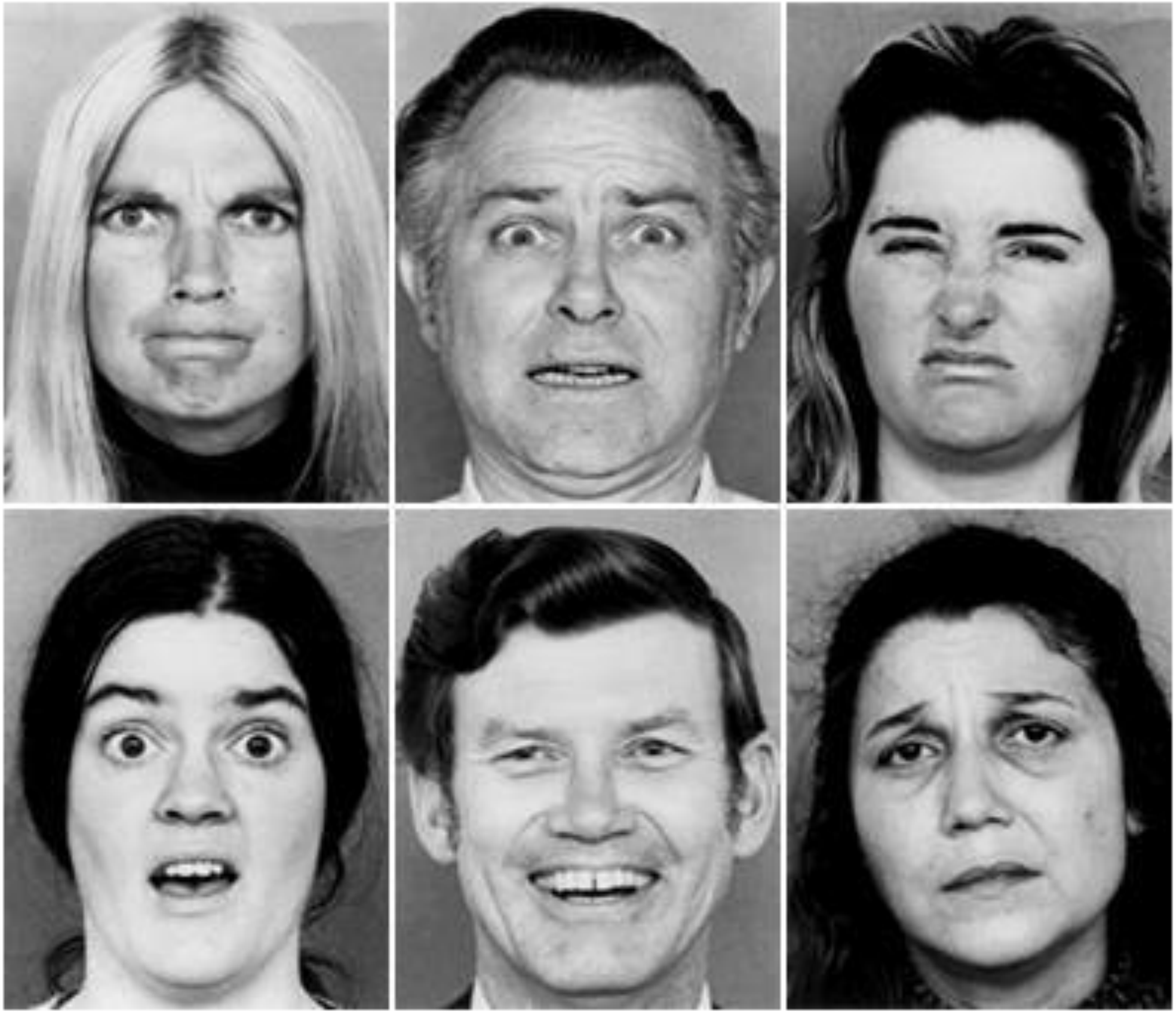
CS 6474/4803 Social Computing: Analyzing Language II

Munmun De Choudhury

munmund@gatech.edu

Week 9 | October 18, 2017

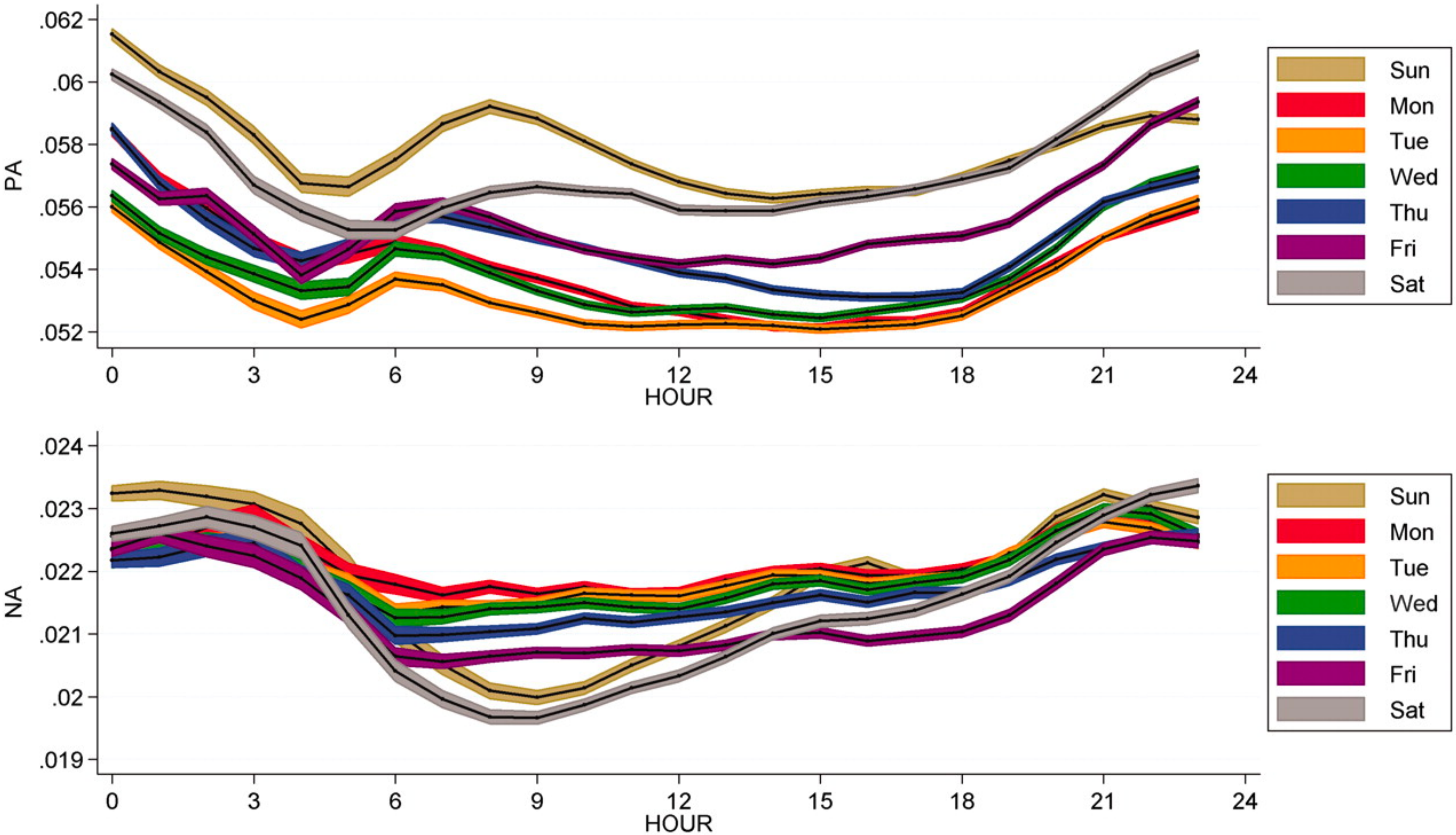
Language is the most common and reliable way for people to translate their internal thoughts and emotions into a form that others can understand. Words and language, then, are the very stuff of psychology and communication -- Tauszczik & Pennebaker

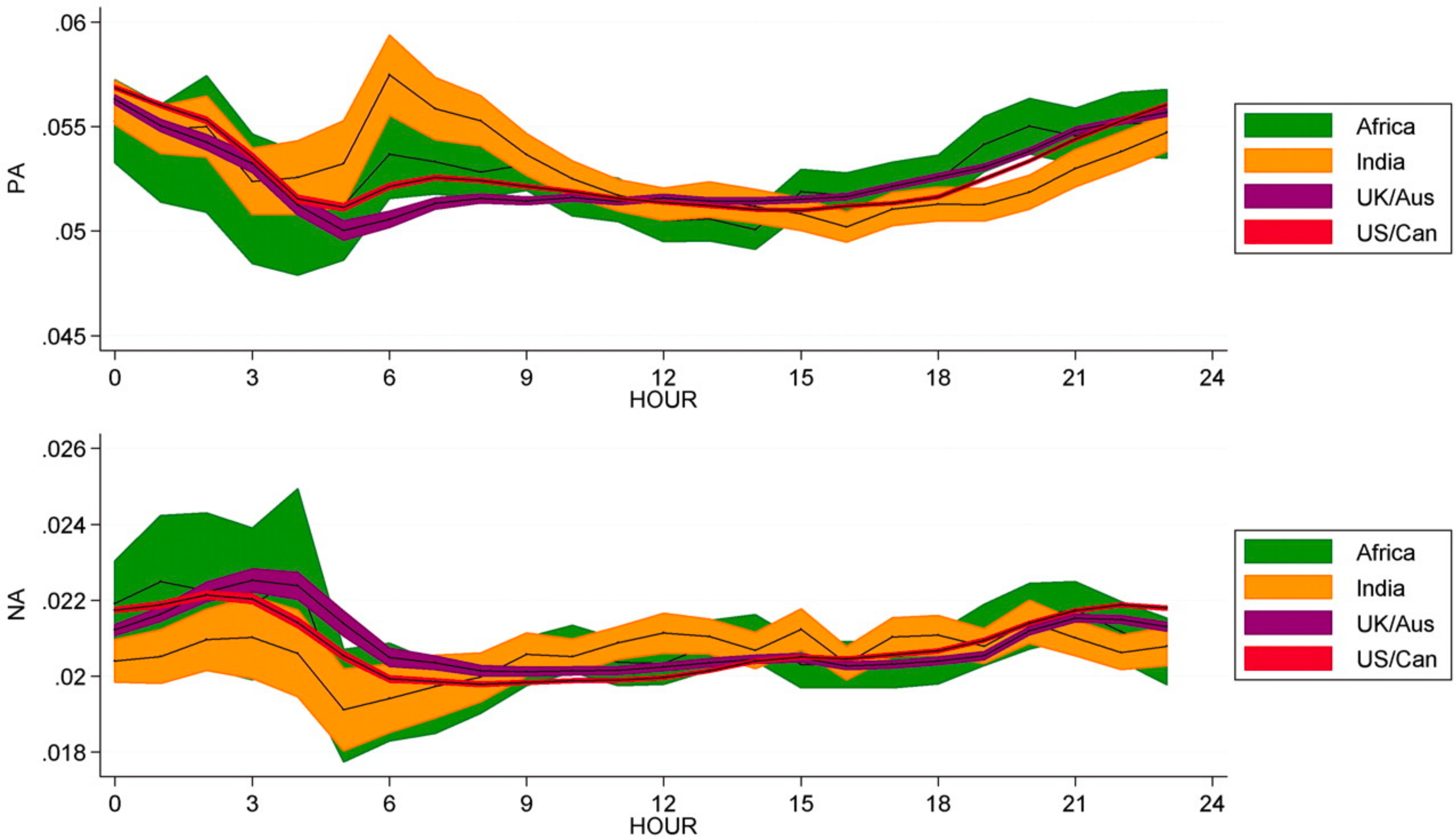


Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures

Summary

- One of the early works examining relationship between social media mood and behavior and psychological theories.
- The potential of online social media to study individual behavior.
- Identify daily and seasonal mood variations and relate it to work, sleep and daylight.
- Validate circadian rhythms in humans.
  - PA spike in the morning, NA increases as the day progresses
- Measure positive affect and negative affect based on the lexicon LIWC.
- PA and NA are not mirror images of each other.


PA
.062
.06
.058
.056
.054
.052
0
3
6
9
12
15
18
21
24
HOUR
Sun
Mon
Tue
Wed
Thu
Fri
Sat
NA
.024
.023
.022
.021
.02
.019
0
3
6
9
12
15
18
21
24
HOUR
Sun
Mon
Tue
Wed
Thu
Fri
Sat



PA
.06
.055
.05
.045
0
3
6
9
12
15
18
21
24
HOUR
Africa
India
UK/Aus
US/Can
NA
.026
.024
.022
.02
.018
0
3
6
9
12
15
18
21
24
HOUR
Africa
India
UK/Aus
US/Can

Twitter is used by millions and both the papers extensively leverage this source of data in measuring mood and affect.

How does use of Twitter for this purpose address limitations in existing mood or affect measurement methods?

Twitter is used by millions and both the papers extensively leverage this source of data in measuring mood and affect.

But could Twitter also have bias?

How do you expect the results relating to mood to be different if the paper used: 1) Facebook 2) Instagram?

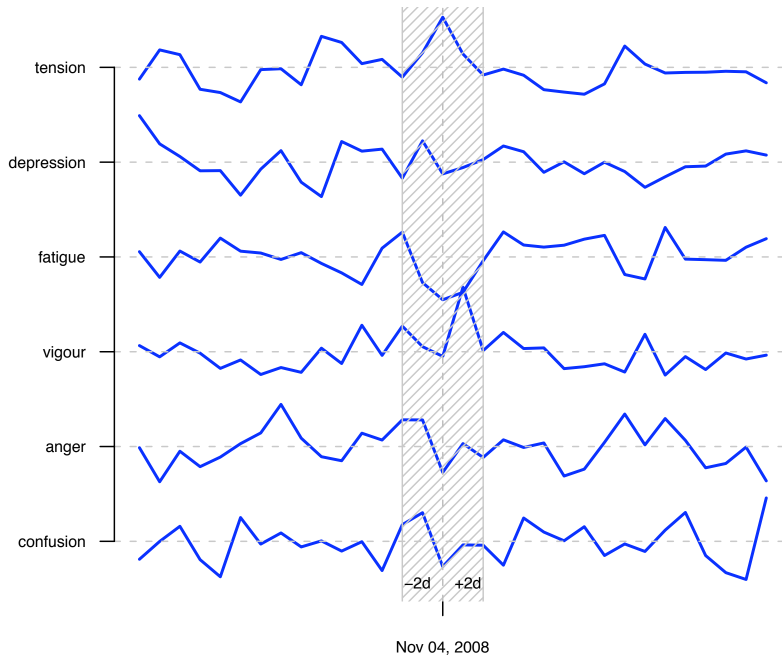
Could platform affordances impact specific moods and their manifestations on social media? How?

Class Exercise

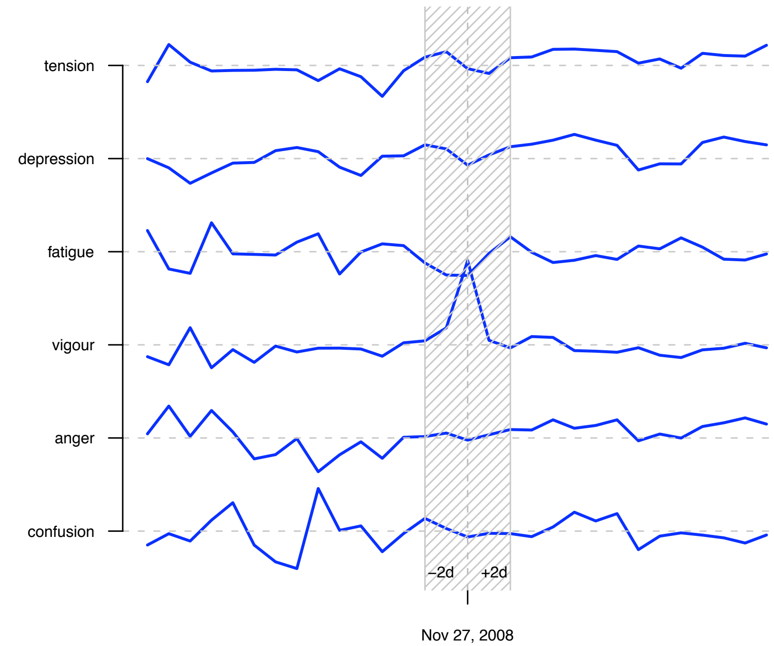
Why is measuring mood useful?
Some examples follow…

Modeling Public Mood and Emotion: Twitter Sentiment and Socioeconomic Phenomena – (Bollen, Pepe, Mao, 2010)

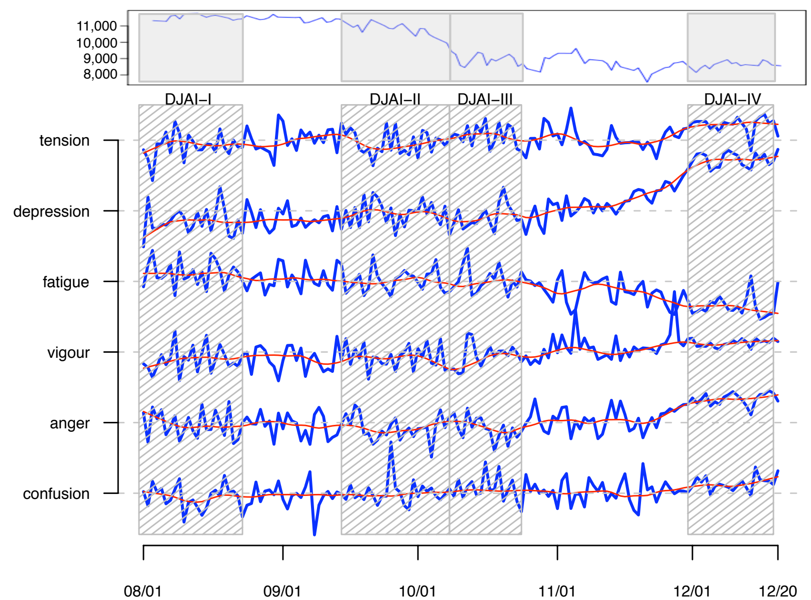
- Examine how Twitter moods reflect social, political, and economic events
- Use POMS (profile of mood states) for detecting mood-indicative twitter posts.
  - POMS dimensions: tension, depression, anger, vigor, fatigue and confusion
- Investigate how a six vector representation of moods deviates during different big scope events.
- High stress/tension during elections; excitement/vigor during thanksgiving.



2008 Presidential elections

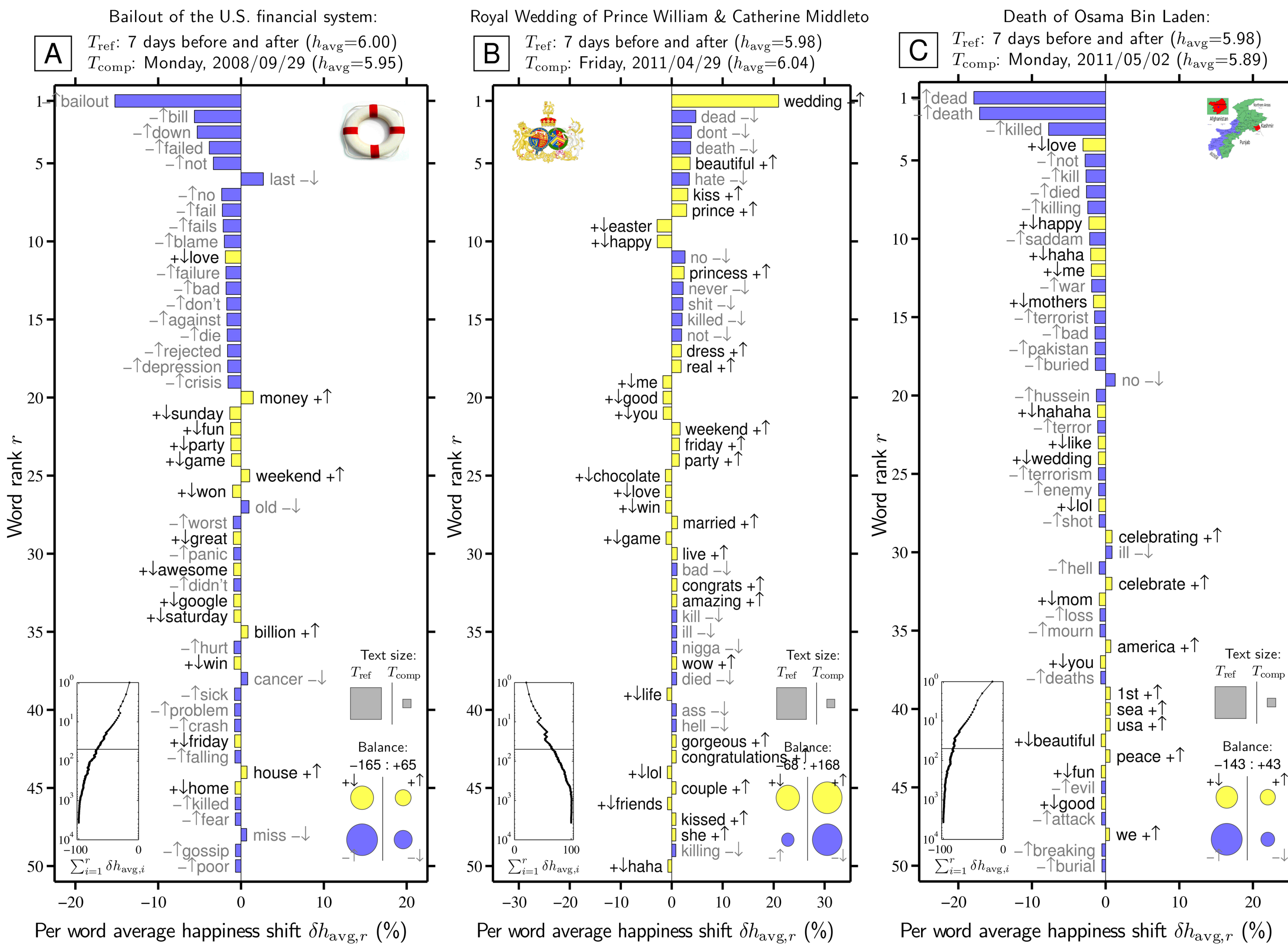


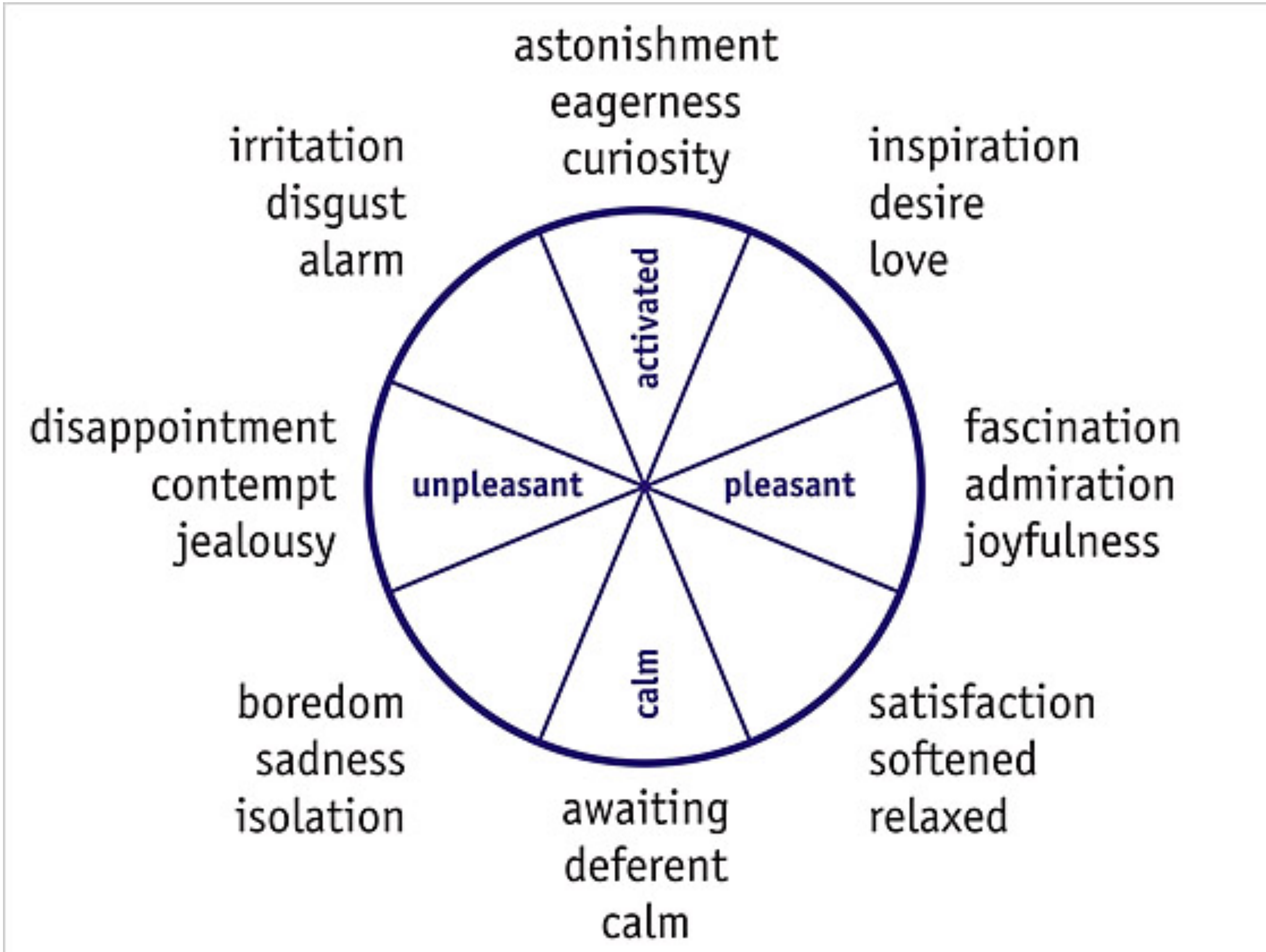
Thanksgiving



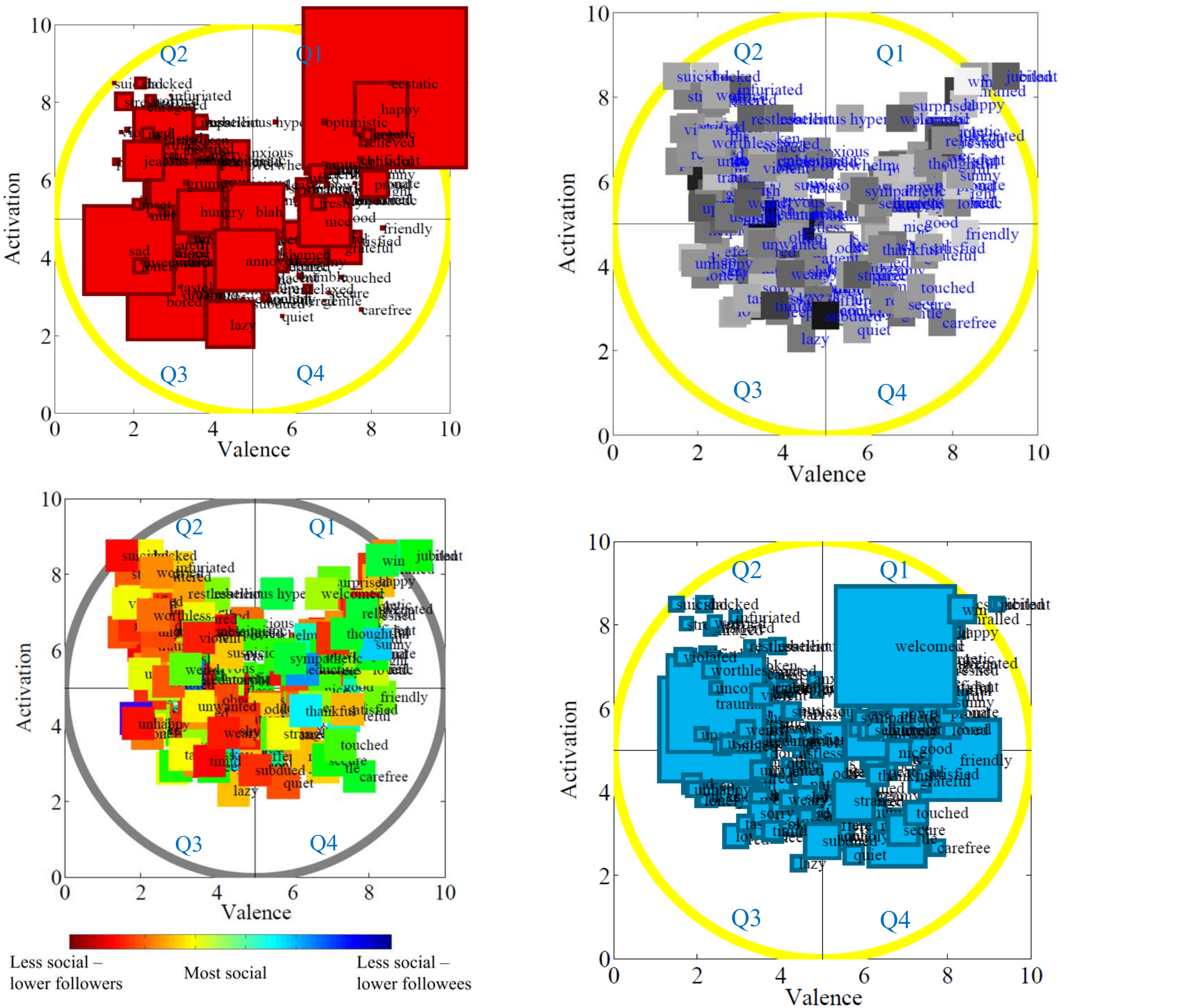
Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter

Peter Sheridan Dodds , Kameron Decker Harris, Isabel M. Kloumann, Catherine A. Bliss, Christopher M. Danforth




astonishment
eagerness
curiosity
inspiration
desire
love
fascination
admiration
joyfulness
satisfaction
softened
relaxed
awaiting
deferent
calm
boredom
sadness
isolation
disappointment
contempt
jealousy
irritation
disgust
alarm
activated
pleasant
calm
unpleasant

Not All Moods Are Created Equal! Exploring Human Emotional States in Social Media, by De Choudhury, Counts, and Gamon 2012



- **Emotion** – brief conscious experience characterized by intense mental activity and a high degree of pleasure or displeasure
- **Affect** – an instinctual reaction to stimulation occurring before the typical cognitive processes considered necessary for the formation of a more complex emotion
- **Mood** – emotional state. Moods differ from emotions or affects in that they are less specific, less intense, and less likely to be triggered by a particular stimulus or event
- **Sentiment** – attitude or opinion with respect to a specific topic, event or situation

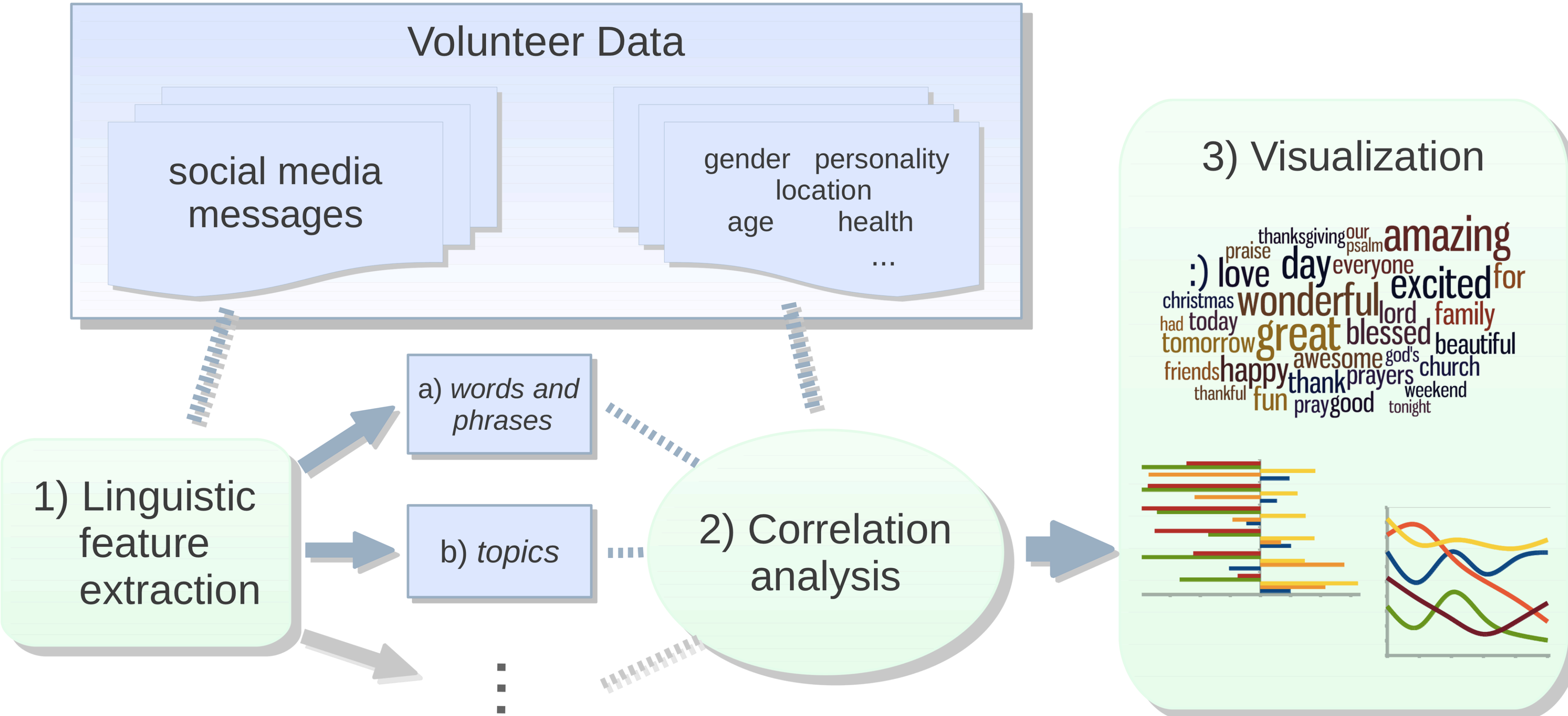
Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach

Summary

- Facebook data of 75K individuals
- Users took personality tests
    - Participants volunteered to share their status updates as part of the My Personality application, where they also took a variety of questionnaires
- Authors found found striking variations in language with personality, gender, and age
    - Use of an open vocabulary approach

- Results confirmed previously known social science findings, suggested new hypotheses, and showed sustained face validity

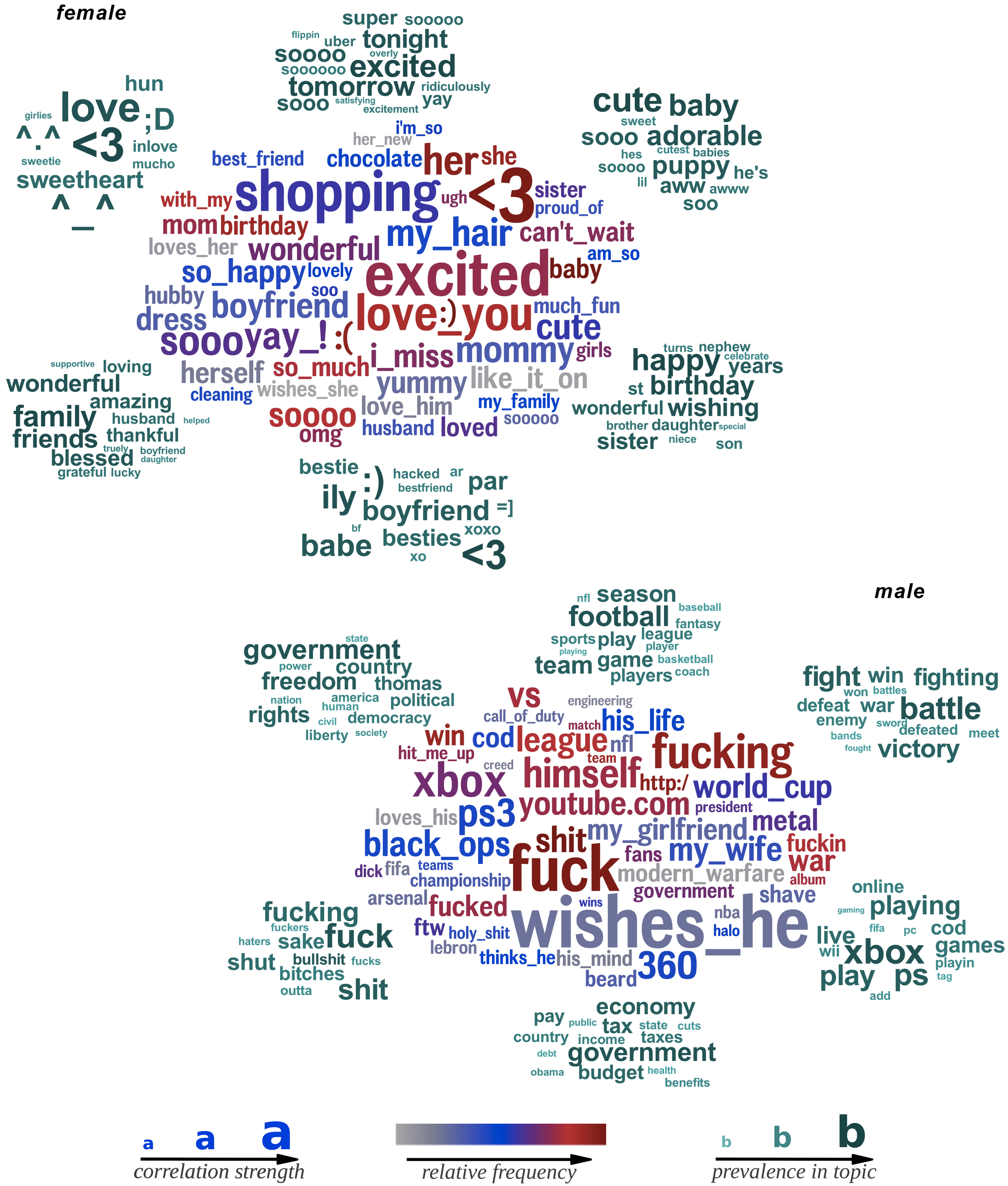
Summary


Volunteer Data
social media messages
gender personality
location
age health
...
a) words and phrases
b) topics
1) Linguistic feature extraction
2) Correlation analysis
3) Visualization
thanksgiving our psalm amazing
praise
:) love day everyone excited for
christmas wonderful lord family
had today
tomorrow great blessed beautiful
friends happy awesome god's church
thank prayers
thankful fun pray good weekend
tonight

Summary

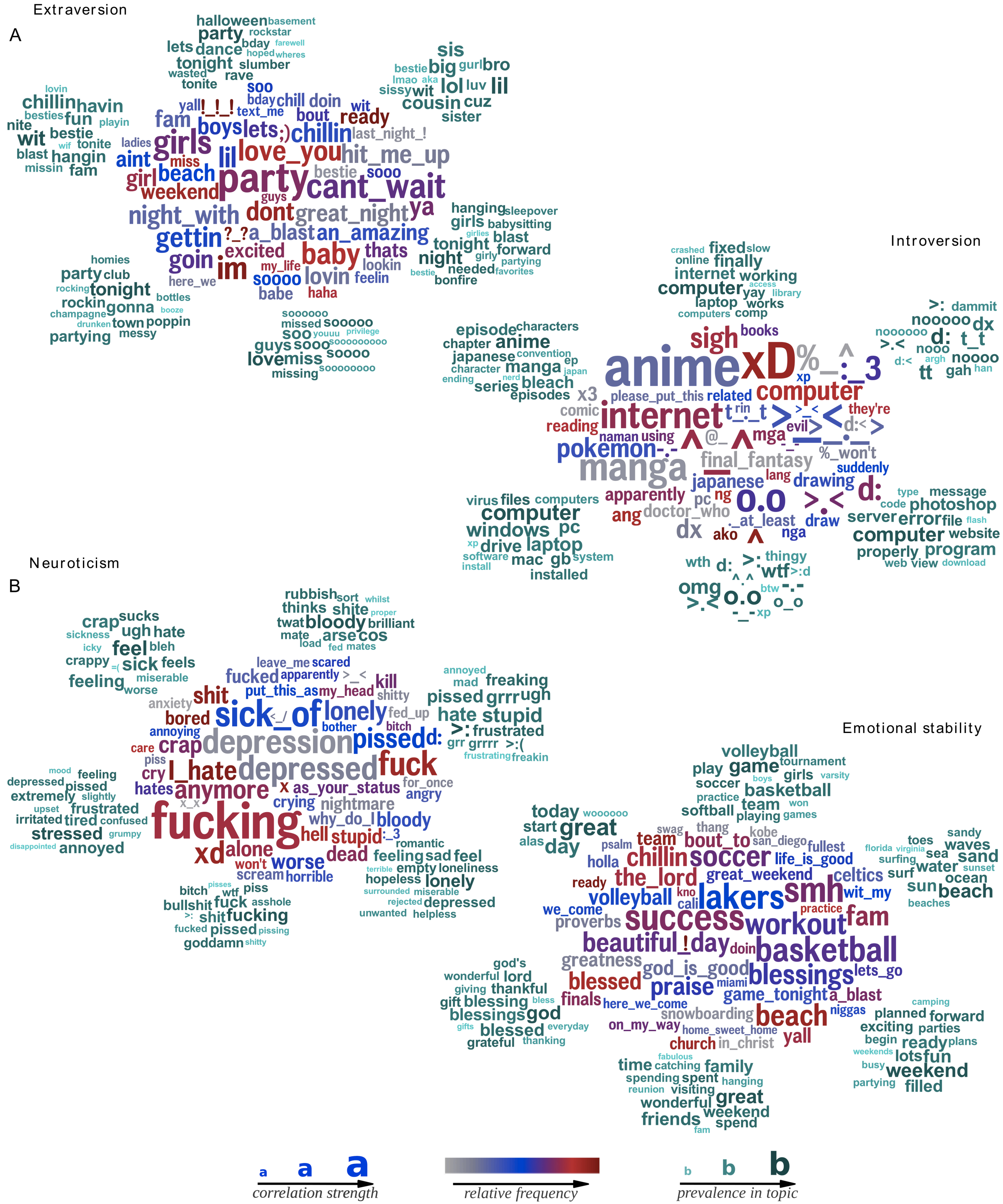
- Open Vocabulary: Differential Language Analysis
- Key characteristics:
  - Open-vocabulary – it is not limited to predefined word lists. Rather, linguistic features including words, phrases, and topics (sets of semantically related words) are automatically determined from the texts. (I.e., it is "data-driven".) This means DLA is classified as a type of open-vocabulary approach.
  - Discriminating – it finds key linguistic features that distinguish psychological and demographic attributes, using stringent significance tests.
  - Simple – it uses simple, fast, and readily accepted statistical techniques.

Summary



female
super sooooo
flippin uber tonight
soooo overly
sooooooo excited
tomorrow ridiculously
sooo satisfying yay
excitement
cute baby
sweet
sooo adorable
hes cutest babies
soooo puppy he's
lil aww awww
soo
hun
girlies love ;D
^.^ <3 inlove
sweetie mucho
sweetheart
^_^
her_new
i'm_so
best_friend chocolate her she
with_my shopping ugh <3 sister
proud_of
mom birthday
my_hair can't_wait
loves_her wonderful am_so
so_happy lovely excited baby
hubby soo
boyfriend love_you much_fun
dress :)
cute
sooo yay_! :(
i_miss mommy girls
herself so_much
yummy like_it_on
cleaning wishes_she
soooo love_him my_family
omg husband loved sooooo
supportive loving
wonderful
amazing
family husband helped
friends thankful
truely boyfriend
blessed daughter
grateful lucky
turns nephew
happy celebrate years
st birthday
wonderful wishing
brother daughter special
sister niece son
bestie
:) hacked ar
bestfriend par
ily
boyfriend =]
bf xoxo
babe besties
xo <3
male
nfl season
baseball
football fantasy
sports play league
playing player
team game basketball
players coach
state
government
power country
freedom thomas
nation america political
human
rights democracy
civil society
liberty
vs engineering
call_of_duty
match his_life
win cod league nfl
hit_me_up team
creed fucking
xbox himself
http:/ world_cup
youtube.com president
loves_his ps3
metal
my_girlfriend
black_ops shit
fans my_wife fuckin
dick fifa teams
war
fuck modern_warfare album
championship
arsenal government
wins shave
fucked
nba
wishes_he
ftw holy_shit halo
lebron
thinks_he his_mind
beard 360
fight win fighting
won battles
defeat war battle
enemy sword
defeated meet
bands
fought victory
online
gaming playing
fifa pc cod
live
wii xbox games
playin
play ps
tag
add
fucking
fuckers
haters sake fuck
shut bullshit fucks
bitches
outta shit
economy
pay public tax state cuts
country income taxes
debt government
obama budget health
benefits
a a a
correlation strength
relative frequency
b b b
prevalence in topic

Summary



Extraversion
A
Introversion
Neuroticism
B
Emotional stability
correlation strength
relative frequency
prevalence in topic

Summary

features	Gender *accuracy*	Age *R*	Extraversion *R*	Agreeableness *R*	Conscientious. *R*	Neuroticism *R*	Openness *R*
LIWC	78.4%	.65	.27	.25	.29	.21	.29
Topics	**87.5%**	**.80**	**.32**	**.29**	**.33**	**.28**	**.38**
WordPhrases	**91.4%**	**.83**	**.37**	**.29**	**.34**	**.29**	**.41**
WordPhrases + Topics	**91.9%**	**.84**	**.38**	**.31**	**.35**	**.31**	**.42**
Topics + LIWC	**89.2%**	**.80**	**.33**	**.29**	**.33**	**.28**	**.38**
WordPhrases + LIWC	**91.6%**	**.83**	**.38**	**.30**	**.34**	**.30**	**.41**
WordPhrases + Topics + LIWC	**91.9%**	**.84**	**.38**	**.31**	**.35**	**.31**	**.42**

accuracy: percent predicted correctly (for discrete binary outcomes). *R*: Square-root of the coefficient of determination (for sequential/continuous outcomes). *LIWC*: *A priori* word-categories from Linguistic Inquiry and Word Count. *Topics*: Automatically created *LDA* topic clusters. *WordPhrases*: words and phrases (n-grams of size 1 to 3 passing a collocation filter). Bold indicates significant ($p < .01$) improvement over the baseline set of features (use of *LIWC* alone).

doi:10.1371/journal.pone.0073791.t002

Why is gender and personality inference useful for social computing researchers and professionals?

How is an open vocabulary approach more suitable for social media language data over closed vocabulary ones?

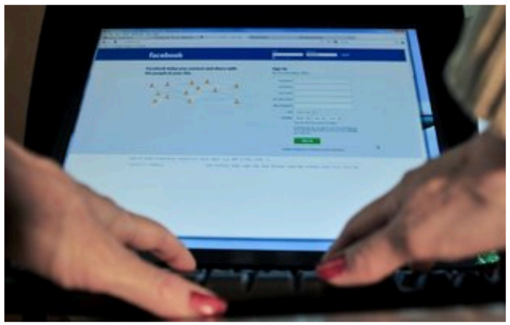
People use social media for all kinds of reasons and purposes. On Facebook in particular, people are heavily concerned about impression management.

Why do you think the assessments of personality are still accurate?

TECHNOLOGY

Facebook Tinkers With Users' Emotions in News Feed Experiment, Stirring Outcry

By VINDU GOEL JUNE 29, 2014



Facebook revealed that it had altered the news feeds of over half a million users in its study.
Karen Bleier/Agence France-Presse — Getty Images

To Facebook, we are all lab rats.

Facebook routinely adjusts its users' news feeds — testing out the number of ads they see or the size of photos that appear — often without their knowledge. It is all for the purpose, the company says, of creating a more alluring and useful product.

But last week, Facebook revealed that it had manipulated the news feeds of over half a million randomly selected users to change the number of positive and negative posts they saw. It was part of a psychological study to examine how emotions can be spread on social media.

The company says users consent to this kind of manipulation when they agree to its terms of service. But in the quick judgment of the Internet, that argument was not universally accepted.

"I wonder if Facebook KILLED anyone with their emotion manipulation stunt. At their scale and with depressed people out there, it's possible," the privacy activist Lauren Weinstein wrote in a Twitter post.

On Sunday afternoon, the Facebook researcher who led the study, Adam D. I. Kramer, posted a public apology on his Facebook page.