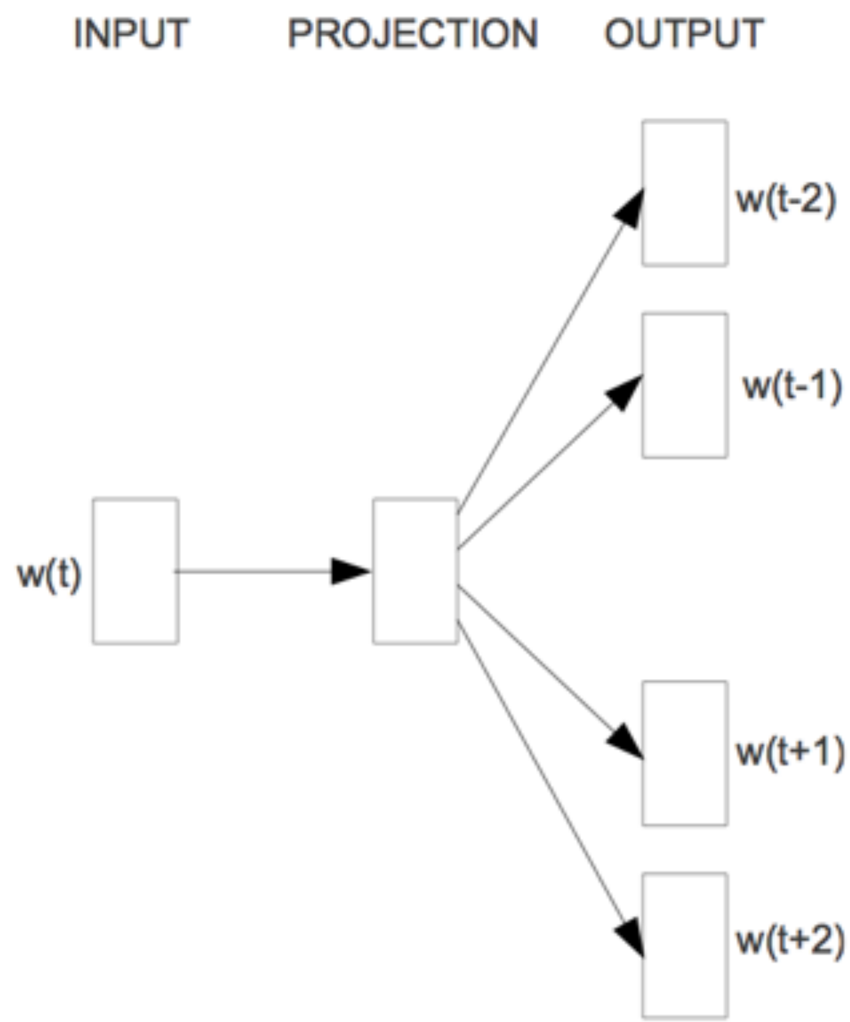
$P(v_{OUT}|v_{IN})$



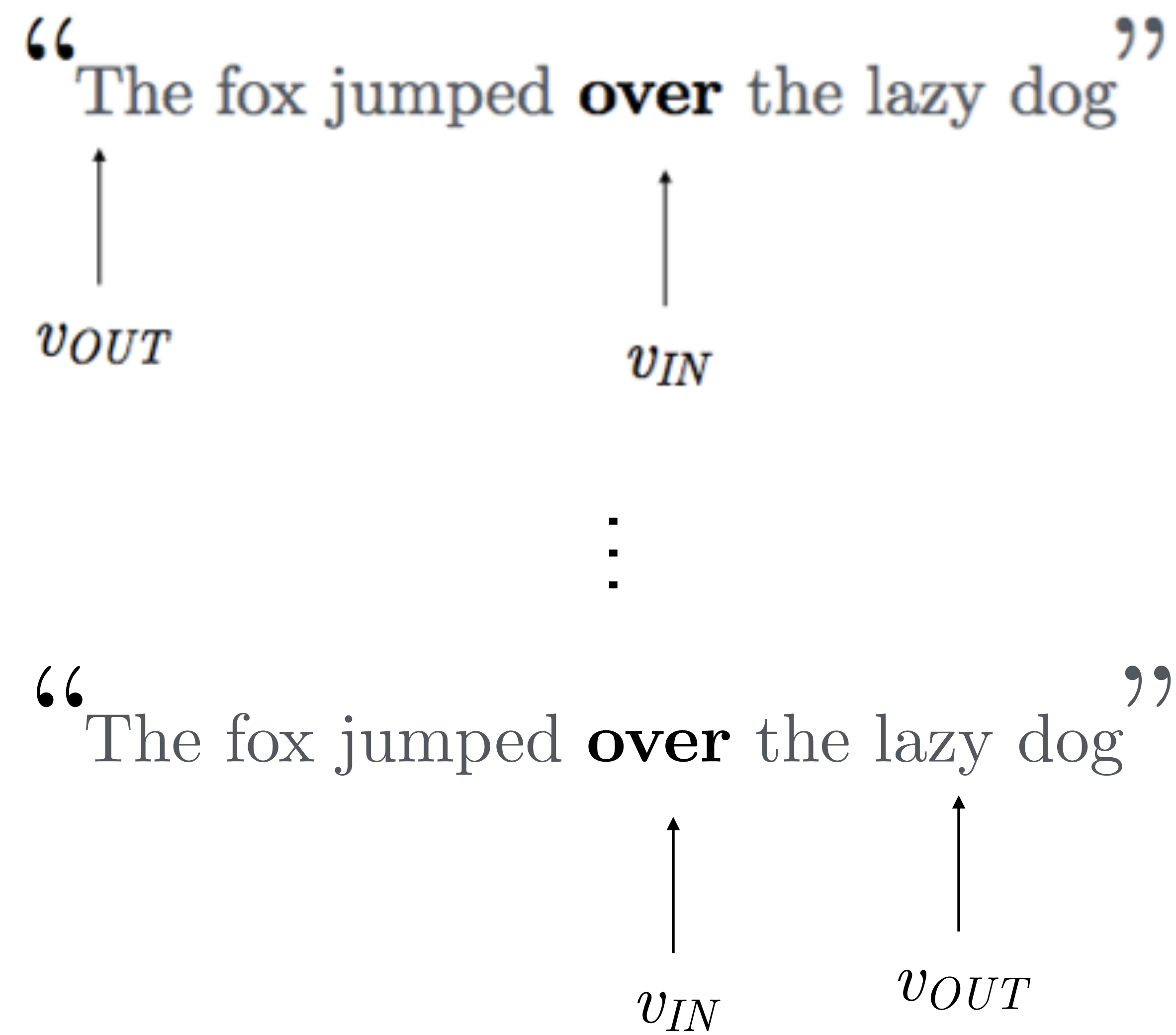
Credit: Tomas Mikolov, Chris Moody

Each word has two vector representations ('in' and 'out') and you learn both.

"Skip-gram model"

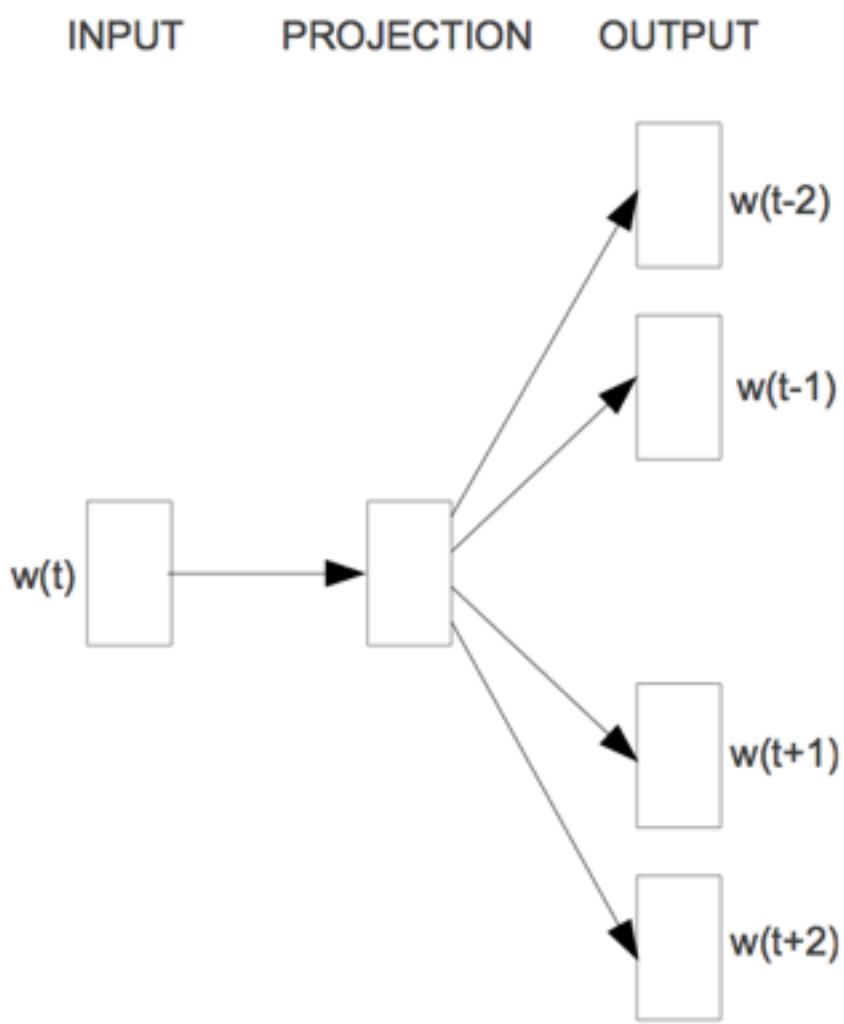


$$P(v_{OUT}|v_{IN})$$

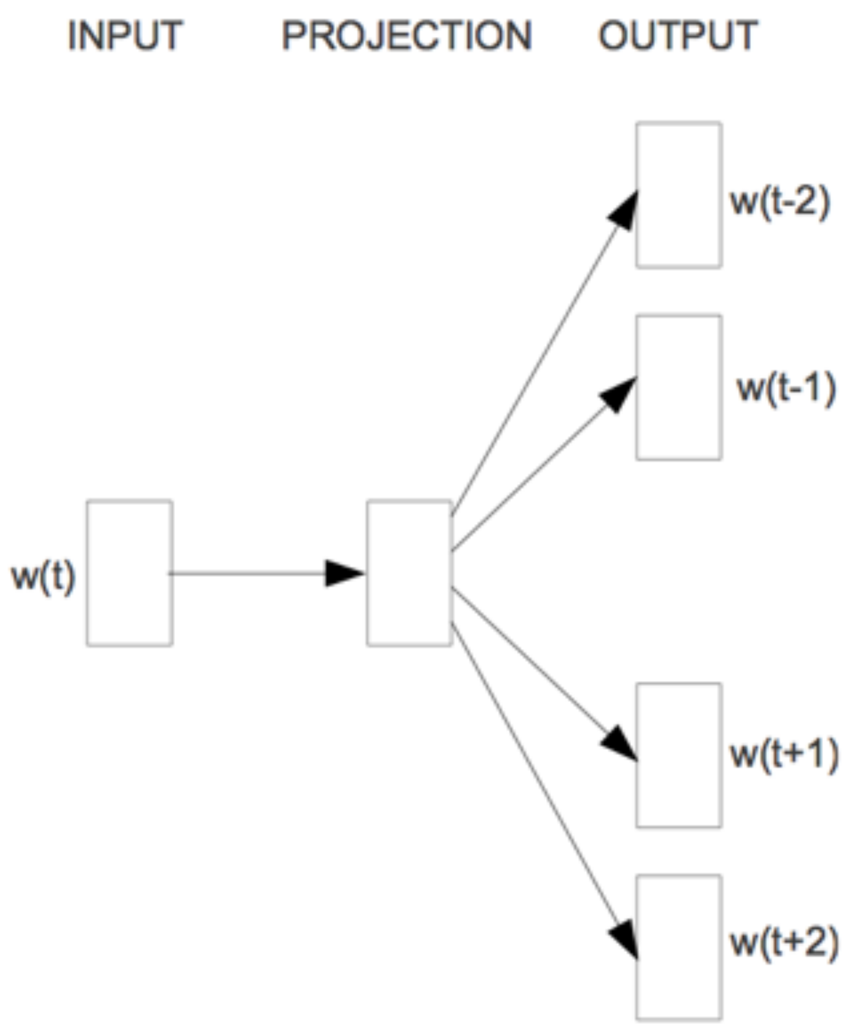


Credit: Tomas Mikolov, Chris Moody

"Skip-gram model"

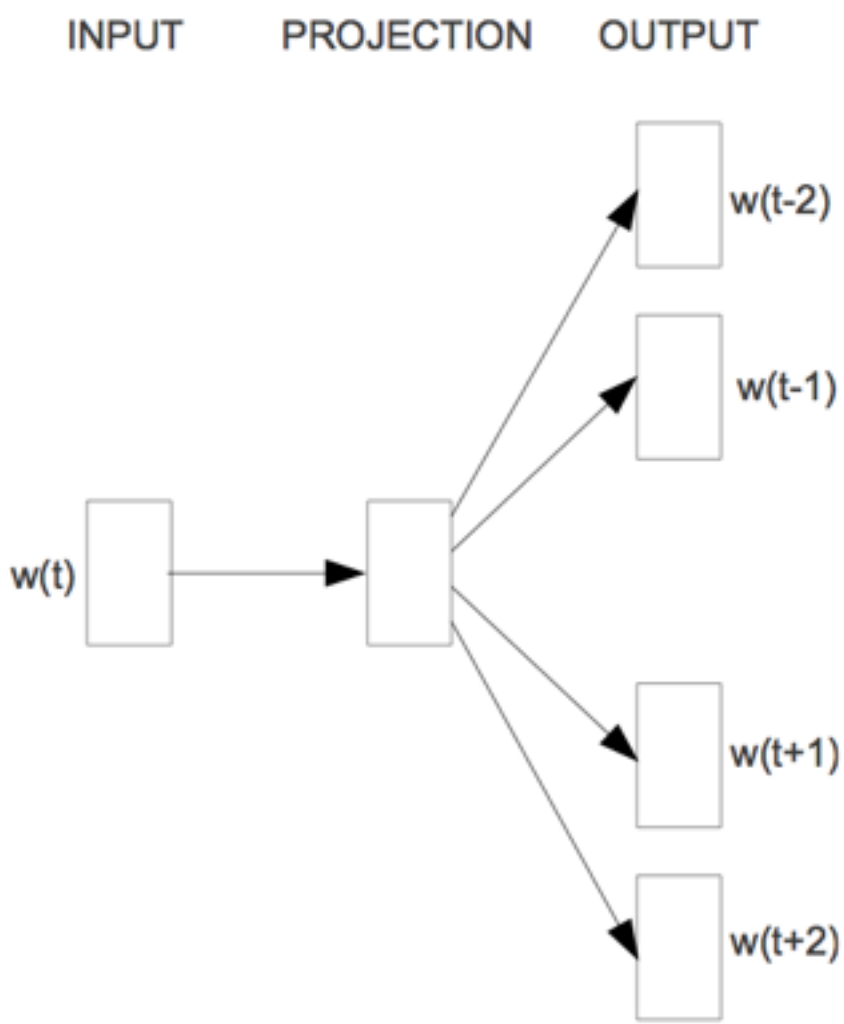


“Skip-gram model”



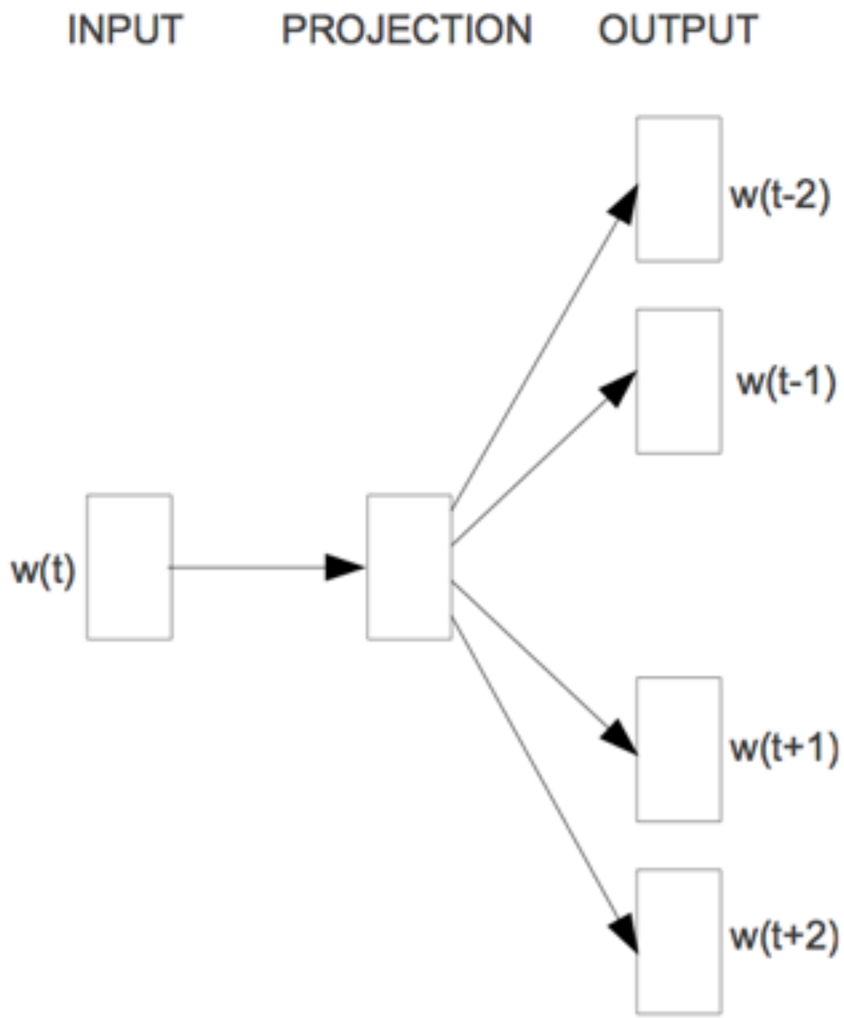
$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0}\log p(w_{t+j}|w_t)$$

"Skip-gram model"



$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0}\log p(w_{t+j}|w_t)$$

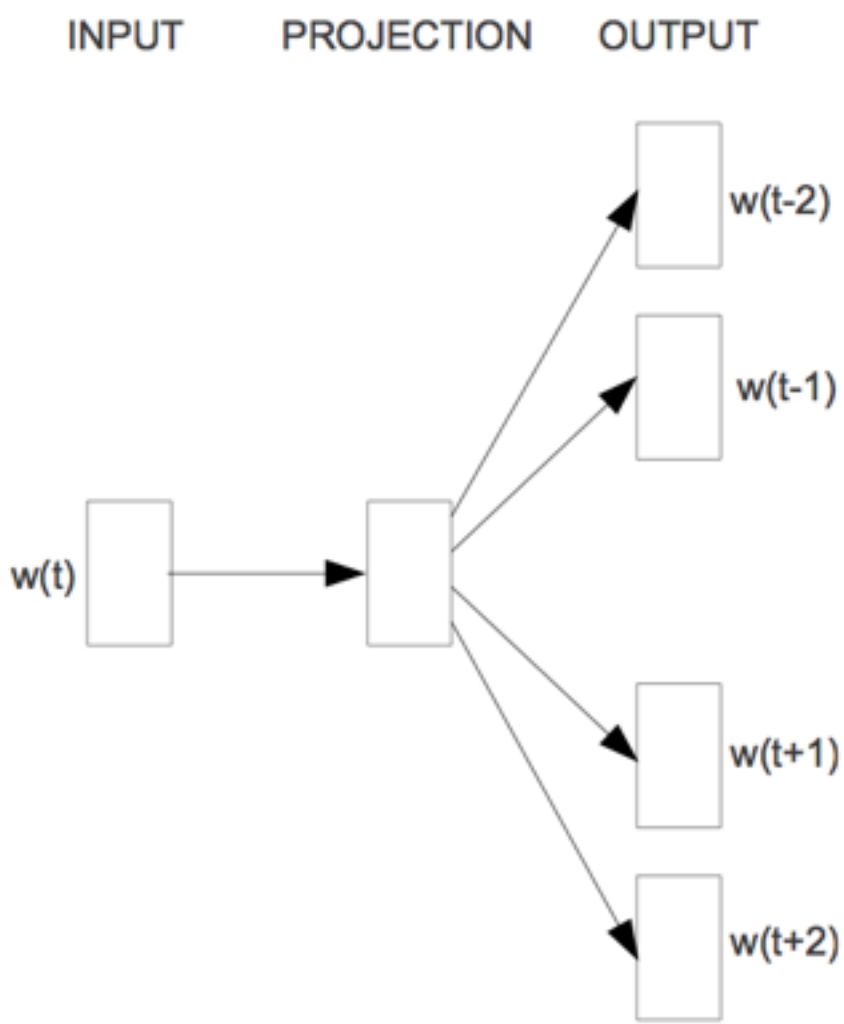
"Skip-gram model"



$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\le j\le c, j\ne 0}\log p(w_{t+j}|w_t)$$

$$p(w_O|w_I) = \frac{\exp\left({v'_{w_O}}^{\top} v_{w_I}\right)}{\sum_{w=1}^{W}\exp\left({v'_w}^{\top} v_{w_I}\right)}$$

"Skip-gram model"

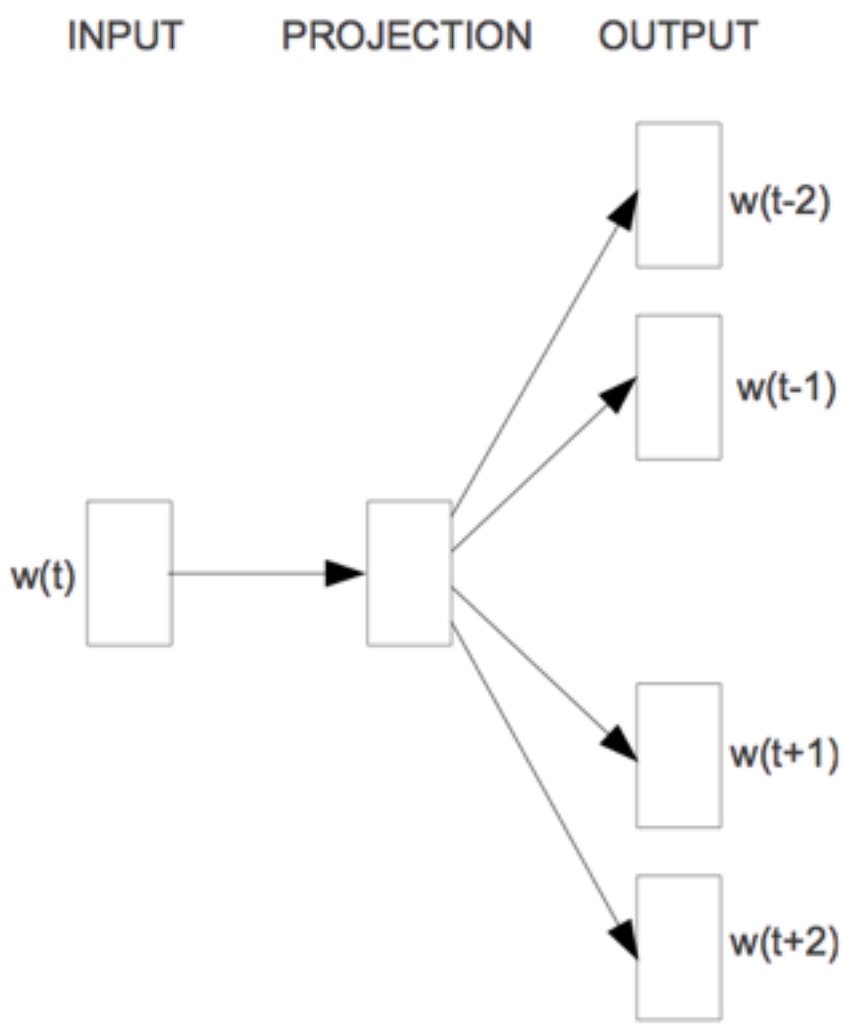


$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0}\log p(w_{t+j}|w_t)$$

$$p(w_O|w_I) = \boxed{\frac{\exp\left({v'_{w_O}}^\top v_{w_I}\right)}{\sum_{w=1}^{W}\exp\left({v'_w}^\top v_{w_I}\right)}}$$

Softmax: Difficult

"Skip-gram model"



$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0}\log p(w_{t+j}|w_t)$$

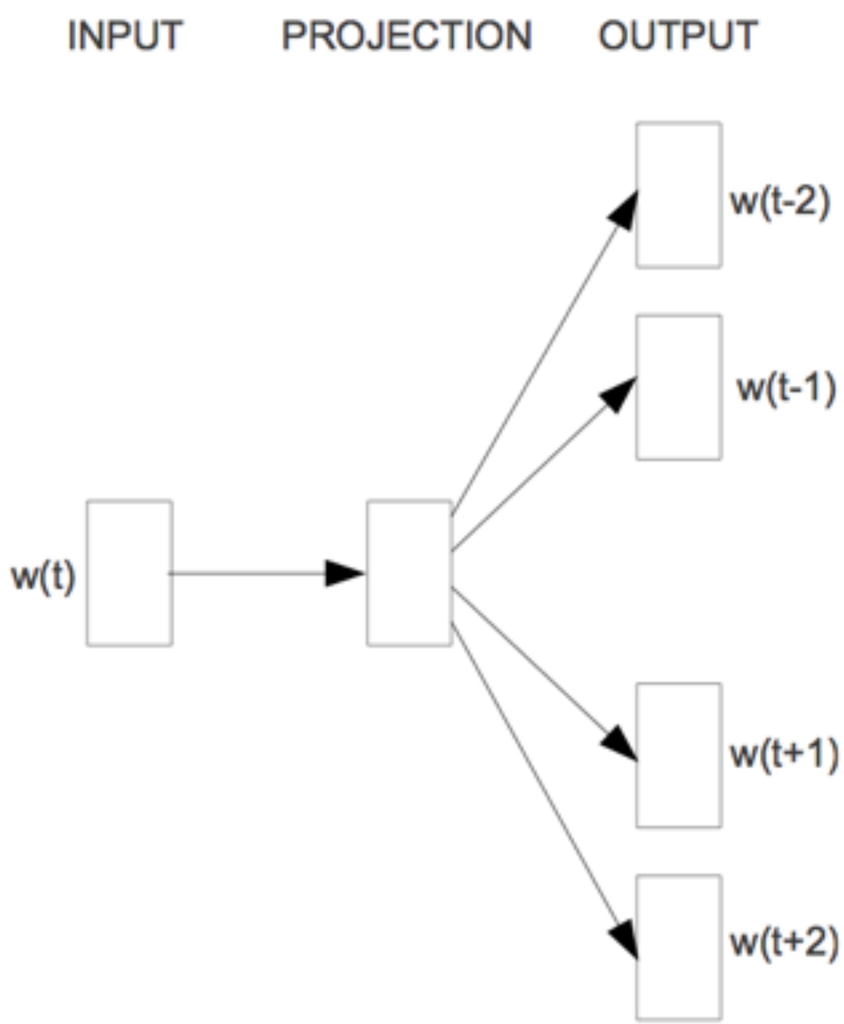
$$p(w_O|w_I) = \boxed{\frac{\exp\left({v'_{w_O}}^\top v_{w_I}\right)}{\sum_{w=1}^{W}\exp\left({v'_w}^\top v_{w_I}\right)}}$$

Softmax: Difficult

Negative sampling (easy)

$$\log\sigma({v'_{w_O}}^\top v_{w_I}) + \sum_{i=1}^{k}\mathbb{E}_{w_i\sim P_n(w)}\left[\log\sigma(-{v'_{w_i}}^\top v_{w_I})\right]$$

"Skip-gram model"



$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0}\log p(w_{t+j}|w_t)$$

$$p(w_O|w_I) = \boxed{\frac{\exp\left({v'_{w_O}}^\top v_{w_I}\right)}{\sum_{w=1}^{W}\exp\left({v'_w}^\top v_{w_I}\right)}}$$

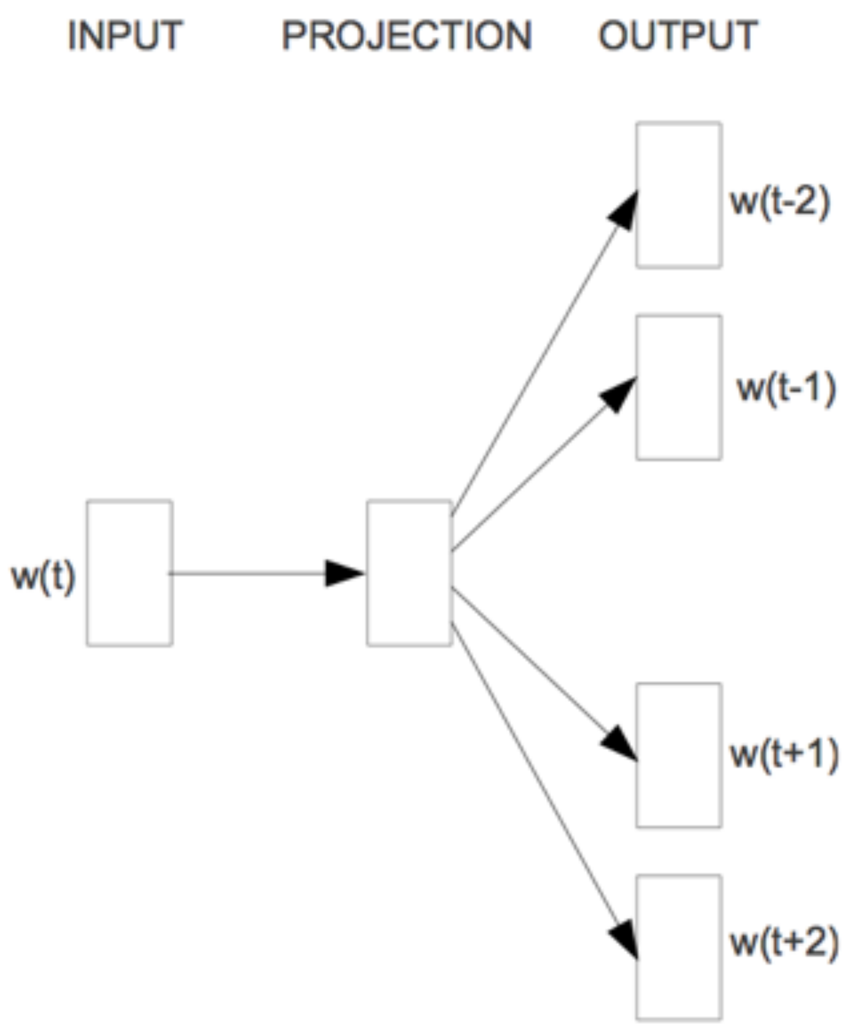
Softmax: Difficult

Negative sampling (easy)

$$\log\sigma({v'_{w_O}}^\top v_{w_I}) + \sum_{i=1}^{k}\mathbb{E}_{w_i\sim P_n(w)}\left[\log\sigma(-{v'_{w_i}}^\top v_{w_I})\right]$$

Actual target word

"Skip-gram model"



$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0}\log p(w_{t+j}|w_t)$$

$$p(w_O|w_I) = \boxed{\frac{\exp\left({v'_{w_O}}^\top v_{w_I}\right)}{\sum_{w=1}^{W}\exp\left({v'_w}^\top v_{w_I}\right)}}$$

Softmax: Difficult

Negative sampling (easy)

random words

$$\log\sigma({v'_{w_O}}^\top v_{w_I}) + \sum_{i=1}^{k}\mathbb{E}_{w_i\sim P_n(w)}\left[\log\sigma(-{v'_{w_i}}^\top v_{w_I})\right]$$

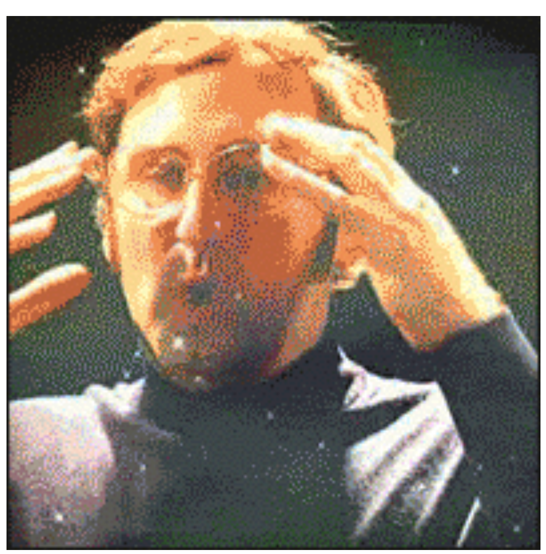
Actual target word

word2vec

If two words *tend to* appear in the **similar context**, they *tend to* have **similar vector representation**

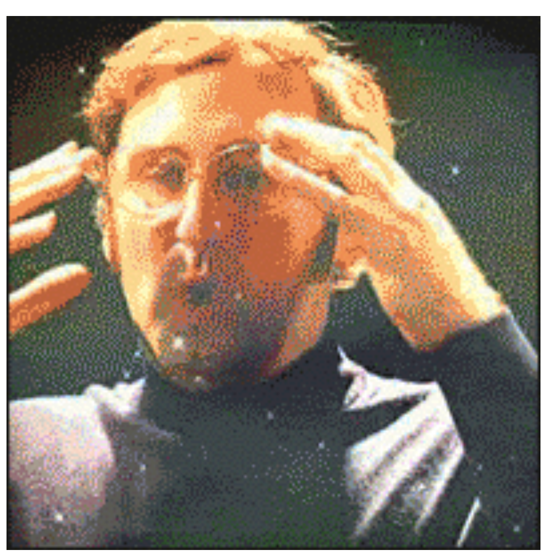
word2vec: analogy

v(**king**) - v(**man**) + v(**woman**) ~ v(**queen**)

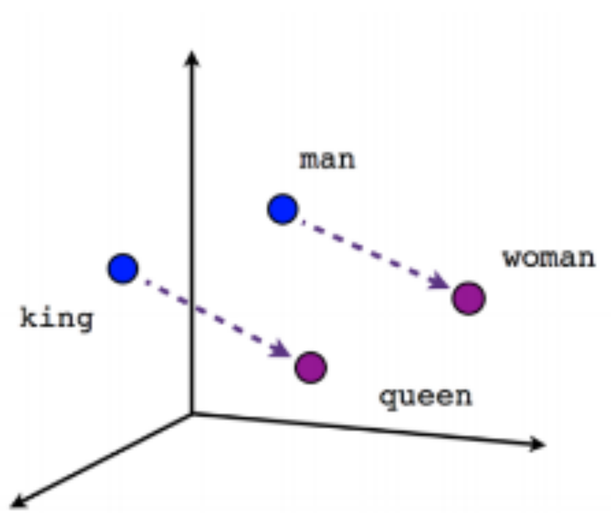


word2vec: analogy

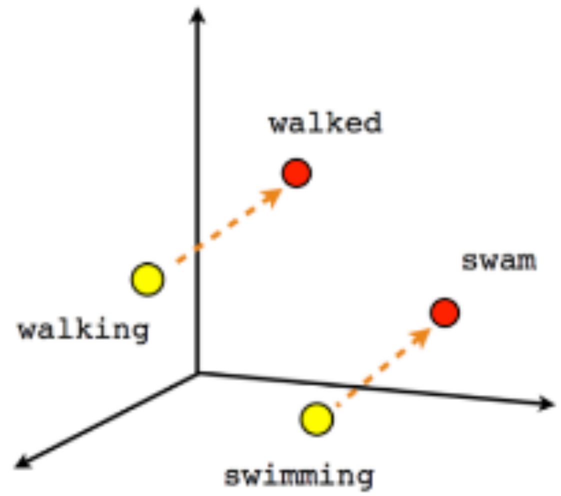
v(**king**) - v(**man**) + v(**woman**) ~ v(**queen**)



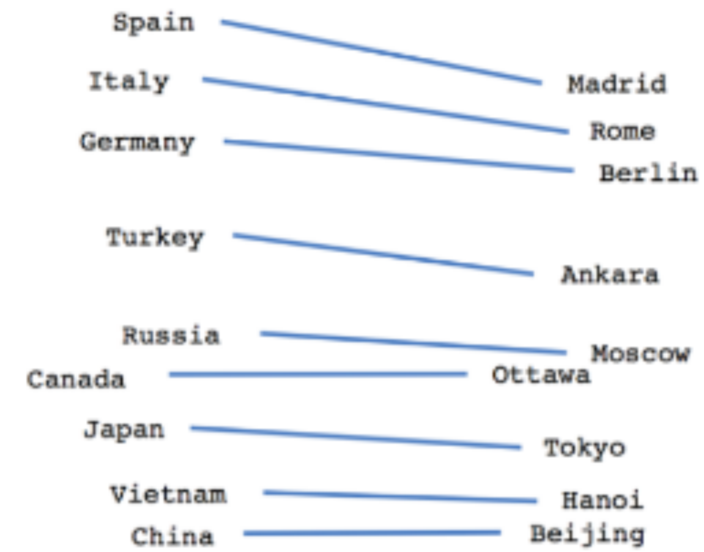
tensorflow.org



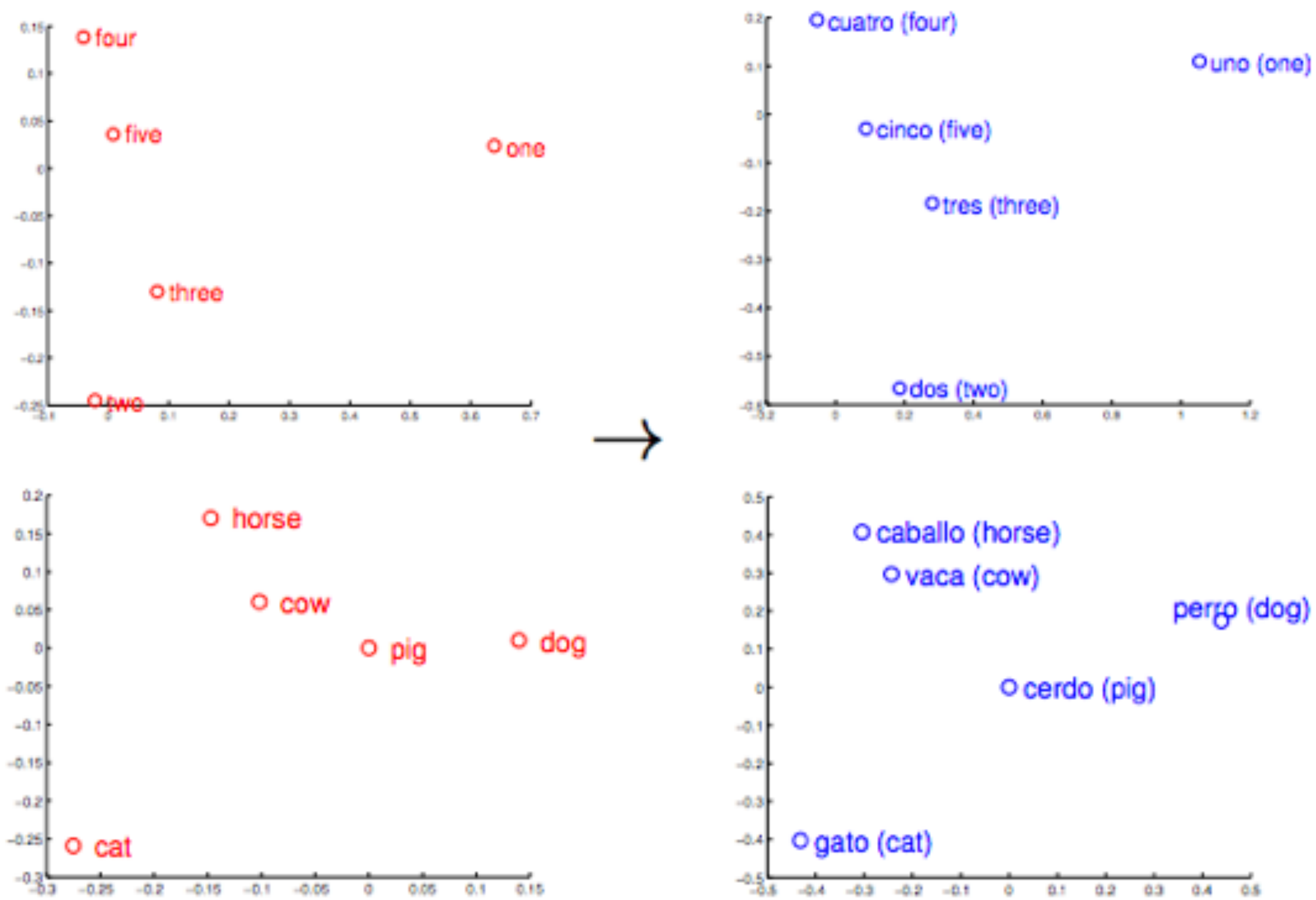
Male-Female



Verb tense



Country-Capital



Mikolov et al. 2013

+ ‘Pregnant’

Credit: Christopher Moody

Linguistic Regularities in Sparse and Explicit Word Representations

Omer Levy* and **Yoav Goldberg**
Computer Science Department
Bar-Ilan University
Ramat-Gan, Israel
{omerlevy,yoav.goldberg}@gmail.com

Abstract

Recent work has shown that neural-embedded word representations capture many relational similarities, which can be recovered by means of vector arithmetic in the embedded space. We show that Mikolov et al.'s method of first adding and subtracting word vectors, and then searching for a word similar to the result, is equivalent to searching for a word that maximizes a linear combination of three pairwise word similarities. Based on this observation, we suggest an improved

word embeddings are designed to capture what Turney (2006) calls *attributional similarities* between vocabulary items: words that appear in similar contexts will be close to each other in the projected space. The effect is grouping of words that share semantic ("dog cat cow", "eat devour") or syntactic ("cars hats days", "emptied carried danced") properties, and are shown to be effective as features for various NLP tasks (Turian et al., 2010; Collobert et al., 2011; Socher et al., 2011; Al-Rfou et al., 2013). We refer to such word representations as *neural embeddings* or just *embeddings*.

Recently, Mikolov et al. (2013c) demonstrated

Linguistic Regularities in Sparse and Explicit Word Representations

Omer Levy* and **Yoav Goldberg**
Computer Science Department

Ab

Recent work has
embedded word r
many relational sir
recovered by mea
in the embedded
Mikolov et al.'s r
and subtracting w
searching for a w
sult, is equivalent t
that maximizes a
three pairwise wor
this observation, w

Neural Word Embedding as Implicit Matrix Factorization

Omer Levy
Department of Computer Science
Bar-Ilan University
omerlevy@gmail.com

Yoav Goldberg
Department of Computer Science
Bar-Ilan University
yoav.goldberg@gmail.com

Abstract

We analyze skip-gram with negative-sampling (SGNS), a word embedding method introduced by Mikolov et al., and show that it is implicitly factorizing

Linguistic Regularities in Sparse and Explicit Word Representations

Omer Levy* and Yoav Gold
Computer Science Departm

Ab

Recent work has
embedded word
many relational sir
recovered by mea
in the embedded
Mikolov et al.'s
and subtracting w
searching for a w
sult, is equivalent
that maximizes a
three pairwise wor
this observation, w

Neur
as Implic

Omer Levy
Department of Computer Sci
Bar-Ilan University
omerlevy@gmail.co

We analyze skip-gram w
method introduced by Mik

Improving Distributional Similarity with Lessons Learned from Word Embeddings

Omer Levy Yoav Goldberg Ido Dagan
Computer Science Department
Bar-Ilan University
Ramat-Gan, Israel
{omerlevy,yogo,dagan}@cs.biu.ac.il

Abstract

Recent trends suggest that neural-network-inspired word embedding models outperform traditional count-based distributional models on word similarity and analogy detection tasks. We reveal that much of the performance gains of word embeddings are due to certain system design choices and hyperparameter optimizations, rather than the embedding

A recent study by Baroni et al. (20
ducts a set of systematic experiments
ing word2vec embeddings to the m
tional distributional methods, such as
mutual information (PMI) matrices (se
and Pantel (2010) and Baroni and Len
for comprehensive surveys). These resul
that the new embedding methods consist
perform the traditional methods by a n
margin on many similarity-oriented tas
ever, state-of-the-art embedding metho

Takeaways

- word2vec != magic. Actually word2vec works similarly to traditional methods (PMI, SVD, …)
- several tweaks in word2vec are actually important and can be transferred to traditional methods.
- Yet, word2vec is a nice method—it's robust, fast, memory efficient.

More takeaways

If two words *tend to* appear in the **similar context**, they *tend to* have **similar vector representation**

- word2vec distance != semantic distance
- it picks up weird word pairs that share weird contexts.

```
>>> model.most_similar(‘teh’)

[(u’ther’, 0.6910992860794067), (u’hte’, 0.6501408815383911), (u’fo’, 0.6458913683891296), (u’tha’, 0.6098173260688782), (u’te’, 0.6042138934135437), (u’ot’, 0.595798909664154), (u’thats’, 0.595078706741333), (u’od’, 0.5908242464065552), (u’tho’, 0.58894944190979), (u’oa’, 0.5846965312957764)]
```

```
>>> model.most_similar(‘pugnacity’)

[(u’pugnaciousness’, 0.6015268564224243), (u’wonkishness’, 0.6014434099197388), (u’pugnacious’, 0.5877301692962646), (u’eloquence’, 0.5875781774520874), (u’sang_froid’, 0.5873805284500122), (u’truculence’, 0.5838015079498291), (u’pithiness’, 0.5773230195045471), (u’irascibility’, 0.5772287845611572), (u’hotheadedness’, 0.5741063356399536), (u’sangfroid’, 0.5715578198432922)]
```

https://quomodocumque.wordpress.com/2016/01/15/messing-around-with-word2vec/

word2vec is not magic!

word2vec is not magic!

word2vec is not magic!

But the one task that word2vec is consistently doing really well (by the formulation):

word2vec is not magic!

But the one task that word2vec is consistently doing really well (by the formulation):

word2vec is not magic!

But the one task that word2vec is consistently doing really well (by the formulation):

pulling out replaceable words.

word2vec is not magic!

But the one task that word2vec is consistently doing really well (by the formulation):

pulling out replaceable words.
(i.e. what are the other words that you'll see in the same context?)

word2vec is not magic!

But the one task that word2vec is consistently doing really well (by the formulation):

pulling out replaceable words.
(i.e. what are the other words that you'll see in the same context?)

word2vec is not magic!

But the one task that word2vec is consistently doing really well (by the formulation):

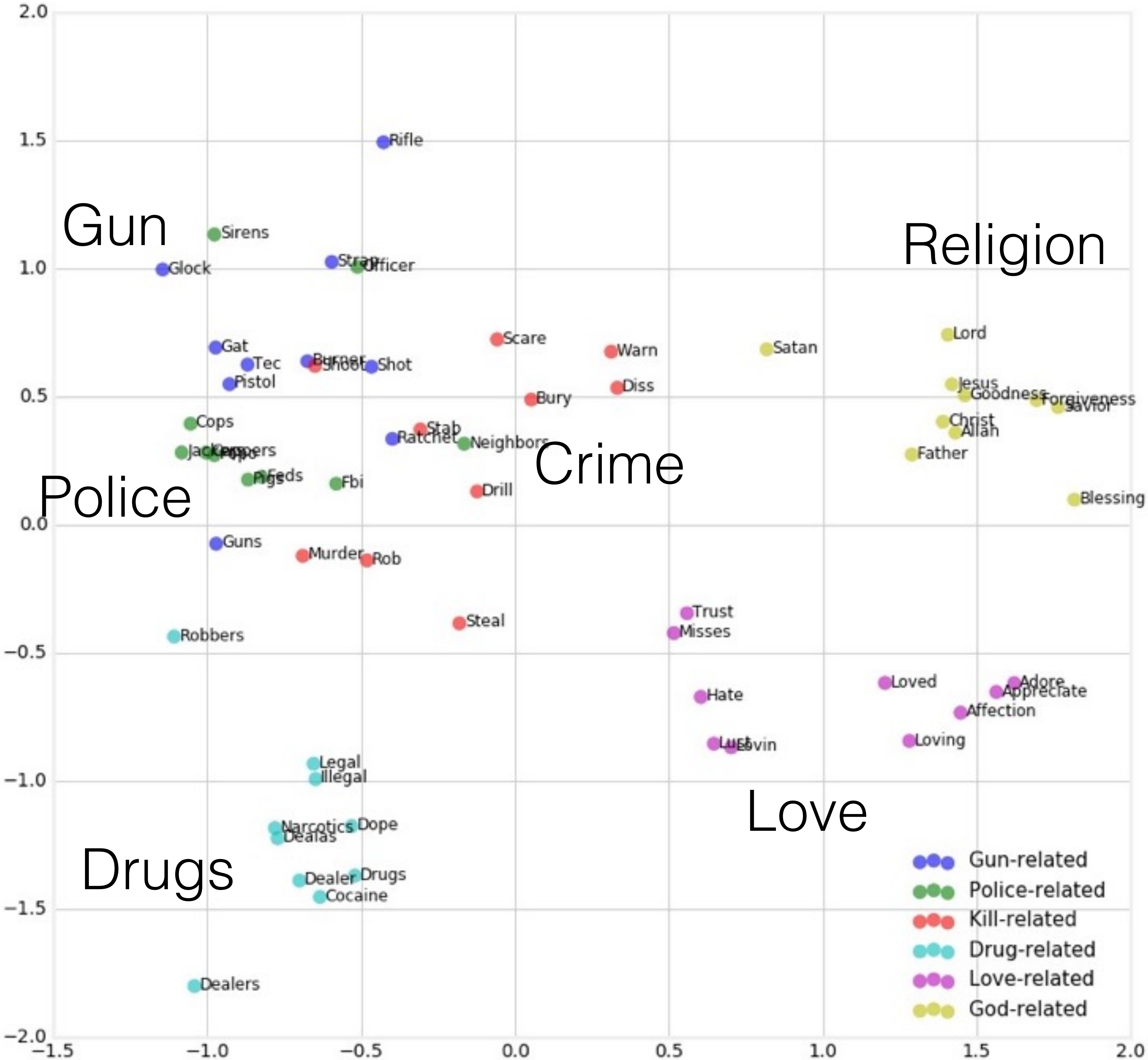
pulling out replaceable words.
(i.e. what are the other words that you'll see in the same context?)

Can be super useful and complement other methods such as topic modeling

Going back to lyrics…

word2vec of hip hop lyrics

gun	money	crack	police
gat	cash	coke	cops
glock	dough	heroin	feds
burner	loot	cocain	coppers
pistol	paper	caine	pigs
rifle	cheese	sacks	popo
ratchet	chips	dope	fbi
guns	cheddar	addicts	sirens
tec	cake	cracks	jackers
strap	scrilla	slingin	officer
shot	dollas	powder	neighbors



Gun
Religion
Crime
Police
Love
Drugs
Rifle
Sirens
Glock
Officer
Scare
Warn
Satan
Lord
Gat
Tec
Shot
Pistol
Bury
Diss
Jesus
Goodness
Forgiveness
Christ
Allah
Father
Cops
Stab
Neighbors
Fbi
Drill
Blessing
Guns
Murder
Rob
Trust
Misses
Robbers
Steal
Loved
Adore
Appreciate
Hate
Affection
Loving
Legal
Illegal
Dope
Narcotics
Dealer
Drugs
Cocaine
Dealers
Gun-related
Police-related
Kill-related
Drug-related
Love-related
God-related

PCA

A sanity check: what are the similar words to “**money**”?

Manual curation	word2vec
cheese	o
cheddar	o
chips	o
dough	o
cream	x
cake	o
scrilla	o
green	x
loot	o
paper	o
Benjamins	o
dead presidents	x
x	cash, dollar, profit, **stacks**, **chedda**, dollars, **funds**, **mail**, **clout**, moneys, **fetti**

REHN, A. L. F., and David Sköld. "‘I Love The Dough’: Rap lyrics as a minor economic literature." Culture and Organization 11.1 (2005): 17-31.

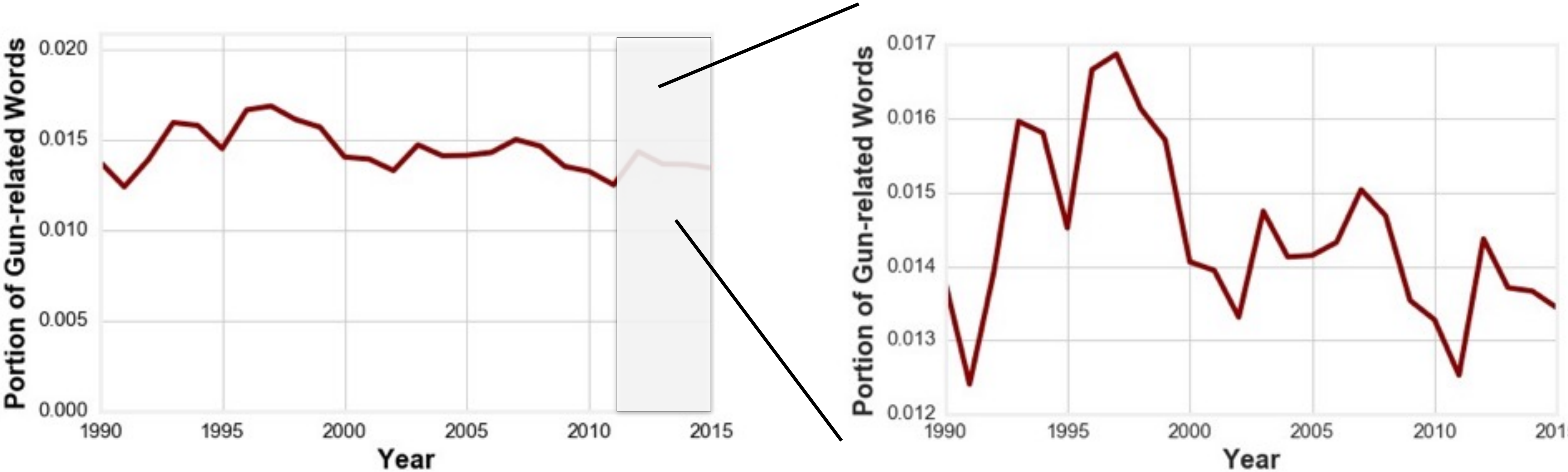
Similar words to “**gun**” (word2vec + manual curation)

heater, uzi, gunshot, ak, tec*, strap**, pistols, grenades, m16, ruger, fired, nines, glocks, 44, 45, gats, magnum, hammers, pistol, guns, sniper, glock, tool, gat, mag, rifle, cap, matic, calico***, blasted, 38, aks, biscuit**, straps, gun, shots, sawedoff, automatic, grenade, sprayed, shot

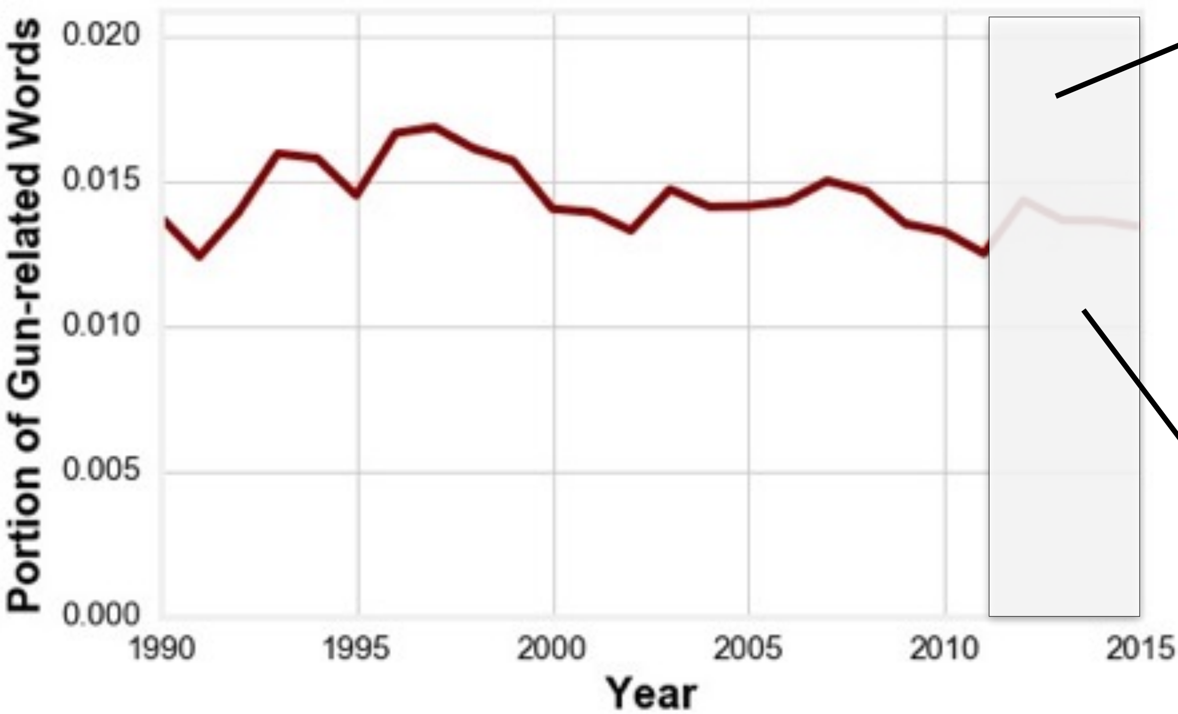
*Representing Tec-9, popular street gun
** A gun or firearm, usually a pistol
*** U.S. weapons manufacturer

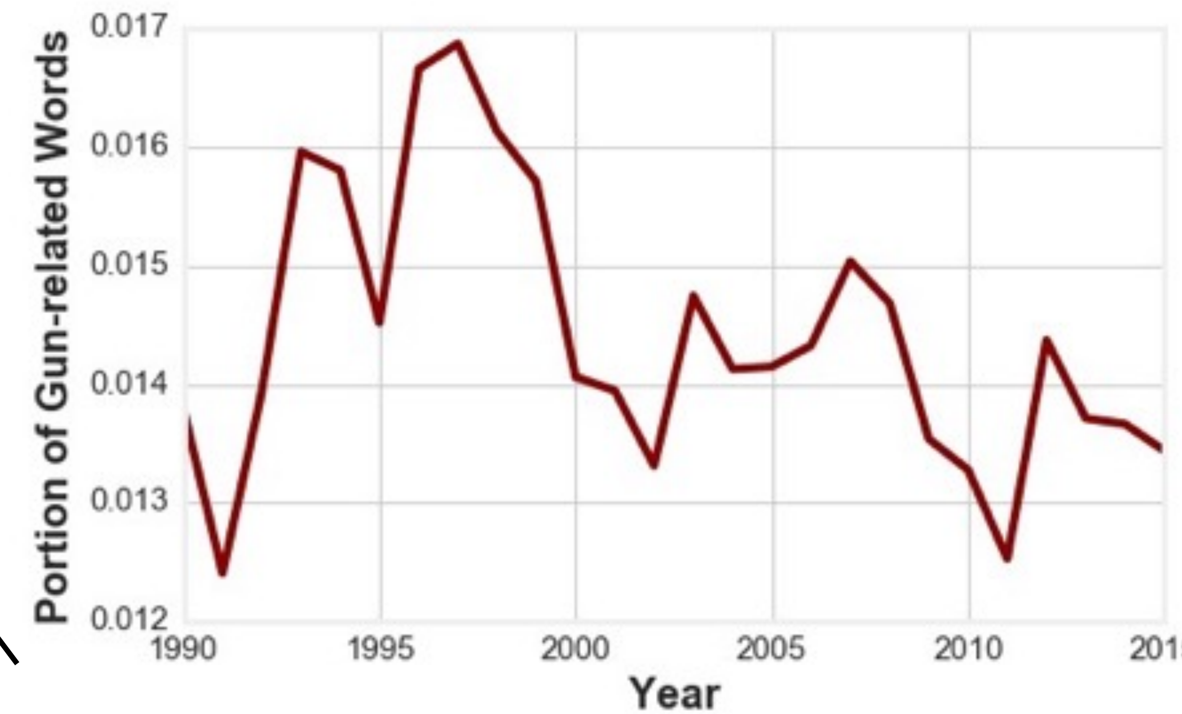
Patterns of individual word usage: what do you see?

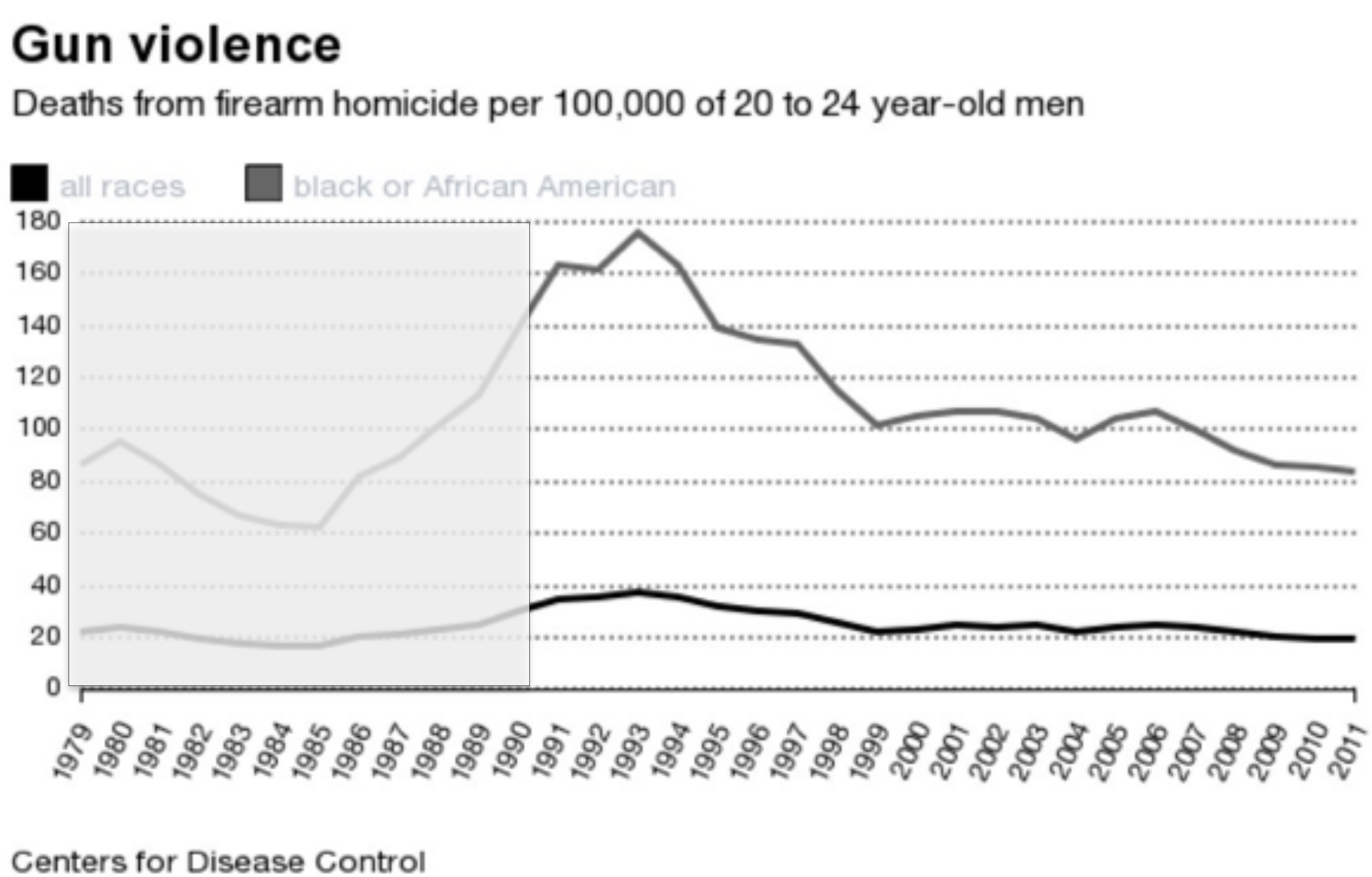
portion of words related to Gun



portion of words related to Gun





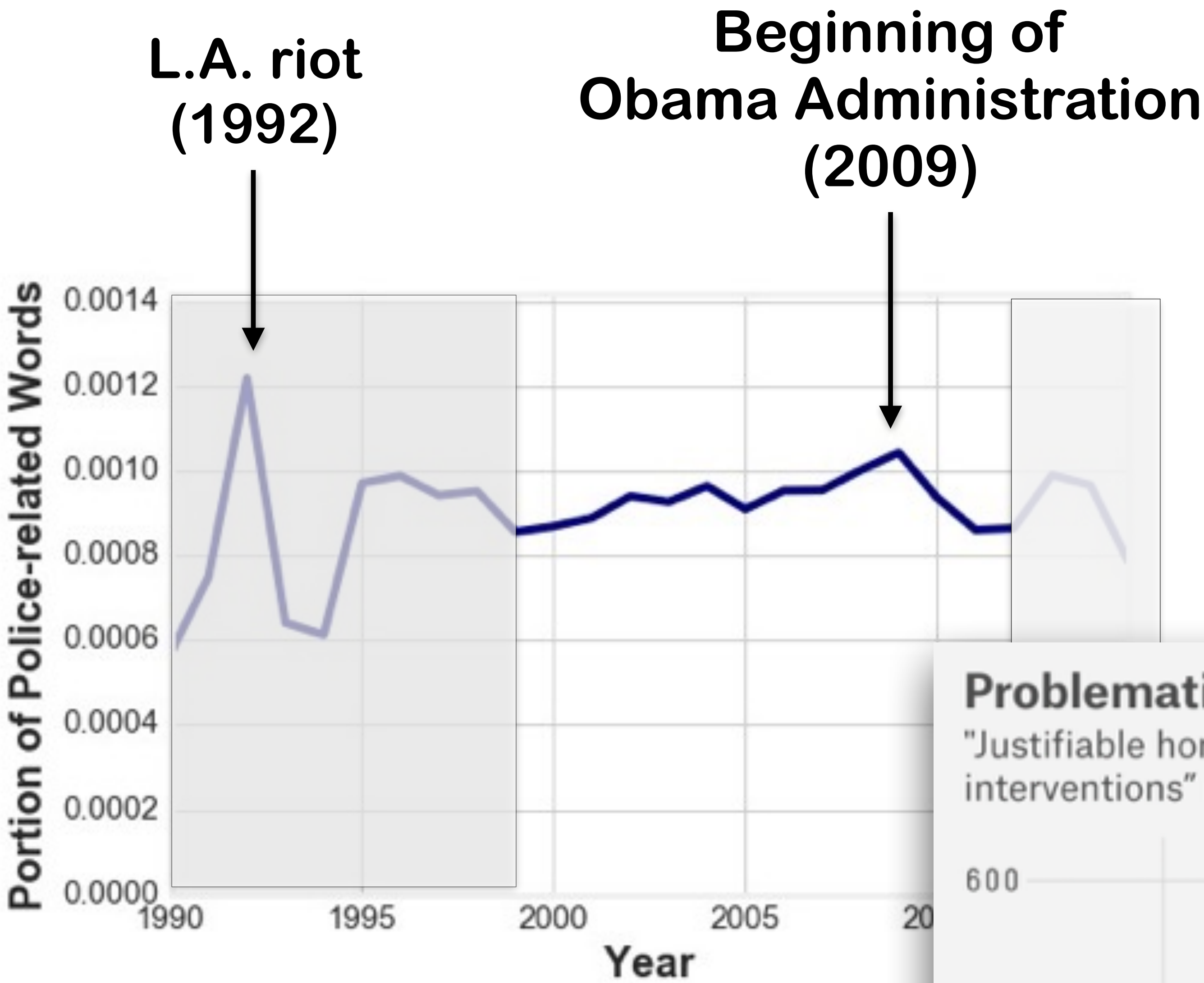


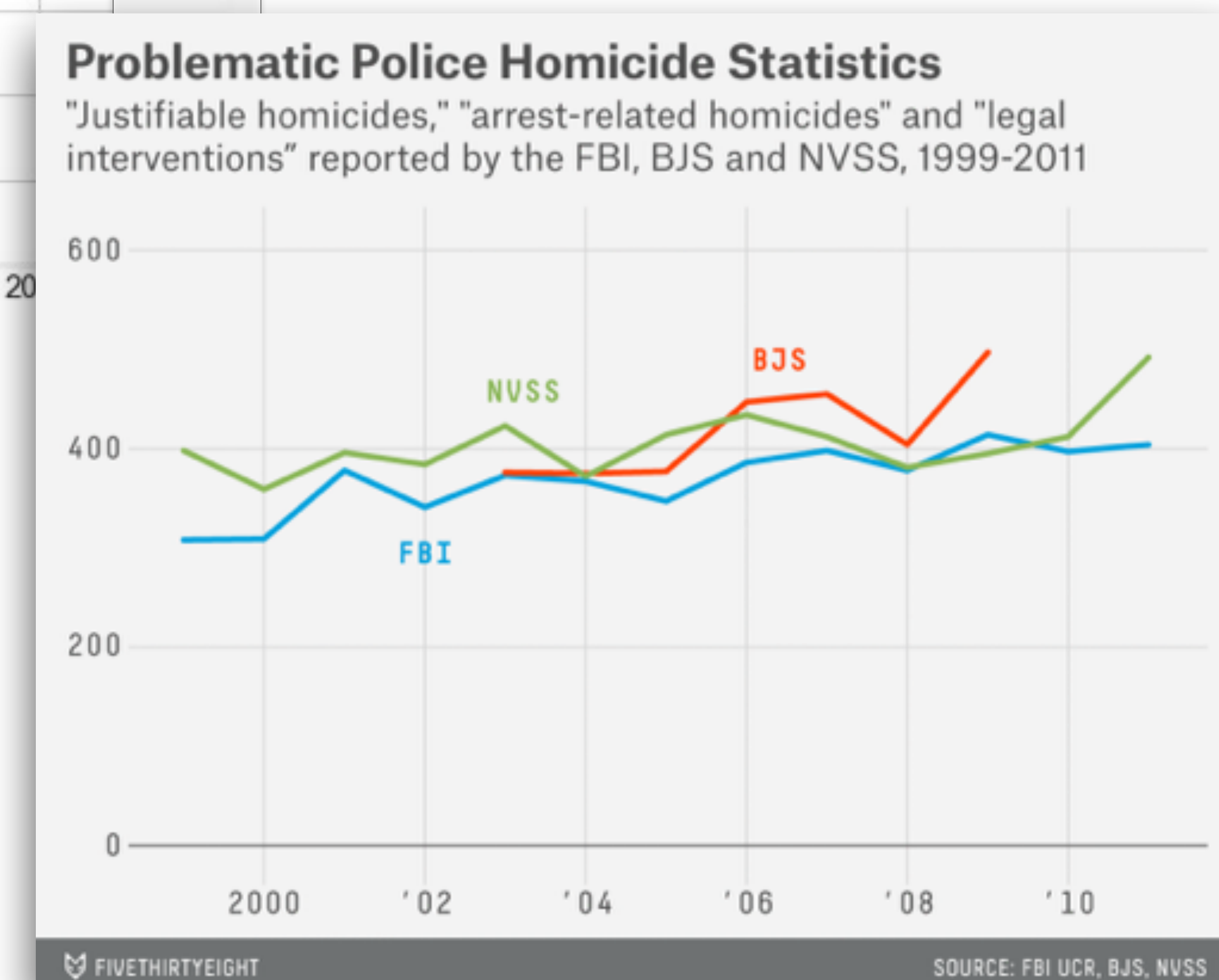
Similar words to “police” (word2vec + manual curation)

popo, official, agents, fbi, badges, coppers, siren, judges, atf*, police, officers, sirens, helicopters, detectives, cops, pigs, feds, cia, popos, fiveoh**, officer

* A part of the government dealing with the control of Alcohol Tobacco and Firearms
** The police in general or a police officer, taken from the 60’s police drama “Hawaii 5-0”

Patterns of individual word usage related to “police”


L.A. riot
(1992)
Beginning of
Obama Administration
(2009)
Portion of Police-related Words
0.0014
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
1990
1995
2000
2005
Year


Problematic Police Homicide Statistics
"Justifiable homicides," "arrest-related homicides" and "legal interventions" reported by the FBI, BJS and NVSS, 1999-2011
600
400
200
0
NVSS
BJS
FBI
2000
'02
'04
'06
'08
'10
FIVETHIRTYEIGHT
SOURCE: FBI UCR, BJS, NVSS

How about the analogy?

n**** - man + woman	money - success + kill	money - kill + success
hoe	rob	moneys
bitch	smack	patience
chick	murda	currency
trick	slap	dough
slut	shoot	potential
yous	n****	wealth
scrub	racks	marriage
lady	kidnap	progress
slug	toss	relationships
girl	fuck	cash

More analogy fun

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']
```

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']
```

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']
```

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']

city - bad + good
['sky', 'hood', 'sun', 'california', 'place', 'skies', 'garden', 'sunshin
e', 'stars', 'winter']

city - good + bad
['town', 'committee', 'kansas', 'boston', 'texas', 'country', 'georgia',
'clique', 'pjs', 'detroit']
```

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']
```

```
city - bad + good
['sky', 'hood', 'sun', 'california', 'place', 'skies', 'garden', 'sunshin
e', 'stars', 'winter']

city - good + bad
['town', 'committee', 'kansas', 'boston', 'texas', 'country', 'georgia',
'clique', 'pjs', 'detroit']
```

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']
```

```
city - bad + good
['sky', 'hood', 'sun', 'california', 'place', 'skies', 'garden', 'sunshin
e', 'stars', 'winter']

city - good + bad
['town', 'committee', 'kansas', 'boston', 'texas', 'country', 'georgia',
'clique', 'pjs', 'detroit']
```

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']
```

```
city - bad + good
['sky', 'hood', 'sun', 'california', 'place', 'skies', 'garden', 'sunshin
e', 'stars', 'winter']

city - good + bad
['town', 'committee', 'kansas', 'boston', 'texas', 'country', 'georgia',
'clique', 'pjs', 'detroit']
```

```
neighborhood - bad + good
['hood', 'spotlight', 'spot', 'homeys', 'window', 'sky', 'projects', 'woo
ds', 'garden', 'turf']

neighborhood - good + bad
['neighbourhood', 'pjs', 'clique', 'stalk', 'stalkin', 'city', 'flocks',
'murderin', 'smackin', 'pd']
```

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']
```

```
city - bad + good
['sky', 'hood', 'sun', 'california', 'place', 'skies', 'garden', 'sunshin
e', 'stars', 'winter']

city - good + bad
['town', 'committee', 'kansas', 'boston', 'texas', 'country', 'georgia',
'clique', 'pjs', 'detroit']
```

```
neighborhood - bad + good
['hood', 'spotlight', 'spot', 'homeys', 'window', 'sky', 'projects', 'woo
ds', 'garden', 'turf']

neighborhood - good + bad
['neighbourhood', 'pjs', 'clique', 'stalk', 'stalkin', 'city', 'flocks',
'murderin', 'smackin', 'pd']
```

More analogy fun

```
government - bad + good
['salvation', 'gods', 'freedom', 'opportunity', 'ourselves', 'highest',
'governments', 'culture', 'community', 'liberty']

government - good + bad
['judges', 'officers', 'federal', 'clones', 'henchmen', 'fbi', 'cia', 'mu
rderin', 'corporate', 'corrupt']
```

```
city - bad + good
['sky', 'hood', 'sun', 'california', 'place', 'skies', 'garden', 'sunshin
e', 'stars', 'winter']

city - good + bad
['town', 'committee', 'kansas', 'boston', 'texas', 'country', 'georgia',
'clique', 'pjs', 'detroit']
```

```
neighborhood - bad + good
['hood', 'spotlight', 'spot', 'homeys', 'window', 'sky', 'projects', 'woo
ds', 'garden', 'turf']

neighborhood - good + bad
['neighbourhood', 'pjs', 'clique', 'stalk', 'stalkin', 'city', 'flocks',
'murderin', 'smackin', 'pd']
```

Can we compare the two cultures?

(*This is still a mixture of many types of music)

“Cash”

Bold: unique in either one dataset

Hip hop	Ultimate guitar
dough	**greenback**
money	**stash**
loot	**nickel**
cheese	dough
checks	money
stacks	**credit**
cheddar	**dope**
chips	**jewelry**
scrill	**dime**
paper	**penny**

“Booze”

Hip hop	Ultimate guitar
henny	**whisky**
bacardi	**morgan**
gin	bologna
cognac	**swig**
alcohol	**pepsi**
cranberry	peanut
brew	sushi
newports	cuban
vodka	hookers
paper	hustlers

“Drug”

Hip hop	Ultimate guitar
dealers	dealer
dealer	**gateway**
drugs	**heroin**
dope	**sniff**
illegal	deparment
cocaine	charity
narcotics	fiend
dealas	**addiction**
robbers	n****
legal	friction

“they”

Hip hop	Ultimate guitar
you	we
feds	**neighbors**
i	unjust
cops	**doctors**
them	**lawyers**
jealous	**kids**
police	**neighbours**
us	**babies**
theyre	they’d
others	**dealers**

“Society”

	HH	UG
0	poverty	lovers'
1	menace	definition
2	political	group
3	hatred	corporate
4	consequence	media
5	conspiracy	dissent
6	media	leader
7	insanity	morality
8	governments	germany
9	prophets	childbirth

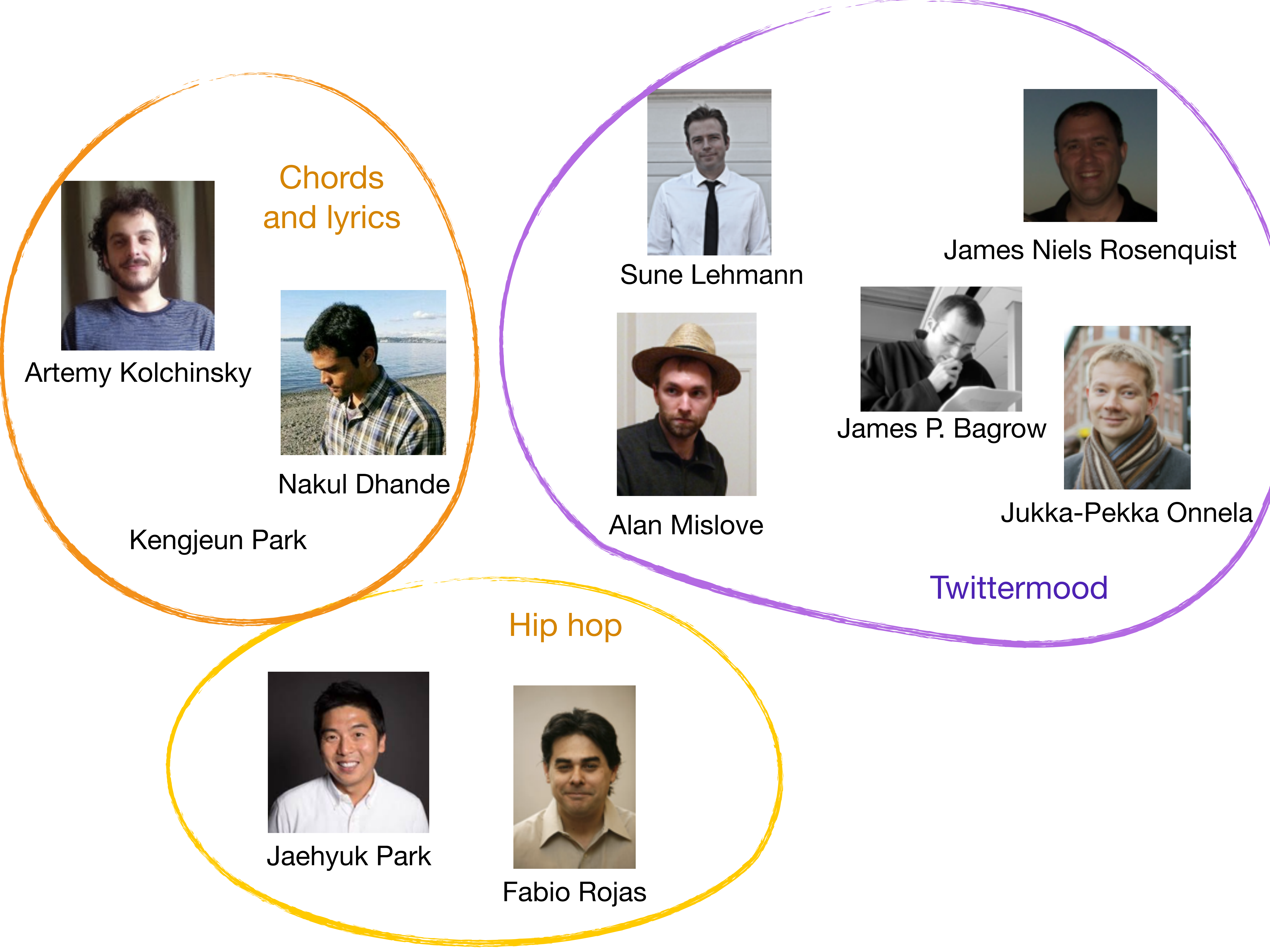
“Food”

	HH	UG
0	stamps	meat
1	dinner	meals
2	steak	fries
3	breakfast	eat
4	meals	pork
5	fridge	taster
6	plate	professional
7	lunch	steak
8	meal	vegetables
9	eggs	chinese

Large-scale Lyrics datasets can tell us so many interesting stories about culture.

JANYSANALYTICS

http://janysanalytics.com/


Chords
and lyrics
Artemy Kolchinsky
Nakul Dhande
Kengjeun Park
Sune Lehmann
James Niels Rosenquist
Alan Mislove
James P. Bagrow
Jukka-Pekka Onnela
Twittermood
Hip hop
Jaehyuk Park
Fabio Rojas

Chords and lyrics

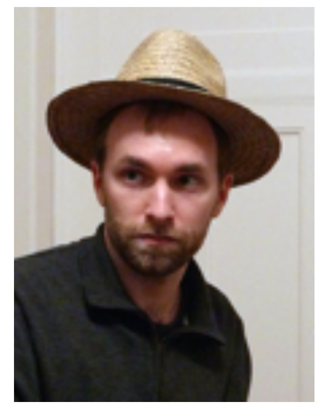
Artemy Kolchinsky

Nakul Dhande

Kengjeun Park

Sune Lehmann

James Niels Rosenquist



Alan Mislove

James P. Bagrow

Jukka-Pekka Onnela

Twittermood

Hip hop

Jaehyuk Park

Fabio Rojas

urban DICTIONARY