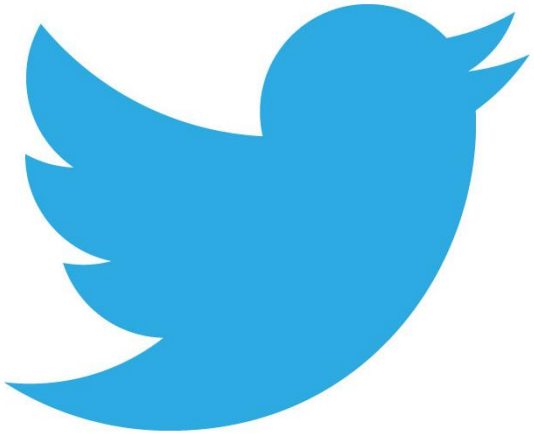
Personality, Gender and Age in the Language of Social Media

Crowdsourcing and Human Computation
12/4/2013

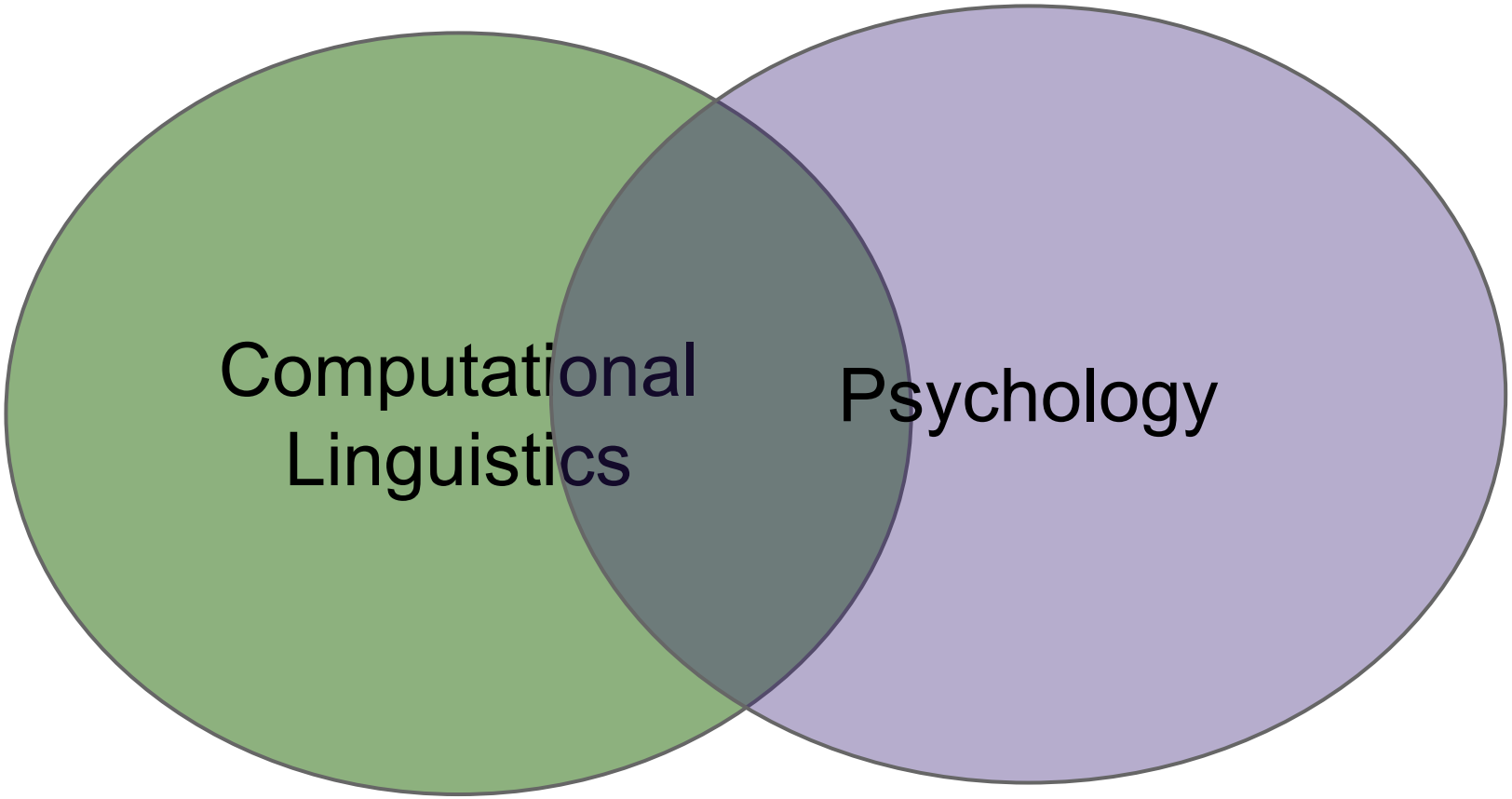
H. Andrew Schwartz

Core Collaborators:
Johannes Eichstaedt, Margaret Kern, Gregory Park, Martin Seligman Lyle Ungar



> 150 million active monthly users
> 350 million messages a day

> 1 billion active monthly users
> 4 billion messages a day


Computational
Linguistics
Psychology

Computational Linguistics traditionally is focused on modeling and understanding language.

- **Results: accurate predictions**

Computational linguistics is traditionally focused on modeling and understanding language.

- **Results: accurate predictions**

Our Goal: Understand the people behind the Language.

- **Results: Novel Insights**

Computational linguistics is traditionally focused on modeling and understanding language.

- **Results: accurate predictions**

Our Goal: Understand the people behind the Language.

- **Results: Novel Insights**

Passive crowdsourcing the human condition?

Psychological language analyses typically limited to apriori language lexica

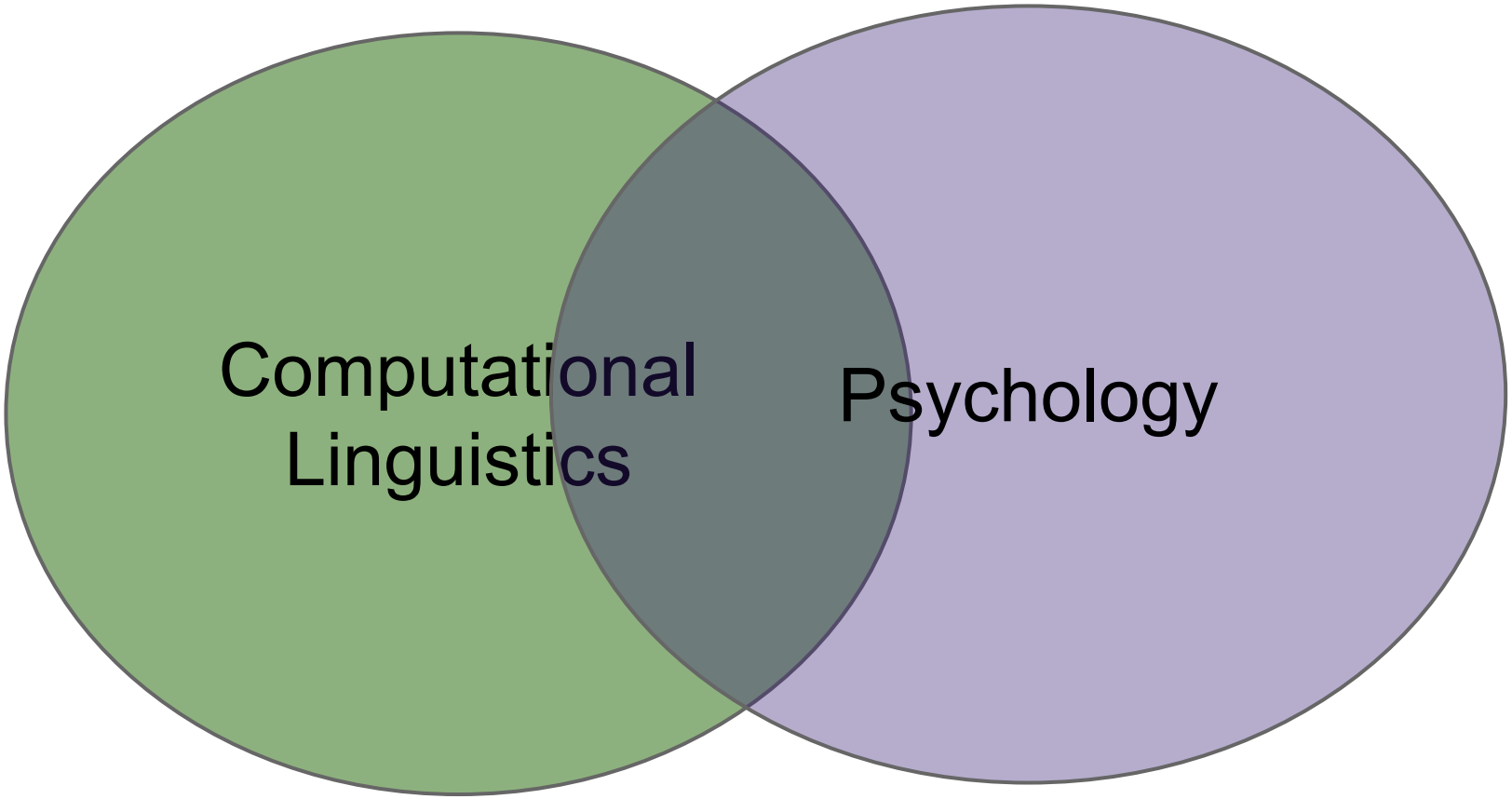
LIWC	Gender		Age		Extraversion		Agreeableness		Conscientious.		Neuroticism		Openness	
Category	[34] *d*	our *β*	[30] *β*	our *β*	[27] *ρ*	our *β*	[27] *ρ*	our *β*	[27] *ρ*	our *β*	[27] *ρ*	our *β*	[27] *ρ*	our *β*
Total function words	-	-0.04	-	0.16	-	-0.04	-	0.02	-	0.02	-	0.03	-	0.09
Total pronouns	0.36	0.07	-	-0.02	*ns*	*ns*	0.11	*ns*	*ns*	-0.03	*ns*	0.04	-0.21	0.07
Personal pronouns	-	0.14	-	-0.08	-	*ns*	-	*ns*	-	-0.04	-	0.04	-	0.05
1st pers singular	0.17	0.13	-0.14	-0.22	*ns*	*ns*	*ns*	-0.03	*ns*	-0.06	0.12	0.05	-0.16	0.05
1st pers plural	*ns*	*ns*	-0.13	0.21	0.11	0.03	0.18	0.05	*ns*	0.05	*ns*	-0.04	-0.1	*ns*
2nd person	-0.06	0.05	-	0.04	0.16	*ns*	*ns*	0.02	*ns*	*ns*	-0.15	*ns*	-0.12	0.02
3rd pers singular	-	0.09	-	0.15	-	*ns*	-	*ns*	-	*ns*	-	0.02	-	*ns*
3rd pers plural	-	-0.05	-	0.26	-	-0.06	-	-0.04	-	*ns*	-	0.02	-	0.03
3rd pers overall	0.2	-	-	-	*ns*	-	*ns*	-	*ns*	-	*ns*	-	*ns*	-
Impersonal pronouns	-	-0.09	-	0.11	-	-0.05	-	*ns*	-	*ns*	-	0.02	-	0.08
Articles	-0.24	-0.24	-	0.28	*ns*	-0.05	*ns*	*ns*	0.09	0.02	-0.11	-0.02	0.2	0.13
Common verbs	-	0.04	-	0.02	-	-0.03	-	*ns*	-	*ns*	-	0.04	-	0.03
Auxiliary verbs	-	0.02	-	0.08	-	-0.06	-	*ns*	-	*ns*	-	0.05	-	0.07
Past tense	0.12	-0.03	-0.16	*ns*	*ns*	-0.04	0.1	0.02	*ns*	-0.02	*ns*	*ns*	-0.16	*ns*
Present tense	0.18	0.08	0.04	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	0.04	-0.16	0.03
Future tense	*ns*	-0.07	0.14	0.09	*ns*	-0.05	*ns*	*ns*	*ns*	*ns*	*ns*	0.03	*ns*	0.05
Adverbs	-	0.05	-	-0.07	-	-0.04	-	*ns*	-	*ns*	-	0.05	-	0.04
Prepositions	-0.17	-0.13	-	0.27	*ns*	-0.04	*ns*	0.03	*ns*	0.06	*ns*	*ns*	0.17	0.06
Conjunctions	-	0.03	-	0.12	-	-0.02	-	0.02	-	0.02	-	0.02	-	0.06
Negations	0.11	*ns*	-	-0.12	*ns*	-0.06	*ns*	-0.05	-0.17	-0.03	0.11	0.07	-0.13	0.02
Quantifiers	-	-0.09	-	0.24	-	-0.02	-	0.03	-	0.05	-	*ns*	-	0.05
Numbers	-0.15	-0.13	-	0.05	-0.12	-0.06	0.11	0.02	*ns*	0.02	*ns*	*ns*	-0.08	0.06
Swear words	-0.22	-0.21	-	-0.17	*ns*	*ns*	-0.21	-0.15	-0.14	-0.09	0.11	0.06	*ns*	*ns*
Social processes	-	0.08	-0.13	0.21	0.15	0.04	0.13	0.02	*ns*	*ns*	*ns*	*ns*	-0.14	*ns*
Family	0.12	0.22	-	0.28	0.09	0.03	0.19	0.03	*ns*	0.03	*ns*	*ns*	-0.17	-0.12
Friends	0.09	0.08	-	0.26	0.15	0.05	0.11	0.04	*ns*	0.02	-0.08	*ns*	*ns*	-0.04
Humans	*ns*	0.04	-	0.06	0.13	0.06	*ns*	-0.05	-0.12	*ns*	*ns*	*ns*	-0.09	*ns*
Affective processes	0.11	0.11	-	-0.05	0.09	0.07	*ns*	0.02	*ns*	*ns*	*ns*	*ns*	-0.12	-0.04
Positive emotion	*ns*	0.21	0.12	0.14	0.1	0.13	0.18	0.13	*ns*	0.1	*ns*	-0.08	-0.15	-0.07
Negative emotion	0.1	-0.12	-0.05	-0.31	*ns*	-0.07	-0.15	-0.17	-0.18	-0.13	0.16	0.15	*ns*	0.03
Anxiety	0.16	0.08	-	-0.13	*ns*	-0.04	*ns*	-0.02	*ns*	-0.02	0.17	0.06	*ns*	0.07
Anger	*ns*	-0.22	-	-0.25	*ns*	-0.05	-0.23	-0.19	-0.19	-0.12	0.13	0.11	*ns*	0.02
Sadness	0.1	0.08	-	-0.15	*ns*	-0.04	*ns*	-0.02	-0.11	-0.04	0.1	0.09	*ns*	*ns*
Cognitive processes	0.07	-0.03	0.07	0.1	*ns*	-0.05	*ns*	0.02	-0.11	*ns*	0.13	0.04	-0.09	0.1
Insight	0.09	-0.05	0.11	0.04	*ns*	-0.09	*ns*	*ns*	*ns*	-0.02	*ns*	0.05	*ns*	0.13
Causation	*ns*	-0.05	*ns*	-0.01	-0.09	-0.06	-0.11	-0.02	-0.12	*ns*	0.11	0.02	*ns*	0.08
Discrepancy	0.07	*ns*	-	0.02	*ns*	-0.05	*ns*	-0.02	-0.13	-0.03	0.13	0.07	-0.12	0.02
Tentative	*ns*	-0.12	-	0.07	-0.11	-0.08	*ns*	*ns*	-0.1	-0.03	0.12	0.06	*ns*	0.07
Certainty	0.14	*ns*	-	0.09	0.1	*ns*	*ns*	0.03	-0.1	0.04	0.13	*ns*	*ns*	0.06
Inhibition	-	0.03	-	0.09	-0.13	*ns*	*ns*	*ns*	*ns*	0.04	0.09	*ns*	*ns*	*ns*
Inclusive	*ns*	0.04	-	0.23	0.09	0.04	0.18	0.05	*ns*	0.05	*ns*	-0.02	0.11	0.06
Exclusive	*ns*	-0.05	*ns*	*ns*	*ns*	-0.07	*ns*	*ns*	-0.16	-0.03	0.1	0.05	*ns*	0.05
Perceptual Processes	0.12	*ns*	-	-0.06	0.09	-0.04	*ns*	*ns*	-0.1	-0.07	*ns*	0.03	-0.11	0.1
See	*ns*	*ns*	-	*ns*	*ns*	-0.02	0.09	*ns*	*ns*	-0.04	*ns*	*ns*	*ns*	0.04
Hear	0.1	-0.07	-	-0.1	0.12	-0.04	*ns*	*ns*	-0.12	-0.06	*ns*	0.02	-0.08	0.08
Feel	0.17	0.04	-	-0.07	*ns*	-0.02	0.1	*ns*	*ns*	-0.04	0.1	0.03	*ns*	0.05
Biological processes	*ns*	0.05	-	-0.06	0.14	0.04	0.09	-0.06	*ns*	-0.06	*ns*	0.05	-0.09	0.02
Body	-	-0.02	-	-0.14	0.1	*ns*	0.09	-0.09	*ns*	-0.09	*ns*	0.06	-0.04	0.04
Health	-	0.05	-	0.07	-	*ns*	-	*ns*	-	*ns*	-	0.06	-	*ns*
Sexual	*ns*	0.05	-	-0.14	0.17	0.1	0.08	-0.04	*ns*	-0.04	*ns*	*ns*	*ns*	*ns*
Ingestion	-	0.02	-	0.12	-	*ns*	-	-0.03	-	-0.03	-	*ns*	-	0.03
Relativity	-	-0.06	-	0.16	-	*ns*	-	0.05	-	0.08	-	-0.03	-	-0.03
Motion	0.07	*ns*	-	0.12	-	0.02	-	0.05	-	0.07	-	-0.04	-	-0.04
Space	*ns*	-0.18	-	0.21	*ns*	*ns*	0.16	*ns*	*ns*	0.02	-0.09	*ns*	-0.11	0.07
Time	*ns*	0.02	-0.19	0.08	*ns*	*ns*	0.12	0.06	0.09	0.09	*ns*	-0.03	-0.22	-0.07
Work	-0.12	-0.08	-	-0.02	-0.08	-0.05	*ns*	0.03	*ns*	0.1	*ns*	-0.03	*ns*	-0.02
Achievement	-	-0.17	-	0.16	-0.09	*ns*	*ns*	0.05	0.14	0.11	*ns*	-0.06	*ns*	-0.02
Leisure	*ns*	-0.08	-	0.03	0.08	0.06	0.15	0.04	*ns*	0.03	*ns*	-0.07	-0.17	*ns*
Home	0.15	0.19	-	0.18	*ns*	*ns*	0.19	0.03	*ns*	0.04	*ns*	-0.02	-0.2	-0.06
Money	-0.1	-0.12	-	0.24	*ns*	*ns*	-0.11	-0.04	*ns*	0.03	*ns*	*ns*	*ns*	0.03
Religion	-	-0.03	-	0.21	0.11	*ns*	*ns*	0.06	*ns*	0.04	*ns*	-0.04	*ns*	*ns*
Death	-	-0.18	-	-0.1	*ns*	-0.08	-0.13	-0.09	-0.12	-0.08	*ns*	0.08	0.15	0.09
Assent	-	0.07	-	-0.22	*ns*	0.05	*ns*	0.04	-0.09	*ns*	*ns*	-0.04	-0.11	-0.05
Nonfluencies	-	-0.03	-	0.02	-	*ns*	-	*ns*	-	*ns*	-	0.03	-	*ns*
Fillers	-	-0.02	-	-0.24	-	*ns*	-	-0.04	-	-0.08	-	0.03	-	0.04
participants (N)	*9,130*	*74,859*	*3,087*	*74,859*	*576*	*72,709*	*576*	*72,772*	*576*	*72,781*	*576*	*71,968*	*576*	*72,809*

Psychological language analyses typically limited to apriori language lexica

- **limited to pre-chosen hypotheses**
- **don’t always measure what is expected**

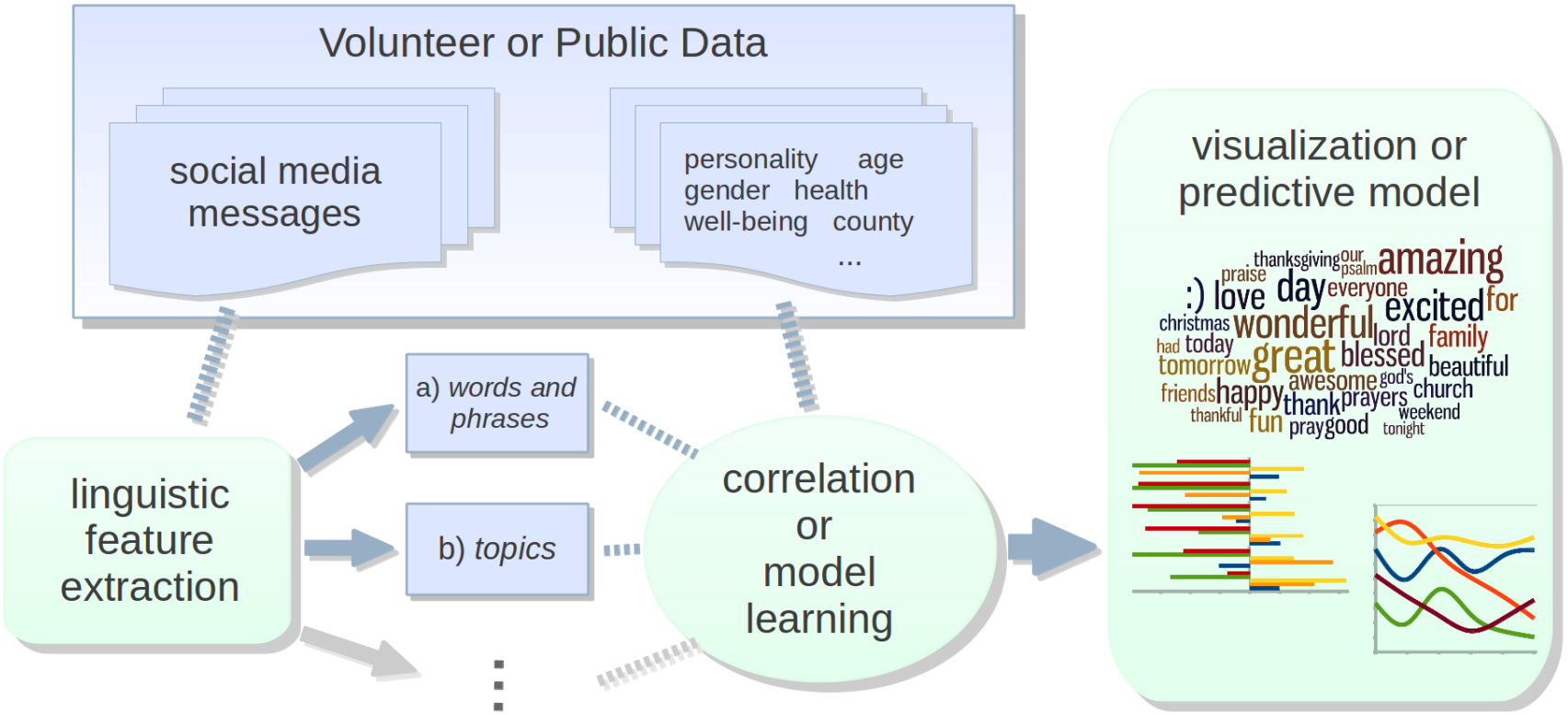
“open vocabulary”: let the data dictate the words, phrases, and linguistic features that matter.

- **transparent results**


Computational
Linguistics
Psychology

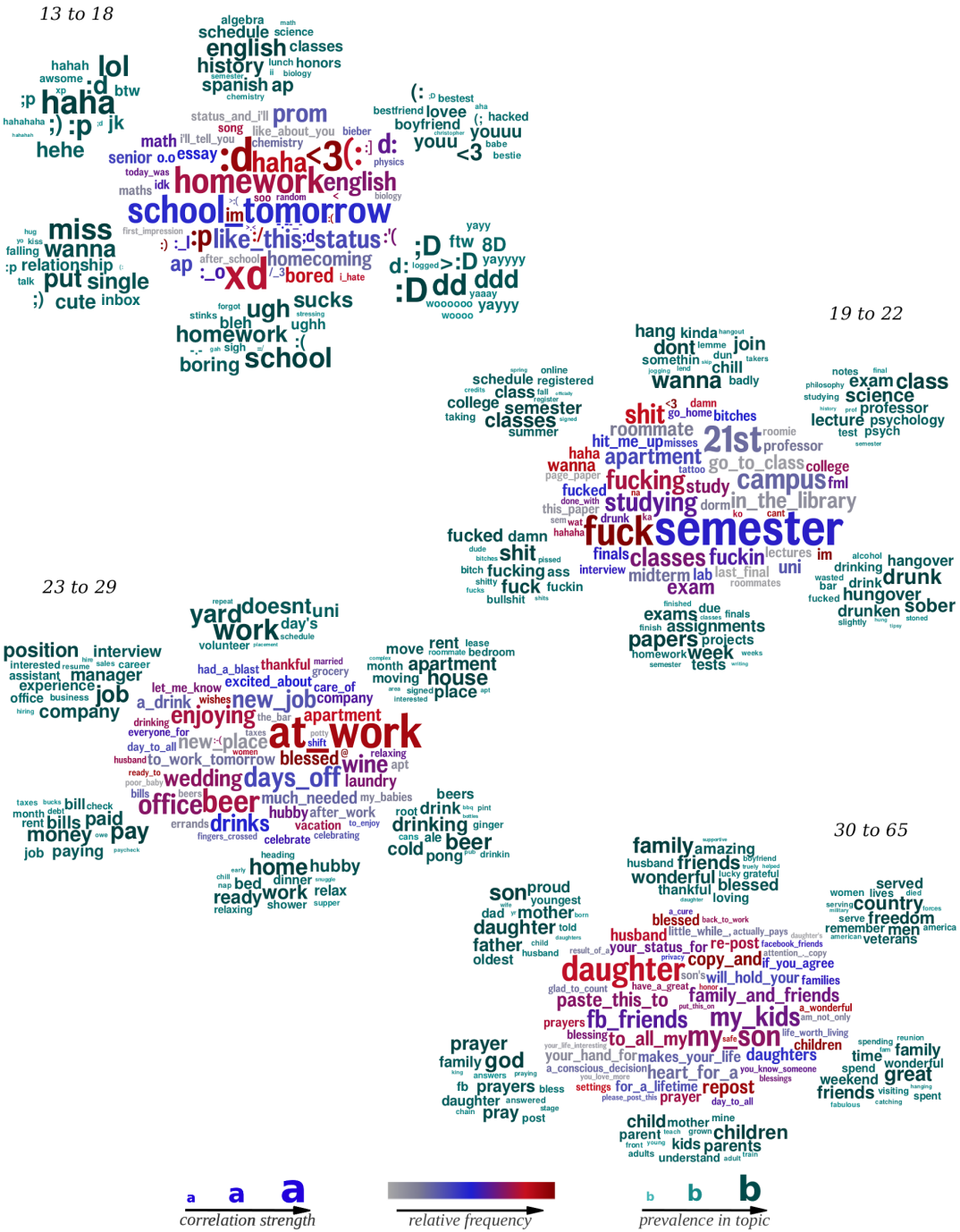
Method

research mostly falls into this framework:

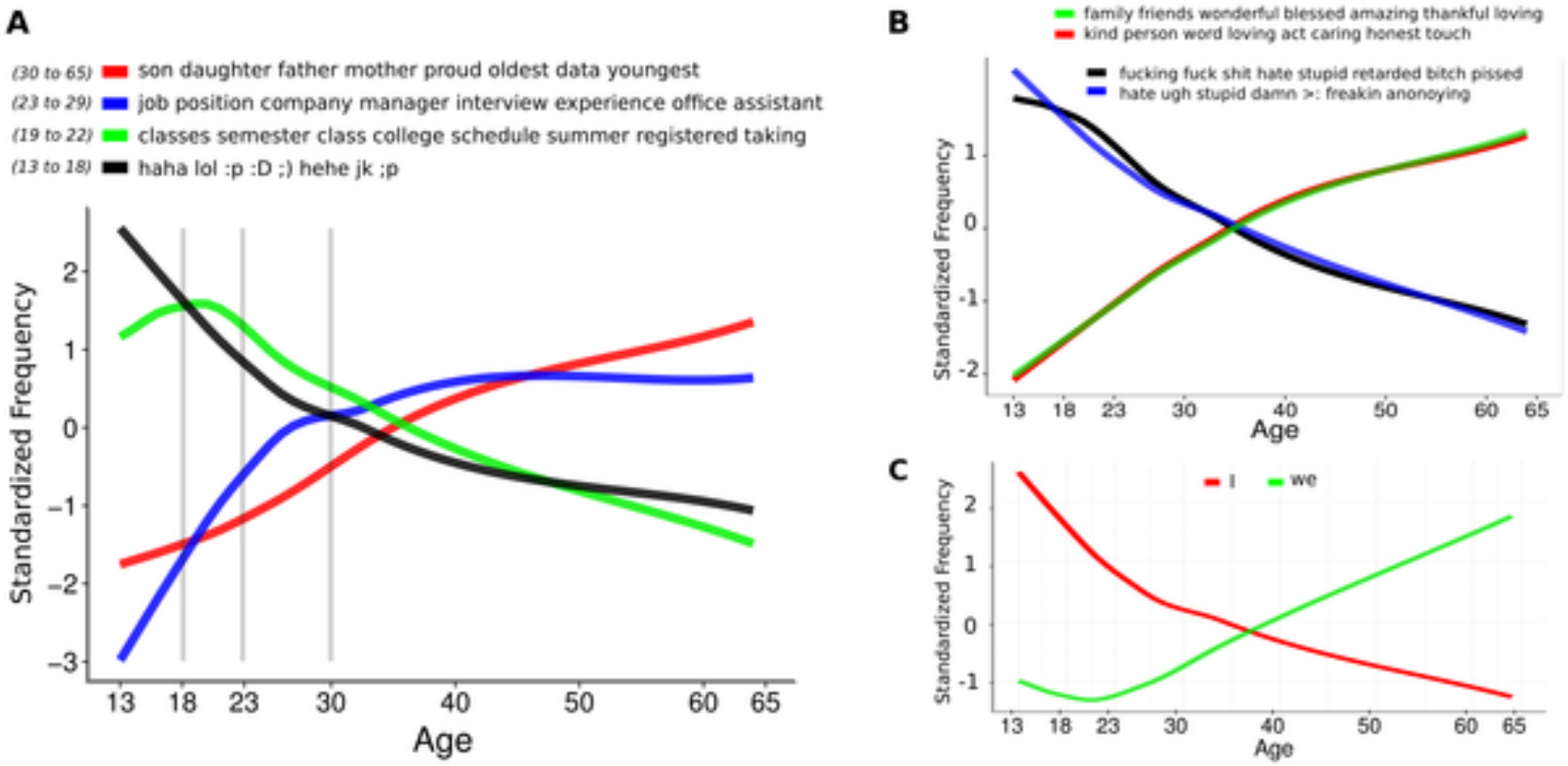


Method

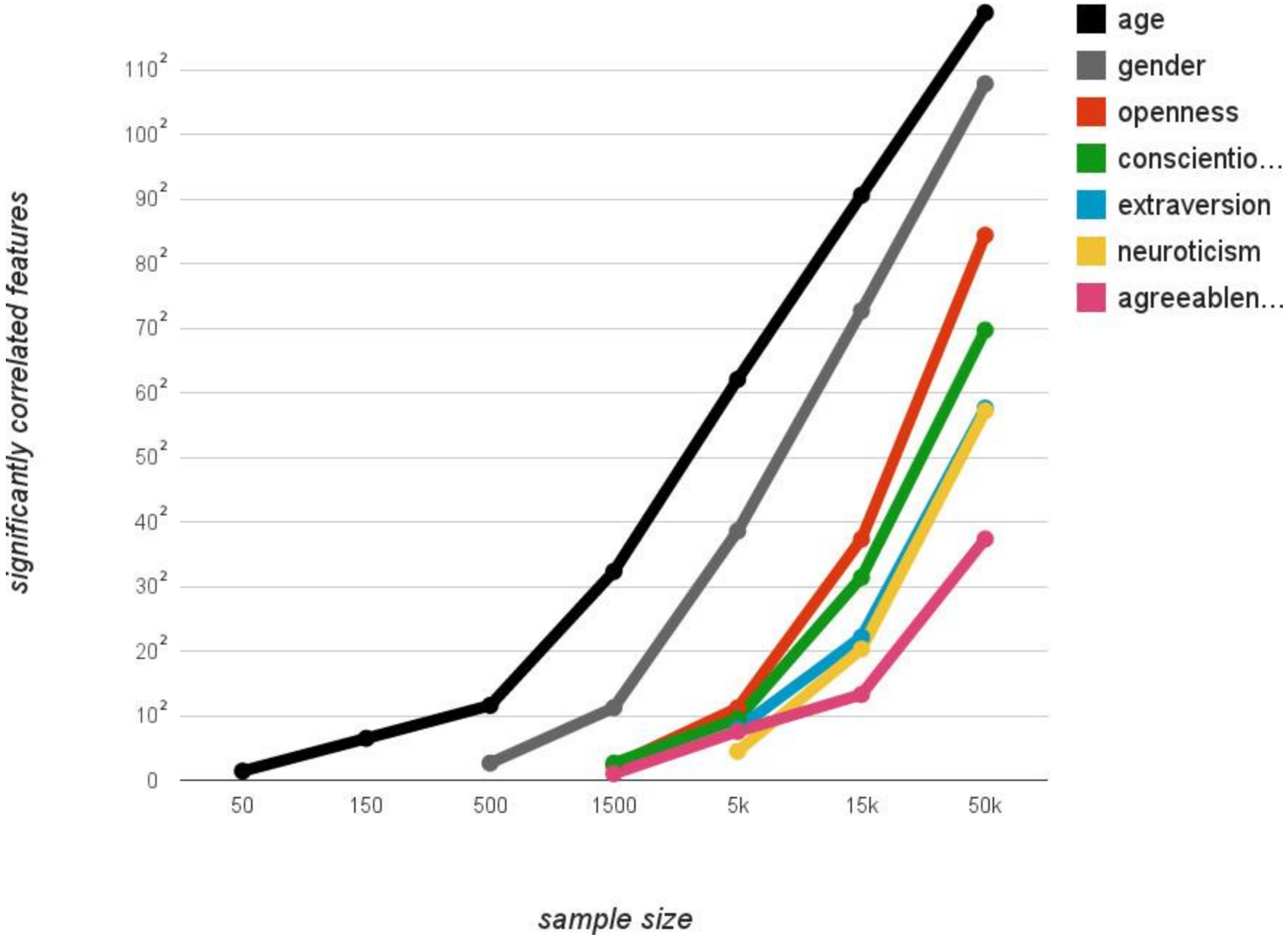
http://prezi.com/rvpue6ucpci1/?utm_campaign=share&utm_medium=copy&rc=ex0share

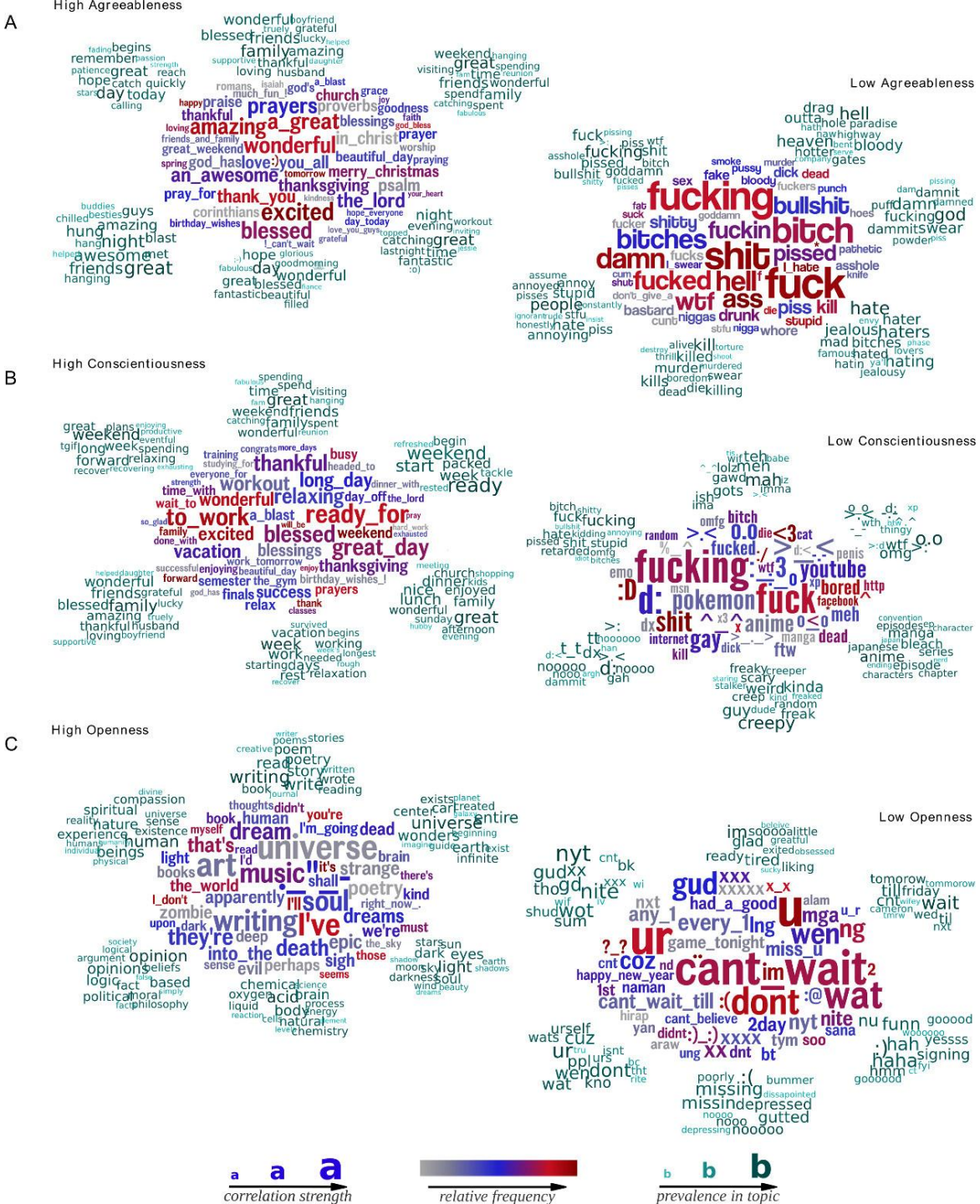

13 to 18
algebra
schedule
science
english
classes
history
lunch
honors
spanish
ap
chemistry
hahah
lol
awsome
:d
btw
;p
haha
hahahaha
;) :p
jk
hehe
status_and_i'll
prom
song
like_about_you
bieber
math
i'll_tell_you
chemistry
senior
o.o
essay
:d
haha
<3(:
d:
today_was
maths
idk
homework
english
soo
random
school
im
tomorrow
miss
wanna
falling
relationship
talk
put
single
cute
inbox
like_this
status
ap
homecoming
xd
bored
i_hate
ugh
sucks
stinks
bleh
ughh
homework
:(
boring
sigh
school
bestfriend
lovee
boyfriend
youuu
<3
babe
bestie
physics
biology
(;
hacked
yayy
;D
ftw
8D
d:
>:D
yayyyy
:D
dd
ddd
yaaay
yayyy
wooooo
19 to 22
hang
kinda
dont
join
somethin
chill
wanna
badly
schedule
registered
class
college
semester
classes
summer
shit
<3
damn
go_home
bitches
roommate
21st
professor
hit_me_up
misses
haha
apartment
wanna
page_paper
tattoo
go_to_class
college
fucked
fucking
study
campus
fml
studying
dorm
in_the_library
this_paper
drunk
cant
hahaha
fucksemester
fucked
damn
shit
finals
classes
fuckin
uni
midterm
lab
last_final
roommates
interview
bitch
fucking
ass
shitty
fuck
fuckin
bullshit
exam
notes
exam
class
philosophy
science
studying
professor
lecture
psychology
test
psych
lectures
im
alcohol
drinking
hangover
drunk
wasted
bar
drink
fucked
hungover
drunken
sober
slightly
23 to 29
yard
doesnt
uni
work
day's
volunteer
position
interview
interested
career
assistant
manager
experience
job
office
company
had_a_blast
thankful
grocery
excited_about
care_of
let_me_know
a_drink
wishes
new_job
company
enjoying
the_bar
apartment
drinking
everyone_for
taxes
day_to_all
new_place
at
work
husband
to_work_tomorrow
blessed
@
relaxing
wine
apt
ready_to
poor_baby
wedding
days_off
laundry
bills
beers
much_needed
my_babies
office
beer
hubby
after_work
errands
drinks
vacation
to_enjoy
fingers_crossed
celebrate
celebrating
taxes
bucks
bill
check
month
debt
bills
paid
rent
money
owe
pay
job
paying
paycheck
move
rent
lease
roommate
bedroom
month
apartment
moving
house
area
signed
interested
place
apt
beers
root
drink
bbq
pint
bottles
drinking
ginger
cans
ale
beer
cold
pong
pub
drinkin
heading
early
home
hubby
chill
bed
dinner
snuggle
nap
ready
work
relax
relaxing
shower
supper
30 to 65
family
amazing
husband
friends
boyfriend
wonderful
lucky
grateful
thankful
blessed
loving
daughter
son
proud
youngest
wife
dad
mother
born
daughter
told
daughters
father
child
husband
oldest
blessed
back_to_work
husband
little_while_
actually_pays
daughter's
your_status_for
re-post
facebook_friends
result_of_a
attention_copy
privacy
daughter
son's
copy_and
if_you_agree
glad_to_count
have_a_great
honor
will_hold_your
families
paste_this_to
family_and_friends
put_this_on
a_wonderful
prayers
fb_friends
my_kids
sm_not_only
prayer
blessing
your_life_interesting
to_all_my
my
son
safe
children
life_worth_living
family
god
your_hand_for
makes_your_life
daughters
time
wonderful
a_conscious_decision
you_know_someone
spend
great
answers
praying
heart_for_a
blessings
you_love_more
settings
for_a_lifetime
repost
weekend
fb
prayers
bless
please_post_this
prayer
day_to_all
friends
daughter
answered
stage
chain
pray
post
fabulous
catching
spent
child
mother
mine
parent
teach
grown
children
front
young
kids
parents
adults
understand
adult
train
served
women
lives
died
serving
military
country
forces
serve
freedom
remember
men
american
america
veterans
a
correlation strength
relative frequency
b
prevalence in topic


female
super
sooooo
tonight
soooo
excited
tomorrow
sooo
yay
hun
love
;D
^_^
<3
inlove
sweetheart
^_^
_
best_friend
chocolate
her
she
i'm_so
her_new
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
dress
boyfriend
love:)you
much_fun
soooyay_!:(
cute
i_miss
mommy
girls
herself
so_much
yummy
like_it_on
cleaning
wishes_she
love_him
my_family
soooo
omg
husband
loved
sooooo
wonderful
amazing
family
husband
friends
thankful
blessed
grateful
lucky
loving
bestie
:)
hacked
par
ily
boyfriend
=]
babe
besties
xoxo
<3
xo
cute
baby
sooo
adorable
soooo
puppy
he's
aww
awww
soo
happy
years
st
birthday
wonderful
wishing
brother
daughter
sister
niece
son
male
season
football
fantasy
sports
play
league
team
game
players
coach
government
country
freedom
thomas
rights
political
democracy
liberty
vs
his_life
win
cod
league
nfl
hit_me_up
fucking
xbox
himself
http:/
world_cup
loves_his
ps3
youtube.com
president
metal
black_ops
shit
my_girlfriend
fans
my_wife
fuckin
war
dick
fifa
fuck
modern_warfare
championship
arsenal
fucked
government
shave
wishes
he
ftw
holy_shit
lebron
thinks_he
his_mind
beard
360
fucking
fuck
sake
shut
bullshit
bitches
shit
outta
fight
win
fighting
defeat
war
battle
enemy
defeated
victory
online
playing
live
xbox
games
play
ps
cod
economy
pay
tax
taxes
country
income
government
budget
benefits
correlation strength
relative frequency
prevalence in topic


A
(30 to 65) son daughter father mother proud oldest data youngest
(23 to 29) job position company manager interview experience office assistant
(19 to 22) classes semester class college schedule summer registered taking
(13 to 18) haha lol :p :D ;) hehe jk ;p
Standardized Frequency
Age
B
family friends wonderful blessed amazing thankful loving
kind person word loving act caring honest touch
fucking fuck shit hate stupid retarded bitch pissed
hate ugh stupid damn >: freakin anonoying
Standardized Frequency
Age
C
I
we
Standardized Frequency
Age


A
Extraversion
halloween party lets dance tonight tonite rave soo chillin havin fun fam boys lets :) chillin girls lil love_you hit_me_up aint girl beach party cant_wait weekend night_with dont great_night ya gettin ?_? a_blast an_amazing goin im excited baby thats sooo lovin babe haha night tonight rockin gonna partying
sis big bro lol lil cousin cuz sister wit ready
Introversion
computer internet sigh books anime xD %_^_:_3 manga pokemon final_fantasy japanese drawing d: >.< 0.0 omg o.o
Neuroticism
B
crap sucks ugh hate feel bleh sick feels feeling
rubbish thinks shite twat bloody arse cos
fucked shit sick_of lonely bored depression pissed d: >:(frustrated crap cry I_hate depressed fuck hates anymore as_your_status crying nightmare fucking why_do_I bloody hell stupid xd alone worse dead stressed annoyed
hate stupid pissed grrr ugh freaking
Emotional stability
volleyball game basketball soccer team today great start day chillin soccer life_is_good the_lord great_weekend volleyball lakers smh success workout fam beautiful !day basketball greatness god_is_good blessings blessed praise beach church
celtics surf sun beach waves sand water ocean
time family spending spent great friends weekend
correlation strength
relative frequency
prevalence in topic


significantly correlated features
110^2
100^2
90^2
80^2
70^2
60^2
50^2
40^2
30^2
20^2
10^2
0
50
150
500
1500
5k
15k
50k
sample size
age
gender
openness
conscientio...
extraversion
neuroticism
agreeablen...

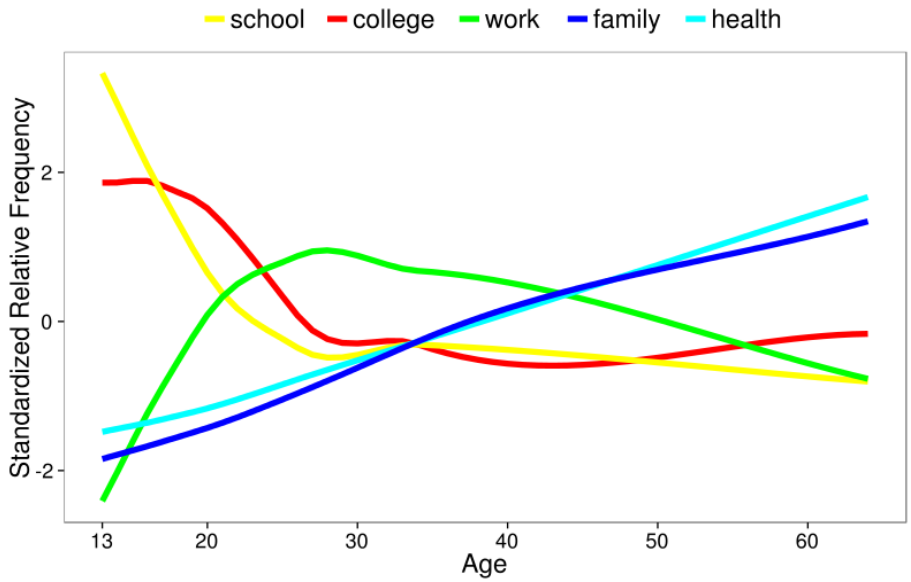

A
High Agreeableness
Low Agreeableness
B
High Conscientiousness
Low Conscientiousness
C
High Openness
Low Openness
correlation strength
relative frequency
prevalence in topic

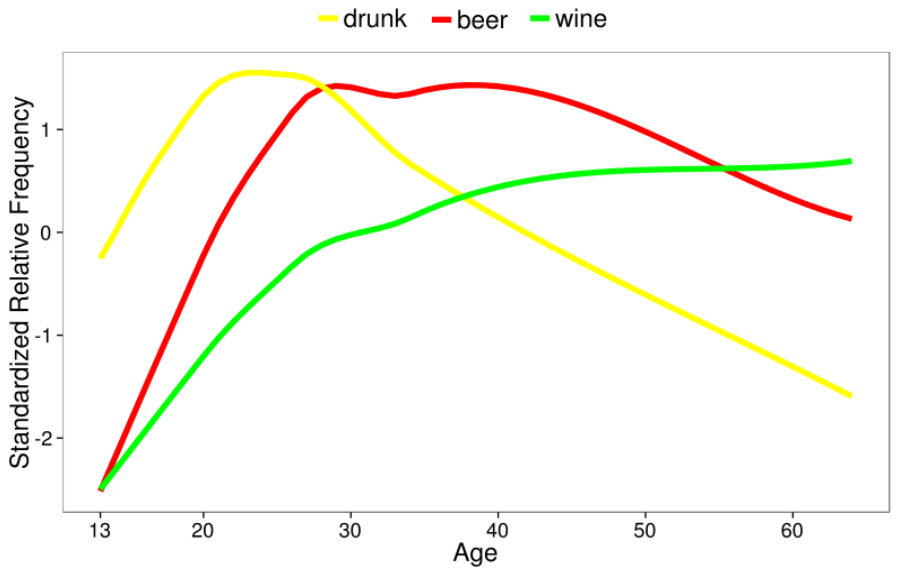
Supporting Table 2. Prediction results when selecting features via differential language analysis.

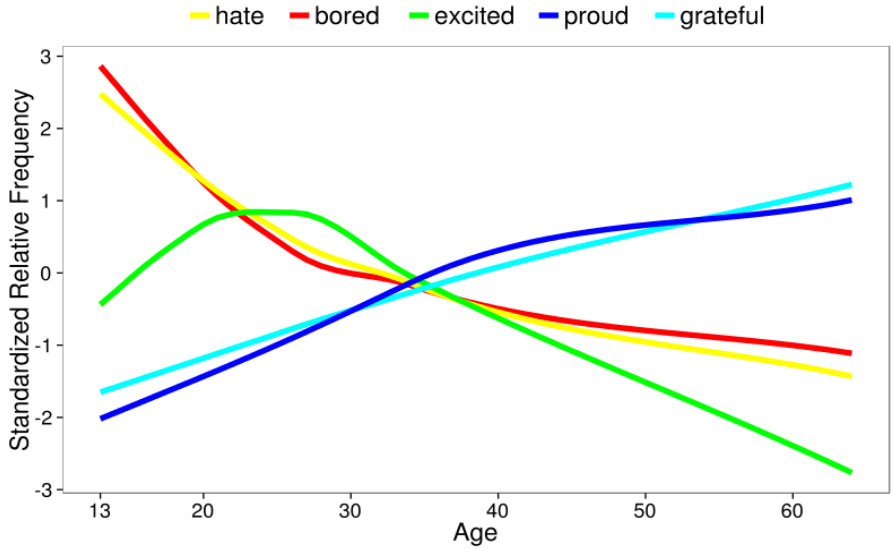
features	Gender *accuracy*	Age *R*	Extraversion *R*	Agreeableness *R*	Conscientious. *R*	Neuroticism *R*	Openness *R*
LIWC	77.7%	.65	.25	.25	.29	.22	.28
Topics	**88.2%**	**.79**	**.34**	**.28**	**.34**	**.28**	**.39**
WordPhrases	**91.8%**	**.81**	**.37**	**.27**	**.34**	**.28**	**.40**
WordPhrases + Topics	**92.0%**	**.82**	**.38**	**.29**	**.35**	**.30**	**.41**
Topics + LIWC	**89.2%**	**.80**	**.35**	**.28**	**.34**	**.28**	**.40**
WordPhrases + LIWC	**91.8%**	**.81**	**.38**	**.28**	**.34**	**.29**	**.40**
WordPhrases + Topics + LIWC	**92.0%**	**.82**	**.38**	**.30**	**.35**	**.30**	**.41**

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). Differential language analysis was run over the training set, and only those features significant at Bonferonni-corrected $p < 0.001$ were included during training and testing. No controls were used so as to be consistent with the evaluation in the main paper, and so one could consider this a univariate feature selection. On average results are just below those of not using *differential language analysis* to select features but there is no significant difference.

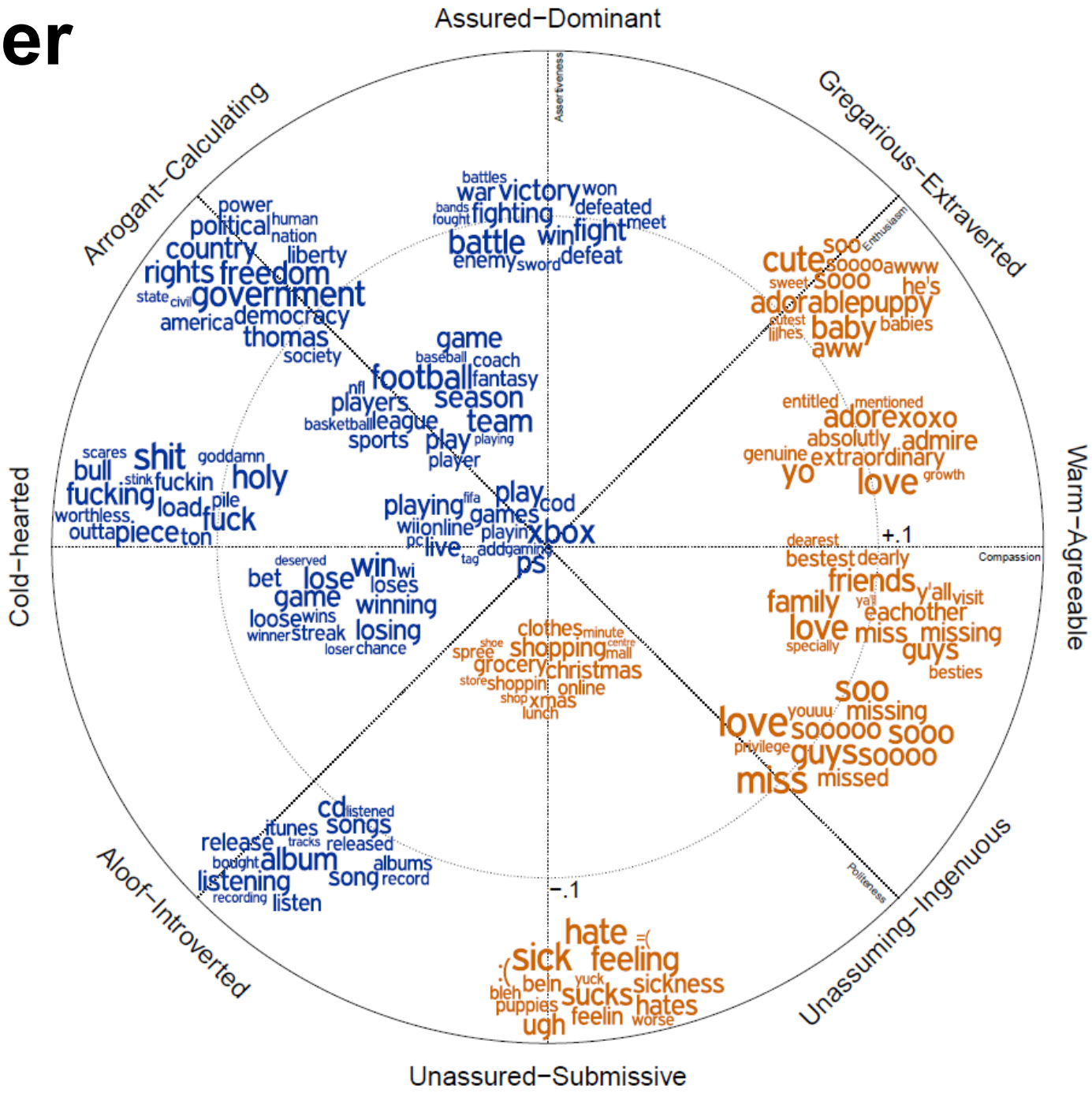
Development



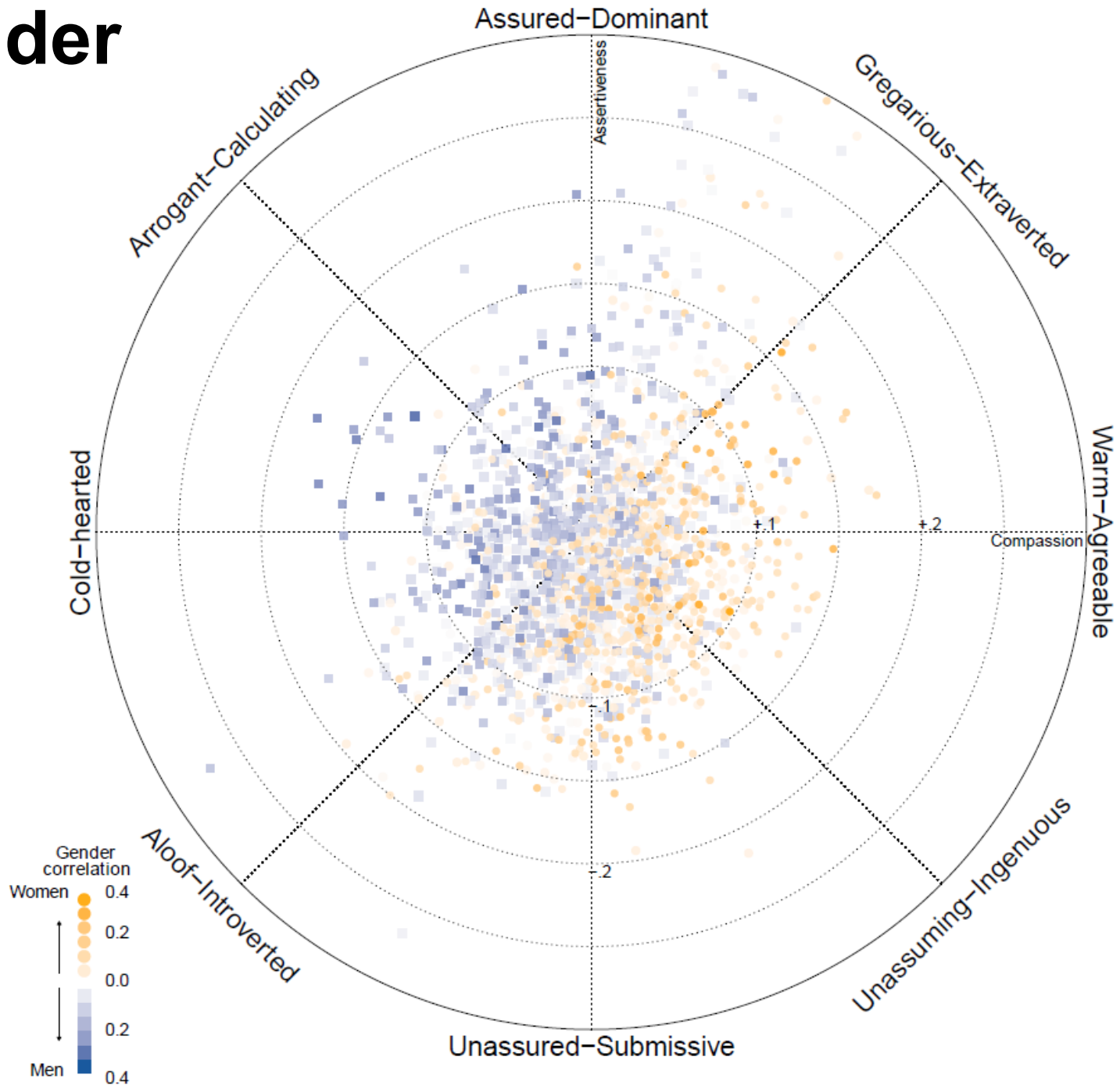




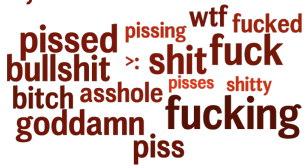
Gender

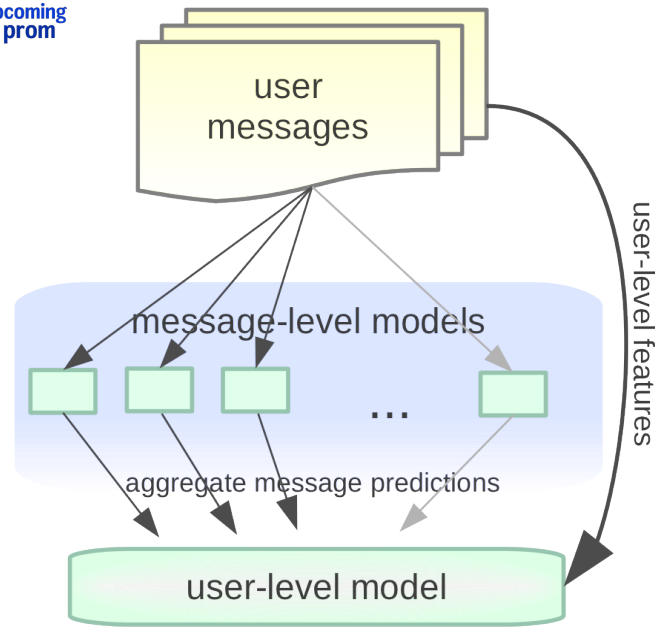

Assured–Dominant
Gregarious–Extraverted
Warm–Agreeable
Unassuming–Ingenuous
Unassured–Submissive
Aloof–Introverted
Cold–hearted
Arrogant–Calculating
Assertiveness
Enthusiasm
Politeness
Compassion
+.1
−.1
power political human nation country liberty rights freedom state civil government democracy america thomas society
battles war victory won bands fought fighting defeated battle win fight meet enemy sword defeat
soo cute sooooo awww sweet sooo he's adorablepuppy cutest lil he's baby babies aww
game baseball coach football fantasy nfl players season basketball league team sports play playing player
entitled mentioned adorexoxo absolutly admire genuine extraordinary yo love growth
scares shit goddamn bull stink fuckin holy fucking load pile worthless fuck outta piece ton
playing fifa play cod games wii online playin xbox pc live tag addgamin ps
dearest bestest dearly friends y'all visit family ya'll eachother love miss missing specially guys besties
deserved bet lose win wi loses game winning loose wins winner streak losing loser chance
clothes minute shoe shopping centre spree mall grocery christmas store shoppin online shop xmas lunch
soo youuu missing love sooooo sooo privilege guyssoooo miss missed
cd listened itunes songs release tracks released album bought albums listening song record recording listen
hate =(sick feeling :(bein yuck sickness bleh sucks hates puppies feelin worse ugh

Gender


Assured–Dominant
Gregarious–Extraverted
Warm–Agreeable
Unassuming–Ingenuous
Unassured–Submissive
Aloof–Introverted
Cold–hearted
Arrogant–Calculating
Assertiveness
Compassion
+.1
+.2
−.1
−.2
Gender correlation
Women
Men
0.4
0.2
0.0
0.2
0.4

Individual Well-Being: message to user-level





baselines	r
(mean)	.000
lexica: GNH	.210
lexica: Hedonometer	.108

message and user-level

Optimism/Pessimism - CAVE

- Uses explanatory style

- Three factors:

  - Internality: whether the cause of the event implicates something about the speaker or is due to situational characteristics

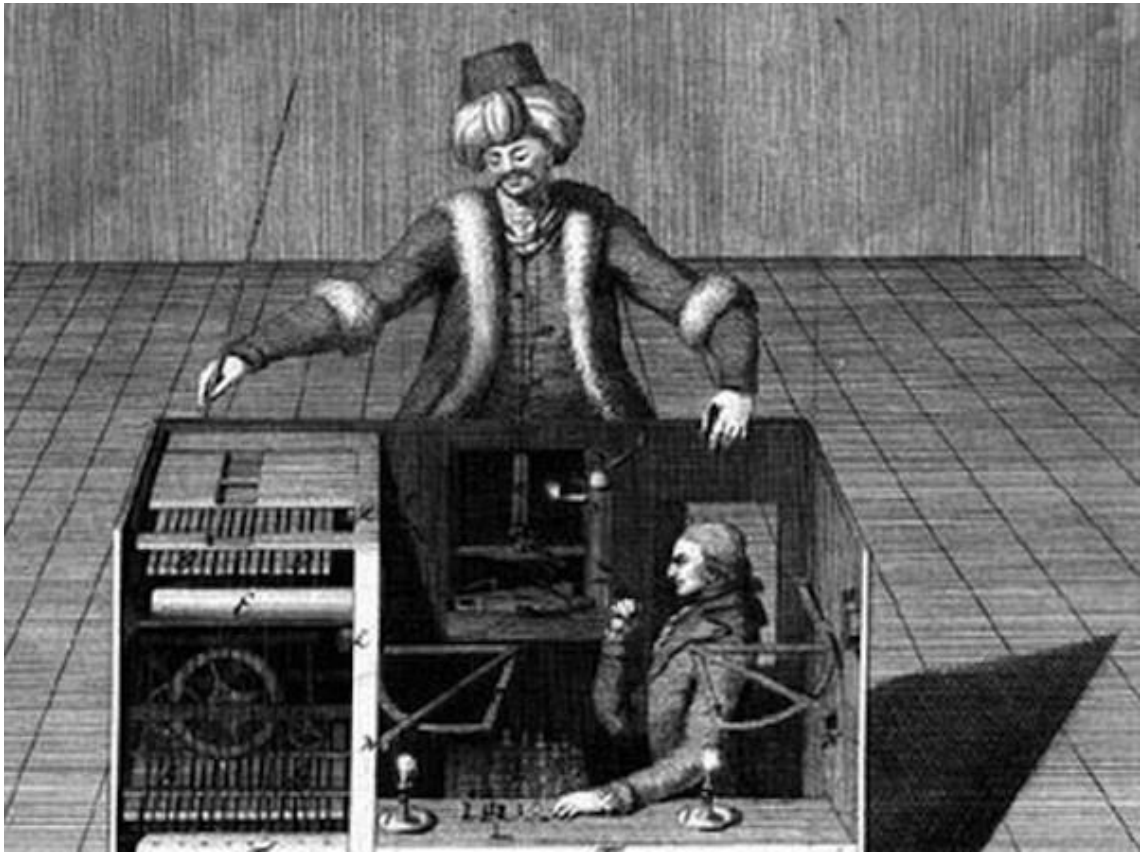
  -Stability: whether the cause of the event persists across time

  -Globality: whether the cause of the event persists across situation (often covaries with stability)

- Positive events: high scores indicate optimism

- Negative events: low scores indicate optimism

Mechanical Turk



Problem: Manually created lists of words are not always used as expected.

Solution: Annotate thousands of samples of real Twitter and Facebook posts for expressions of well-being.

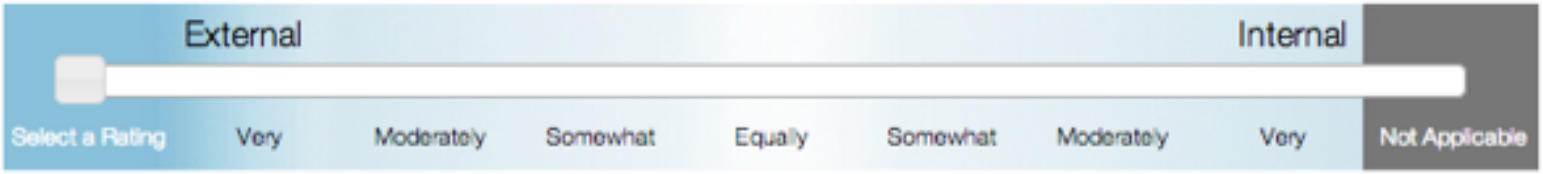
RT @XSTROLOGY: #Leo can not stand you bad-mouthing their lover.<~true its goin be a fight if u talkin sideways bout my boo lol

Does the above message include a causal explanation?

(•) Yes ( ) No

To what extent does the causal explanation of this message indicate **that the cause of the event is internal (e.g. cause of the event placed on speaker)** or **that the cause of the event is external (e.g. cause of the event placed on environment):**

If you cannot assess a degree of internality or externality, rate as 'Not Applicable'.



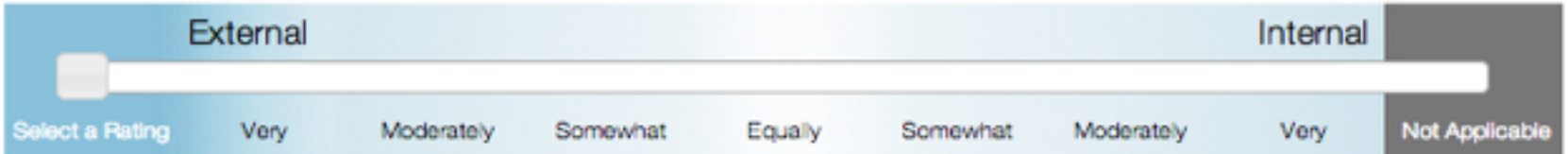
RT @XSTROLOGY: #Leo can not stand you bad-mouthing their lover.<~true its goin be a fight if u talkin sideways bout my boo lol

Does the above message include a causal explanation?

( ) Yes (•) No

To what extent does this message indicate **internality** or **externality:**

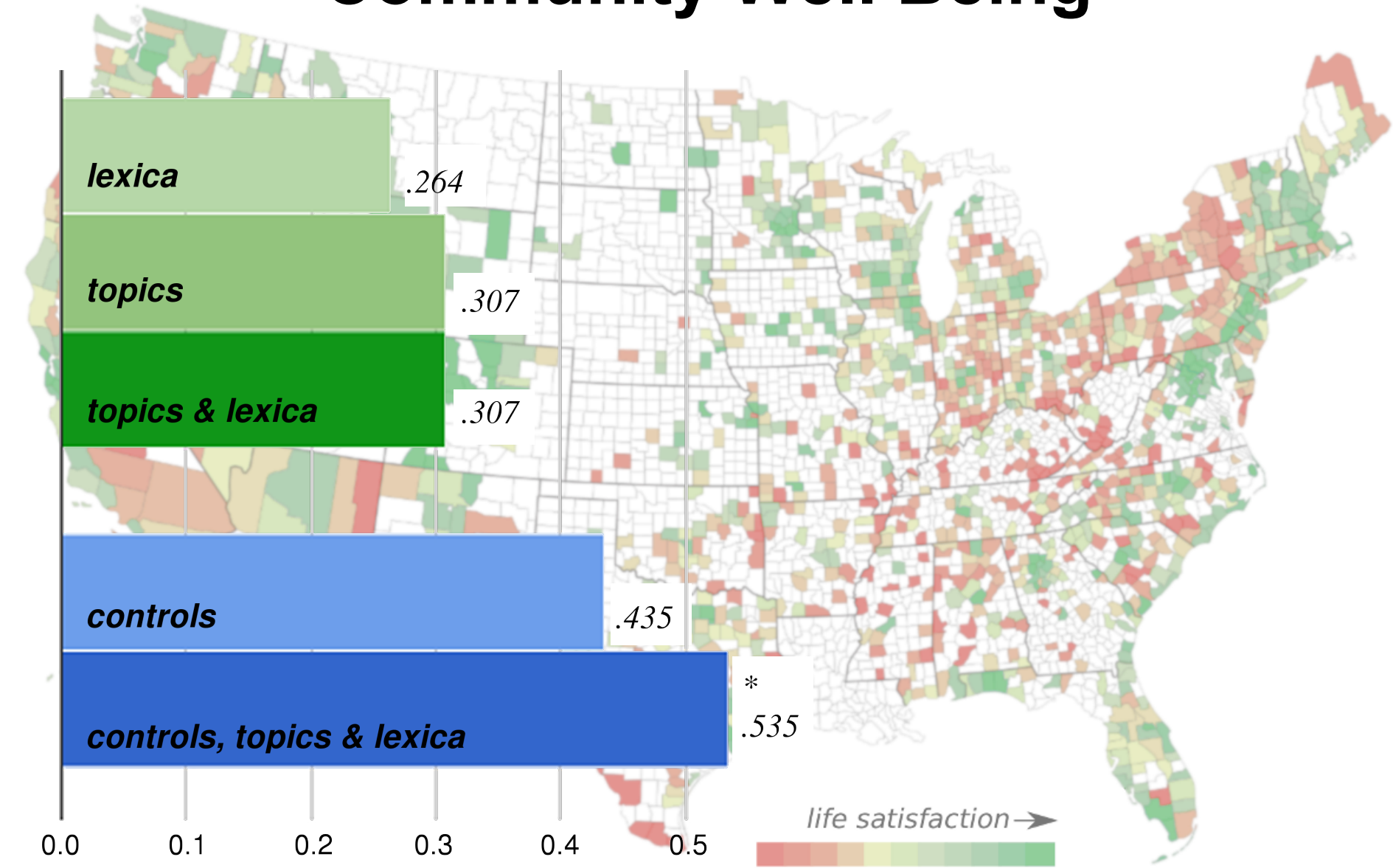
If you cannot assess a degree of internality or externality, rate as 'Not Applicable'.



End of talk

(remaining slides were not presented but might look fun!)

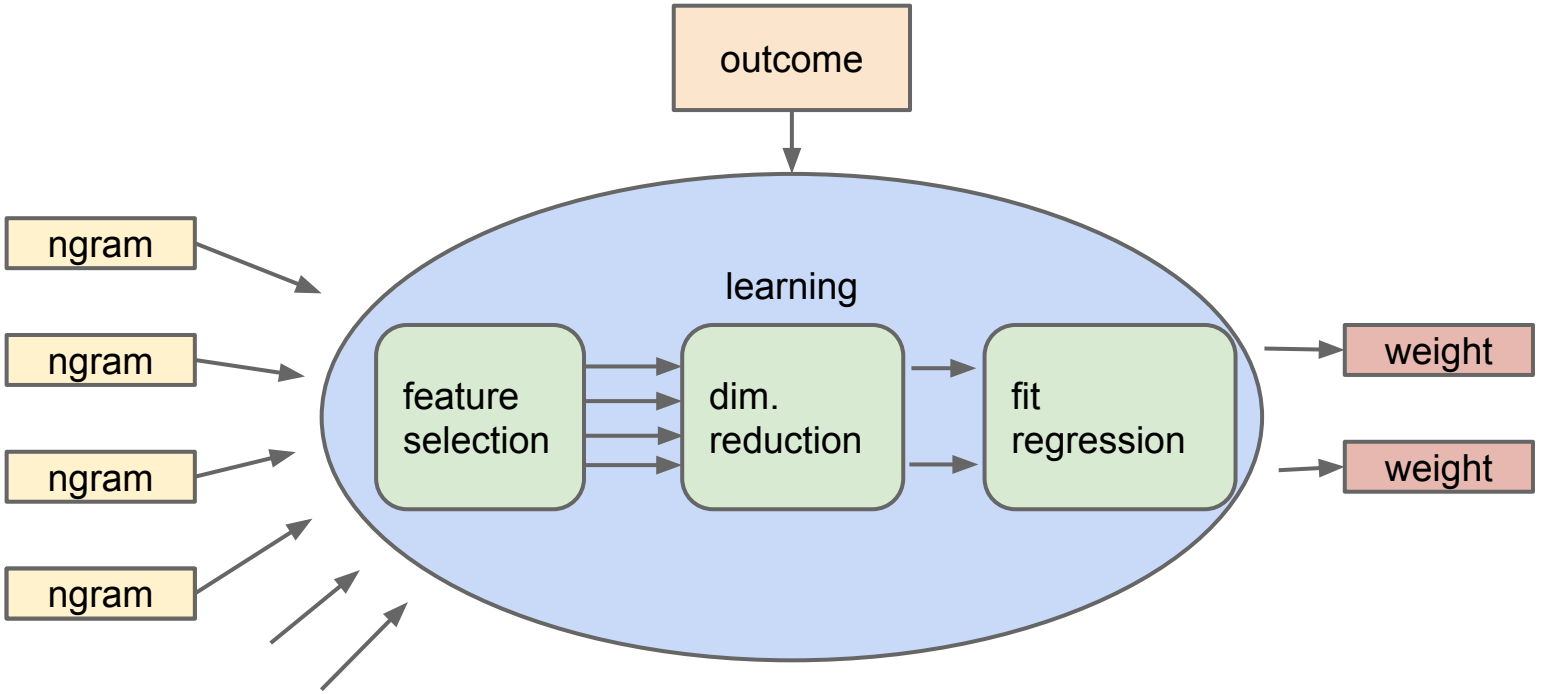
Community Well Being



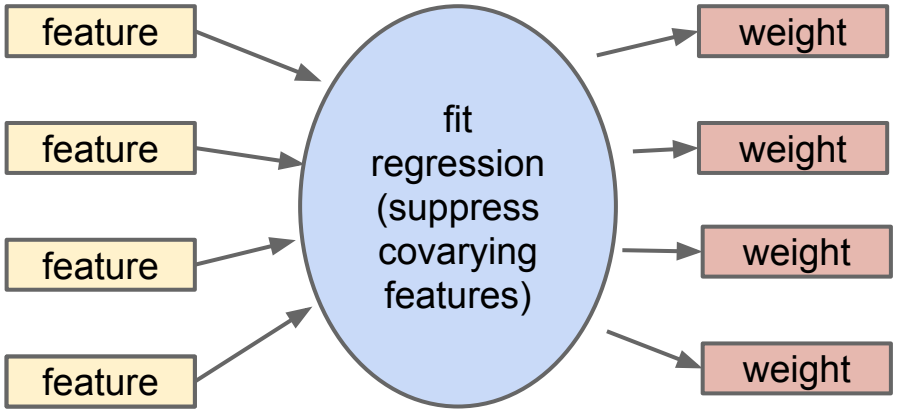
*significant improvement over controls alone

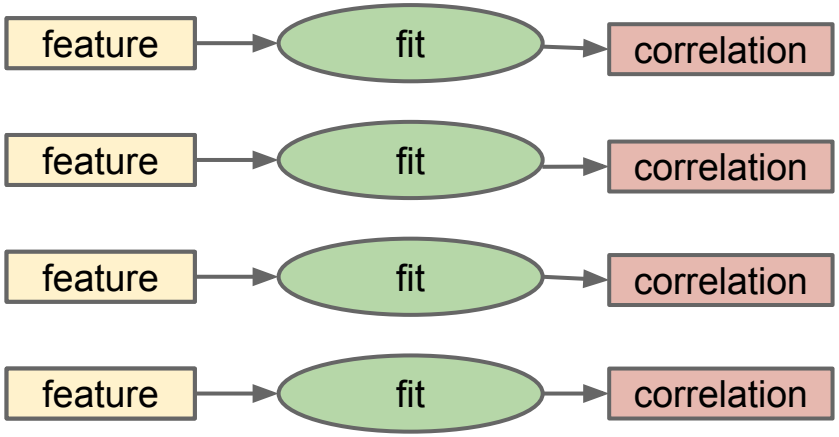
Generating Lexica from *most* Supervised n-gram Models

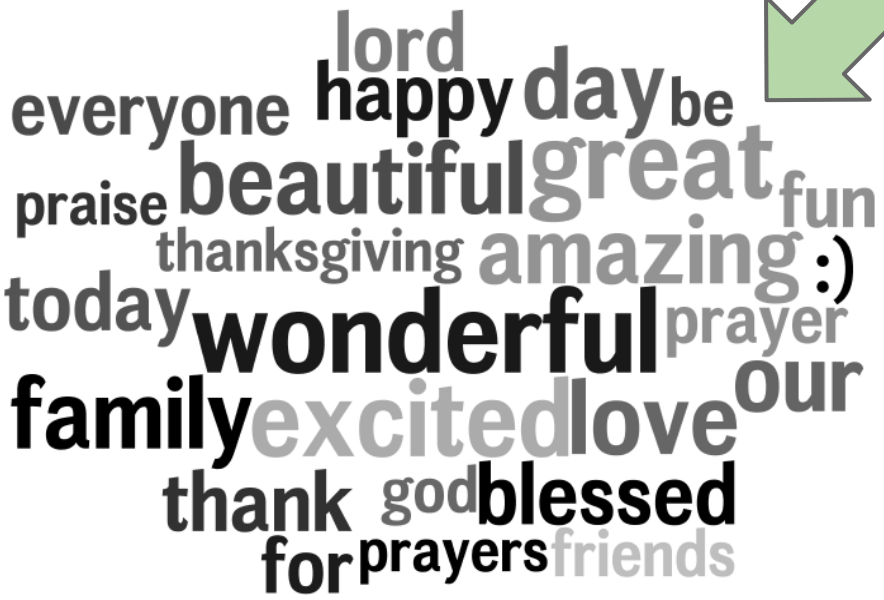
- Generalize multi-variate regression model into lexica.
- Works at multiple levels:
  Hand annotated messages or users
  - OCEAN: User-level Cambridge data set
  - PERMA: MTurk-ed messages



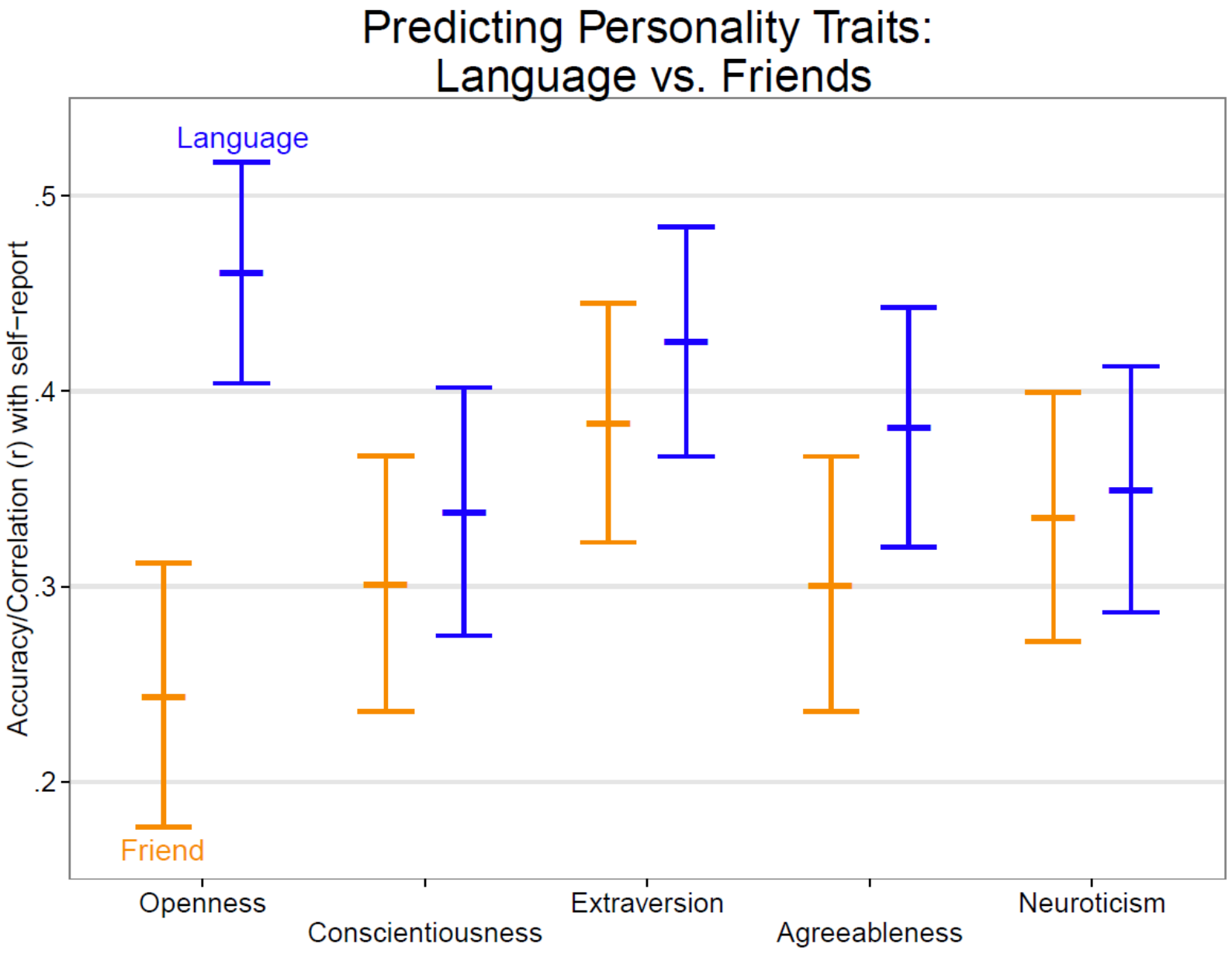
Multivariate or Univariate for *Insights*?





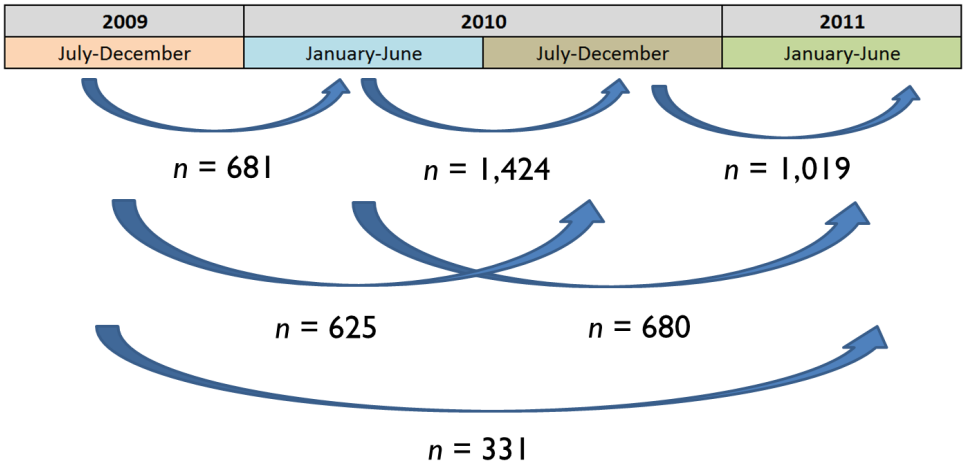


Language-Based Psychometrics

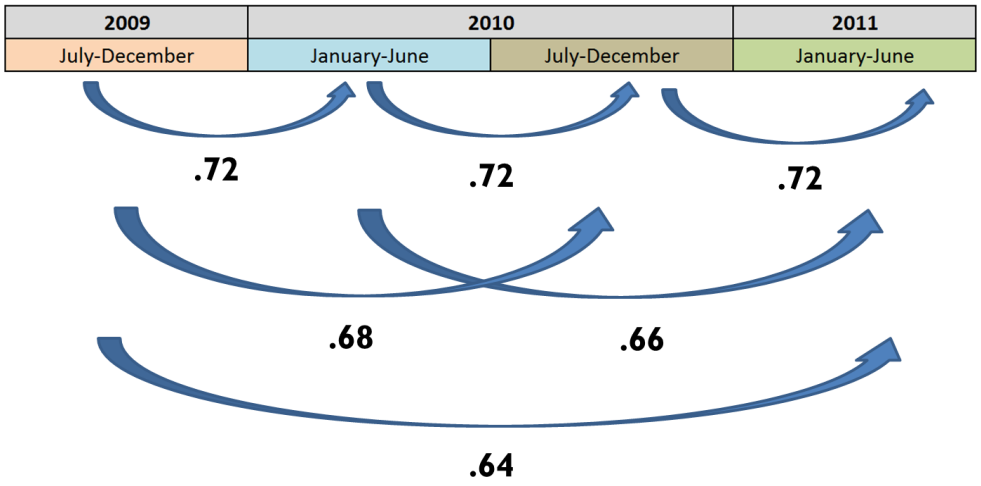


Language-Based Psychometrics

Sample sizes



Extraversion



		Time 2	Time 3	Time 4
O	Time 1	.71	.68	.64
	Time 2		.74	.71
	Time 3			.76
C	Time 1	.75	.74	.70
	Time 2		.76	.72
	Time 3			.76
E	Time 1	.72	.68	.64
	Time 2		.72	.66
	Time 3			.72
A	Time 1	.65	.61	.55
	Time 2		.64	.57
	Time 3			.65
N	Time 1	.62	.57	.51
	Time 2		.62	.61
	Time 3			.63