The Role of Personality, Age and Gender in Tweeting about Mental Illnesses

Daniel Preoţiuc-Pietro, Johannes Eichstaedt, Gregory Park, Maarten Sap
Laura Smith, Victoria Tobolsky, **H. Andrew Schwartz** and Lyle Ungar

Penn | World Well-Being Project

Problem

- Mental illnesses are massively underdiagnosed
- Goal is unobtrusive measurement via Social Media
- Focus on depression and PTSD

This presentation:

- Study the predictive power of multiple features
- Study the insights provided by each feature

Data

- based on Twitter self-reports

  ‘I have been diagnosed with depression’

  483 users with depression

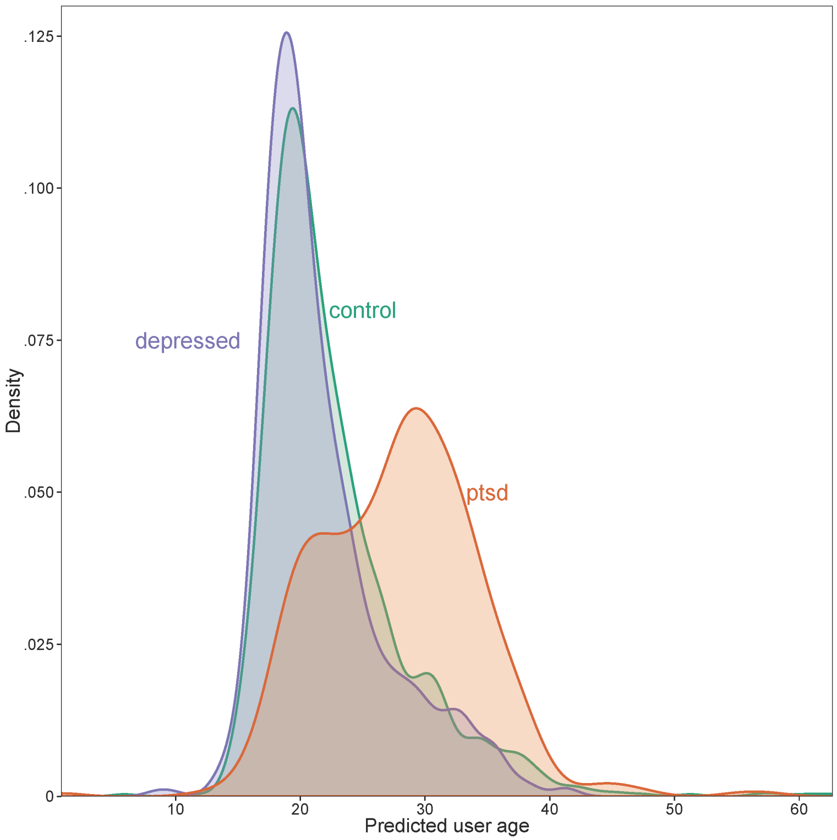
  370 users with PTSD

  1104 control users
- each user has avg. 3400 messages

(Coppersmith et. al, CLPsych 2014)

Age, Gender

- automatically inferred from Twitter posts
  (Sap et. al, EMNLP 2014)
- age predictive language use plays a massive role in separating PTSD users
- gender differences are smaller but still significant



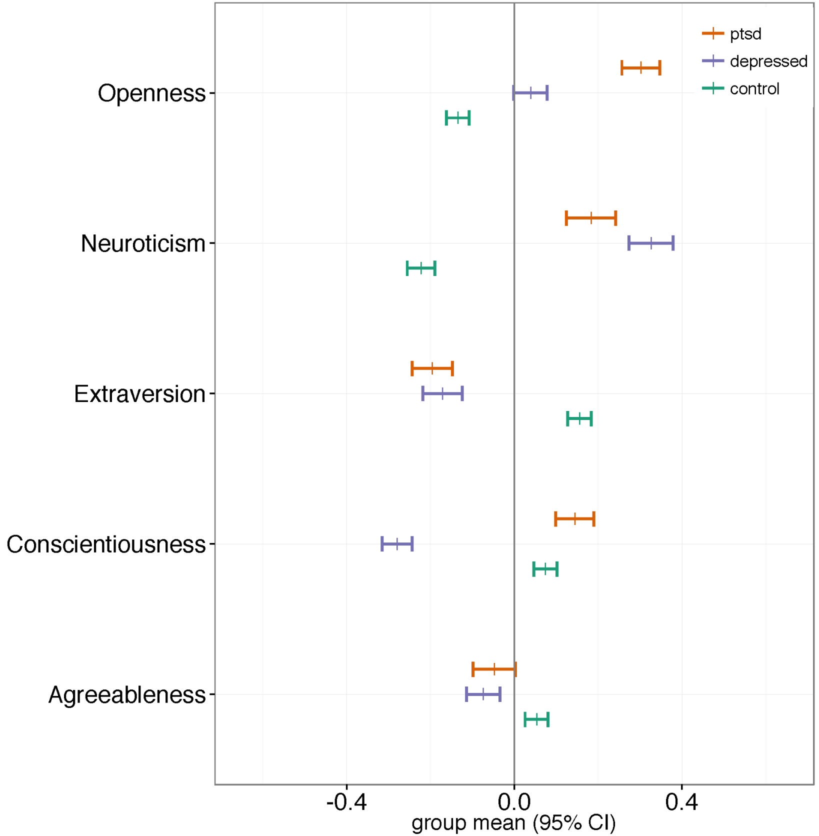
Penn | World Well-Being Project

Age, Gender

Penn | World Well-Being Project

Personality

- automatically inferred from Twitter posts
  (Park et. al 2014)
- Big Five model of personality
- showing group means of each trait



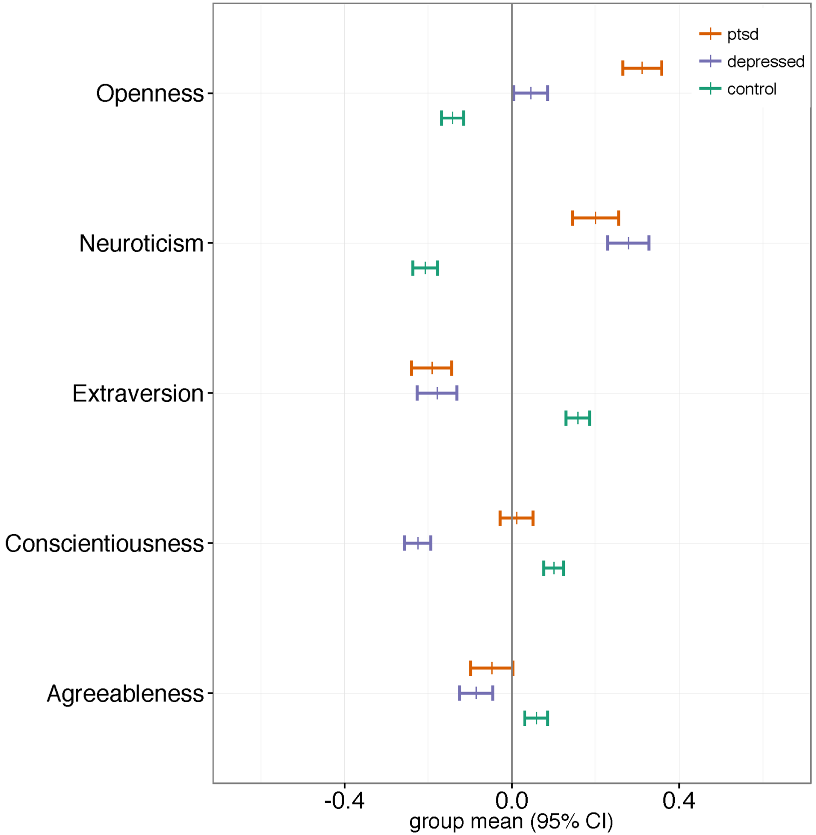
Penn | World Well-Being Project

Personality

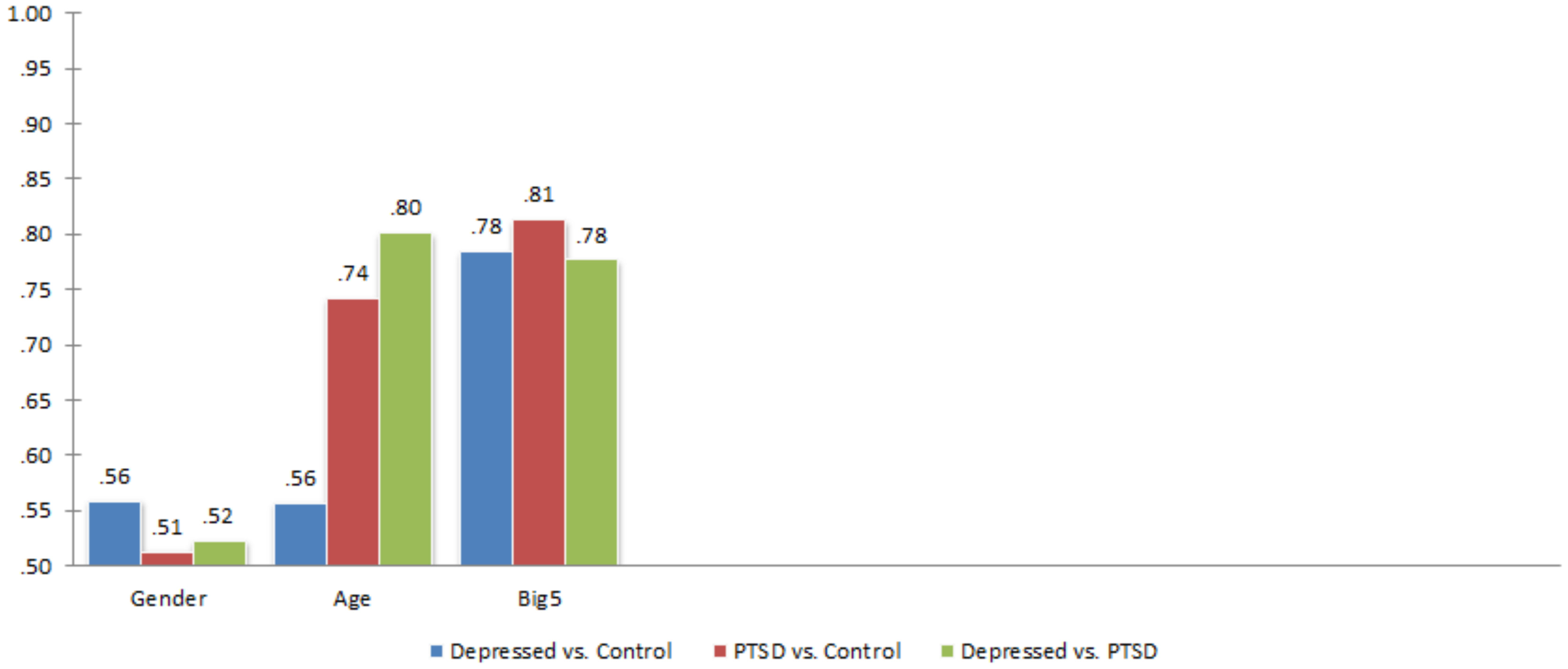
- controlling for age/gender
- mentally ill users:

1. high on neuroticism
2. more introverted
3. less agreeable

- age removes some conscientiousness and neuroticism differences

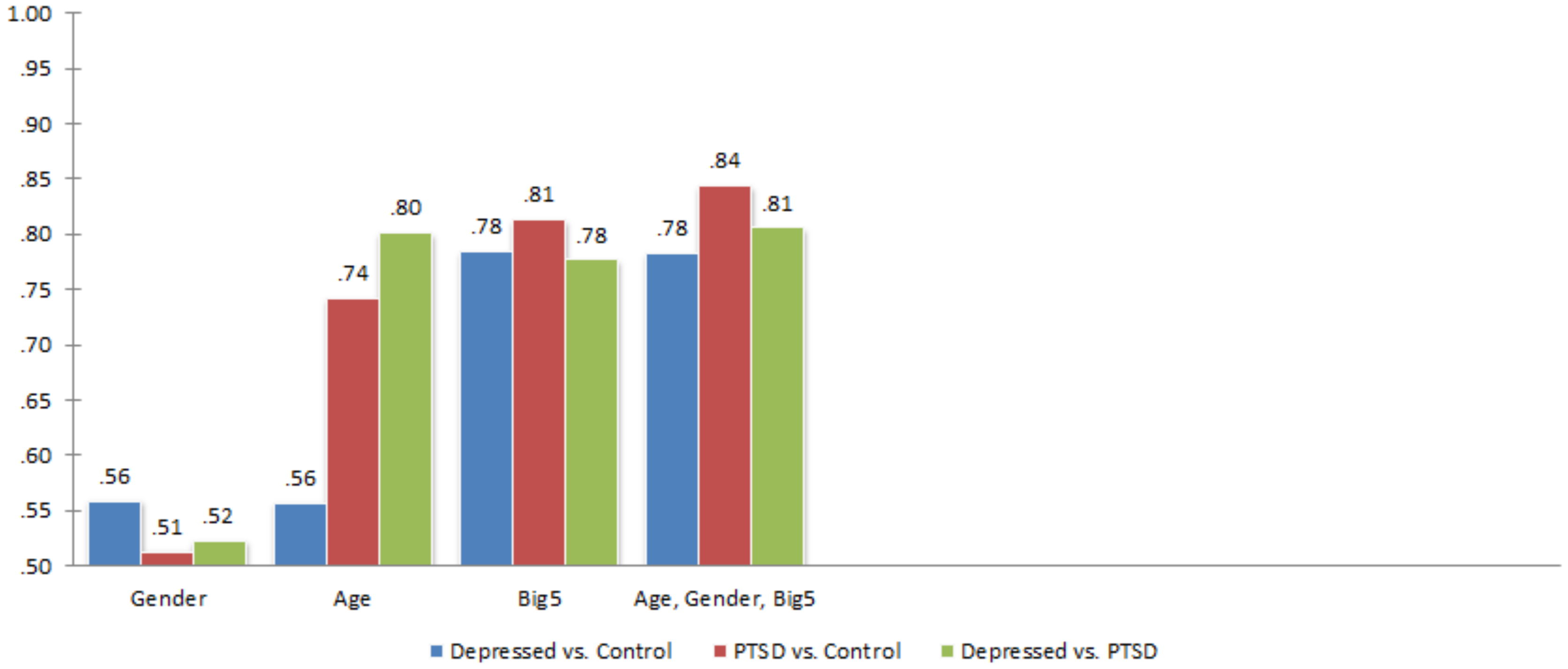


Personality



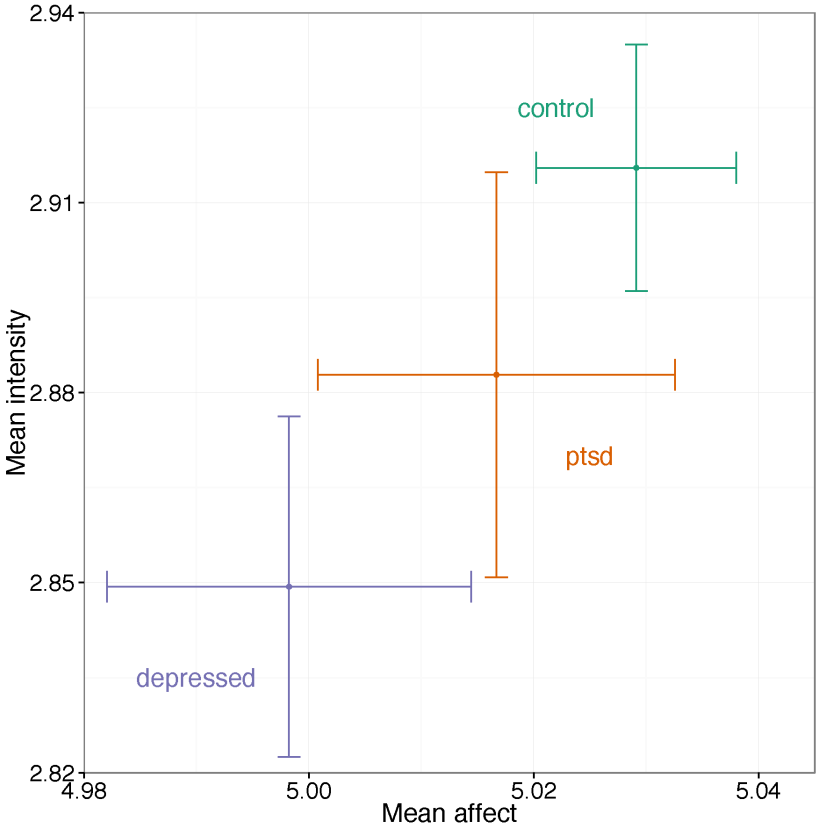
Penn | World Well-Being Project

Age, Gender, Personality



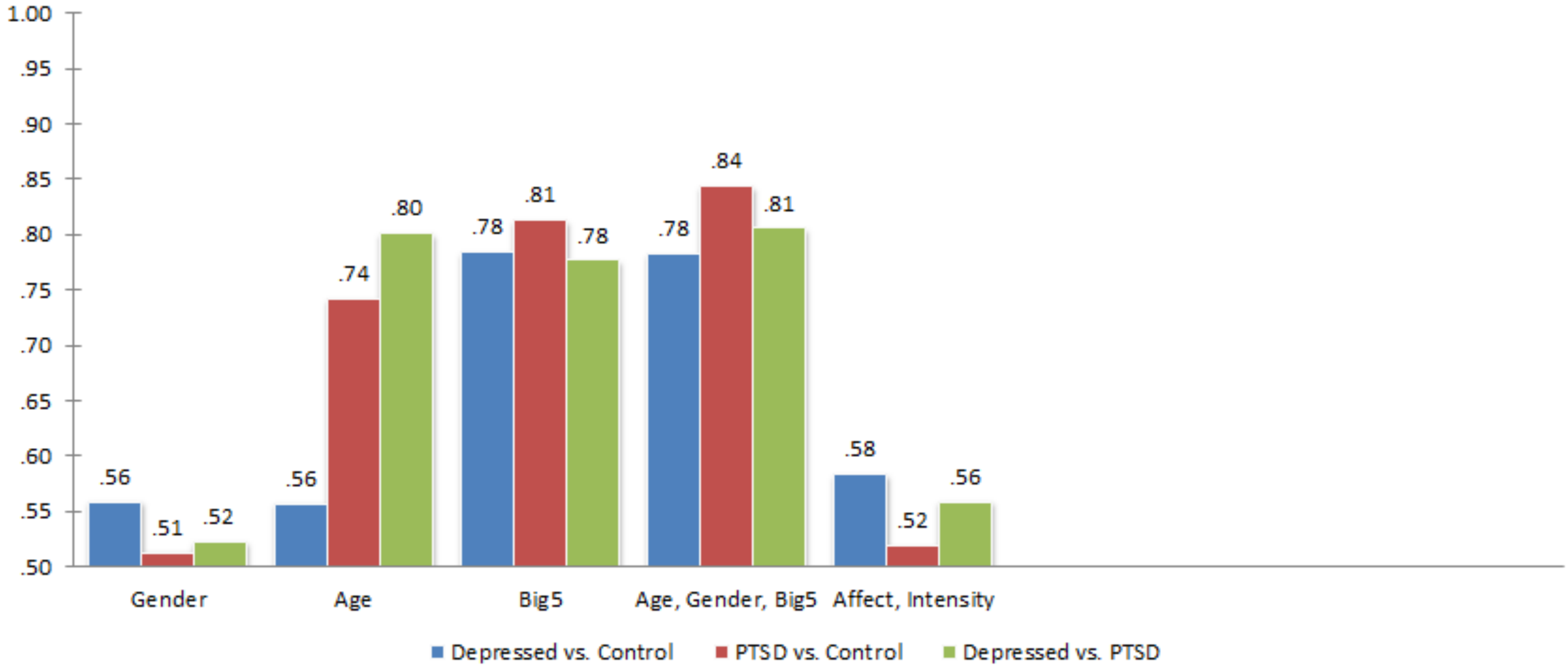
Affect and Intensity

- Model trained on 3000 annotated FB posts and applied to all user posts
- circumplex model similar to valence & arousal (ANEW)
- mentally ill users are less aroused and less positive



Penn | World Well-Being Project

Affect and Intensity

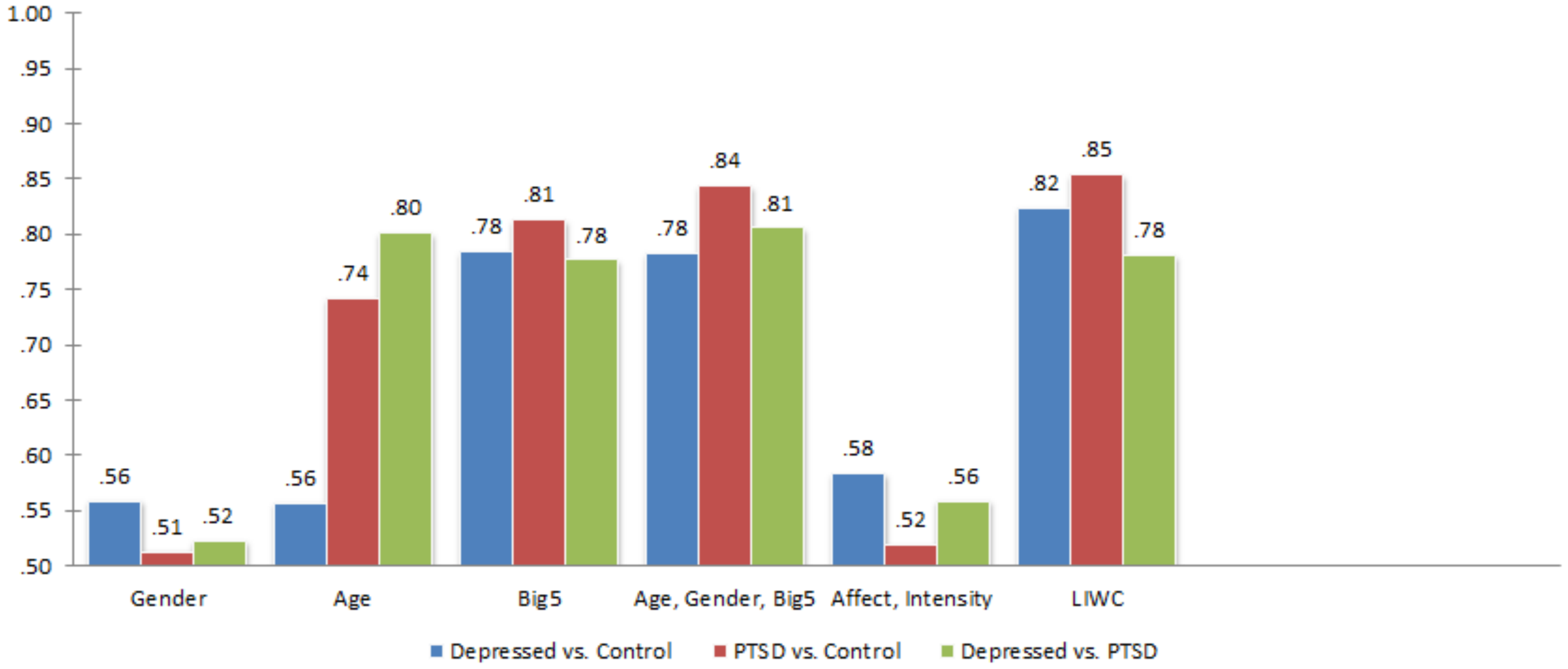


Penn | World Well-Being Project

LIWC

- standard psychologically inspired dictionaries
- 64 categories such as:
    - parts-of-speech
    - topical categories
    - emotions
- standard baseline for open vocabulary approaches

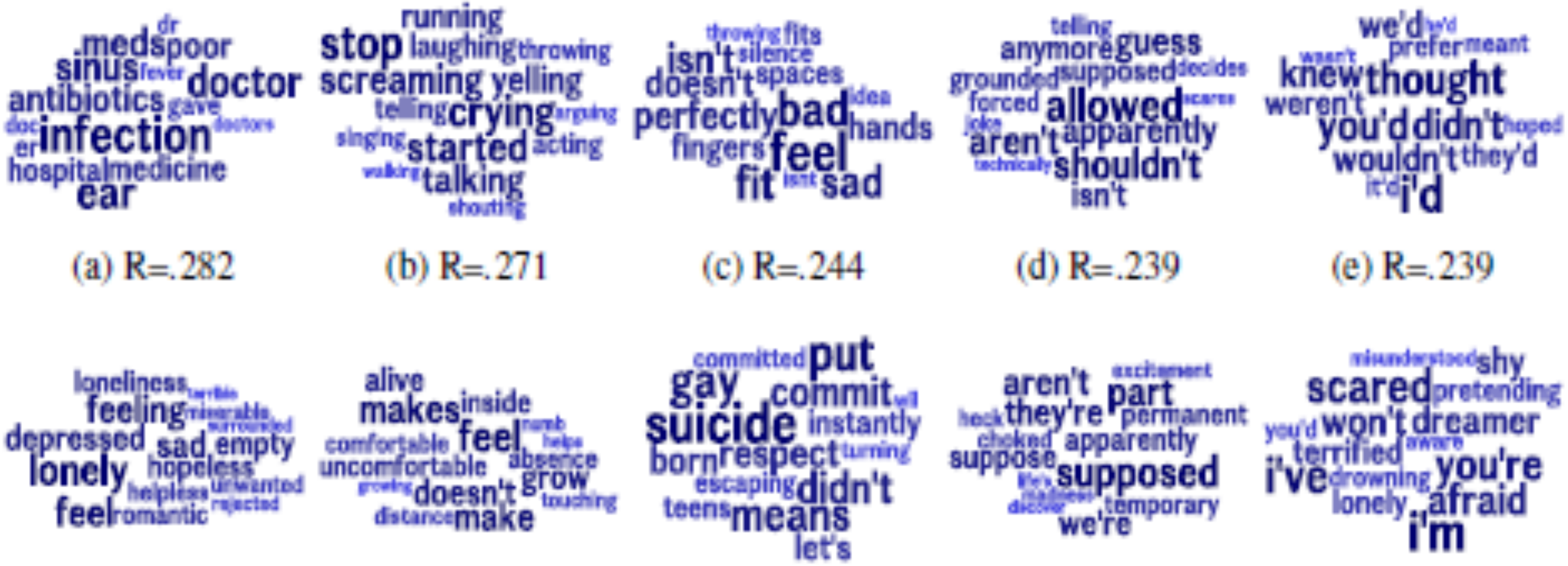
LIWC


1.00
.95
.90
.85
.80
.75
.70
.65
.60
.55
.50
.56
.51
.52
.56
.74
.80
.78
.81
.78
.78
.84
.81
.58
.52
.56
.82
.85
.78
Gender
Age
Big5
Age, Gender, Big5
Affect, Intensity
LIWC
Depressed vs. Control
PTSD vs. Control
Depressed vs. PTSD

Topics

- computed using Latent Dirichlet Allocation (LDA)
- underlying set of Facebook statues
- 2000 topics in total
- soft clusters: a word can belong to multiple topics with a probability

Topics: Depression



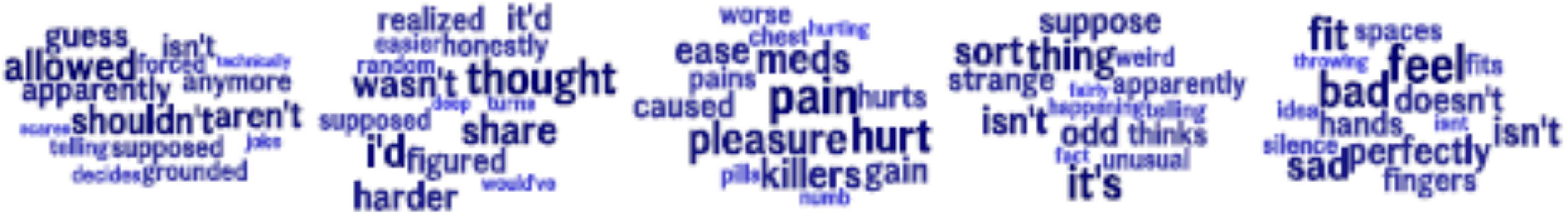
(a) R=.282 (b) R=.271 (c) R=.244 (d) R=.239 (e) R=.239

(f) R=.238 (g) R=.237 (h) R=.235 (i) R=.230 (j) R=.229

Topics controlled for age and gender

Penn | World Well-Being Project

Topics: PTSD



(a) R=.280 (b) R=.280 (c) R=.277 (d) R=.266 (e) R=.255

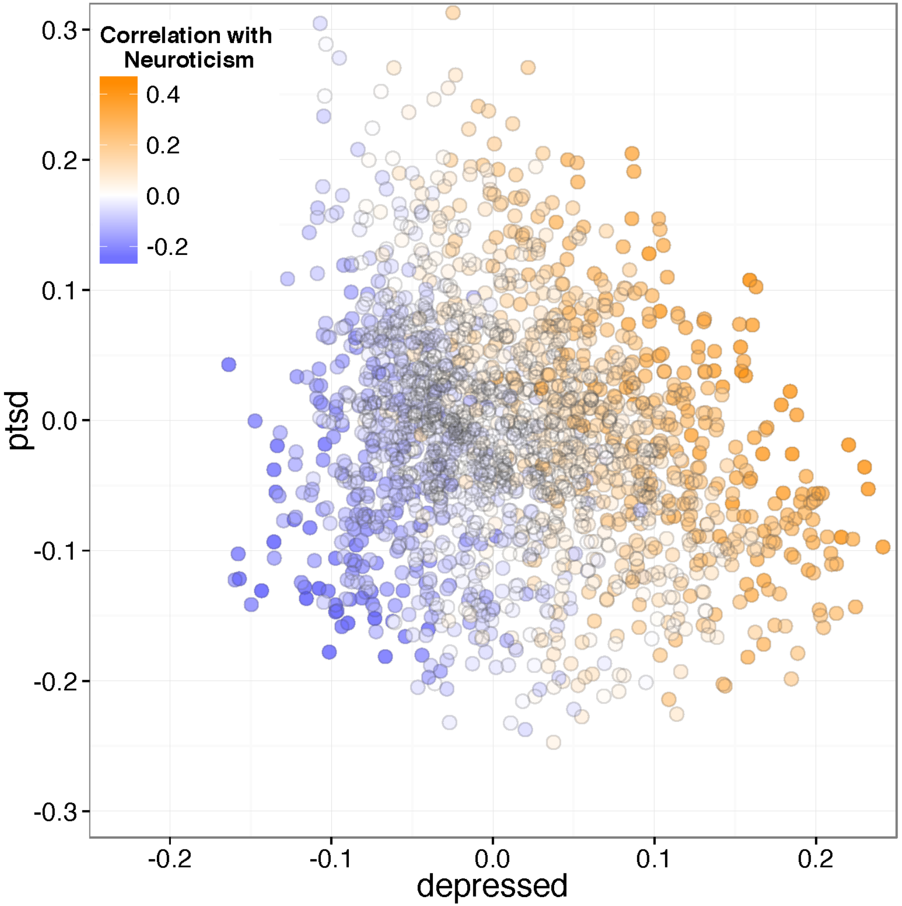
(f) R=.254 (g) R=.254 (h) R=.248 (i) R=.241 (j) R=.237

Topics controlled for age and gender

Penn | World Well-Being Project

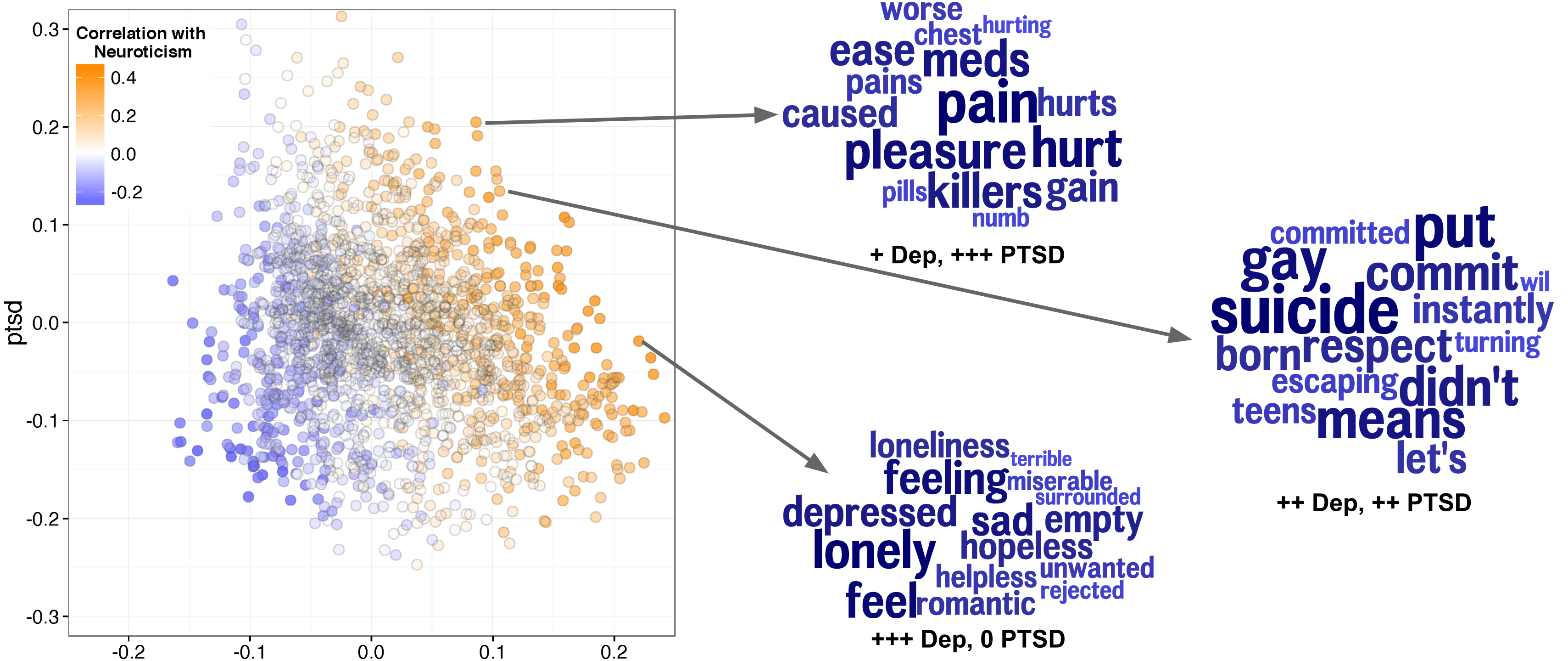
Topics: PTSD, Depression & Neuroticism

- Analysis of high neuroticism topics
- Neuroticism topics are high in both illnesses, but different ones are activated in each illness

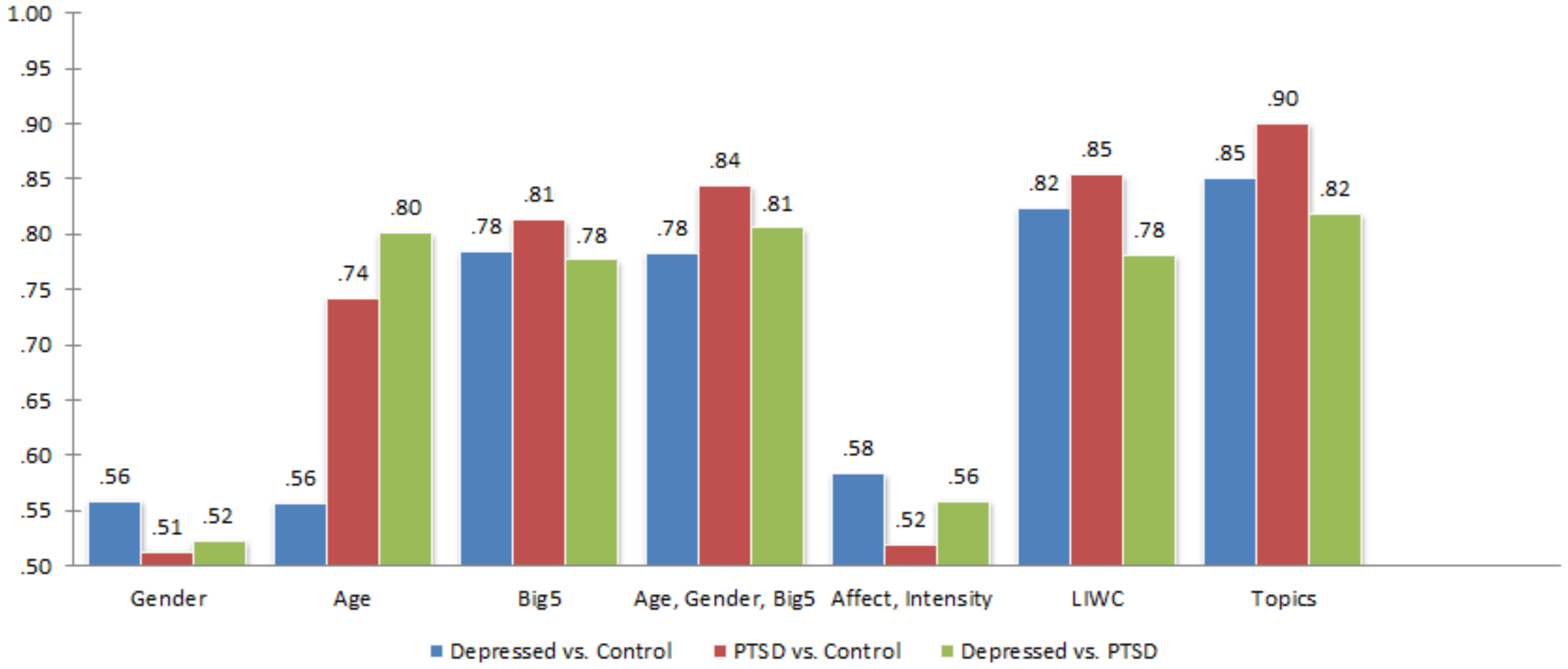


Penn | World Well-Being Project

Topics: PTSD, Depression & Neuroticism



Topics

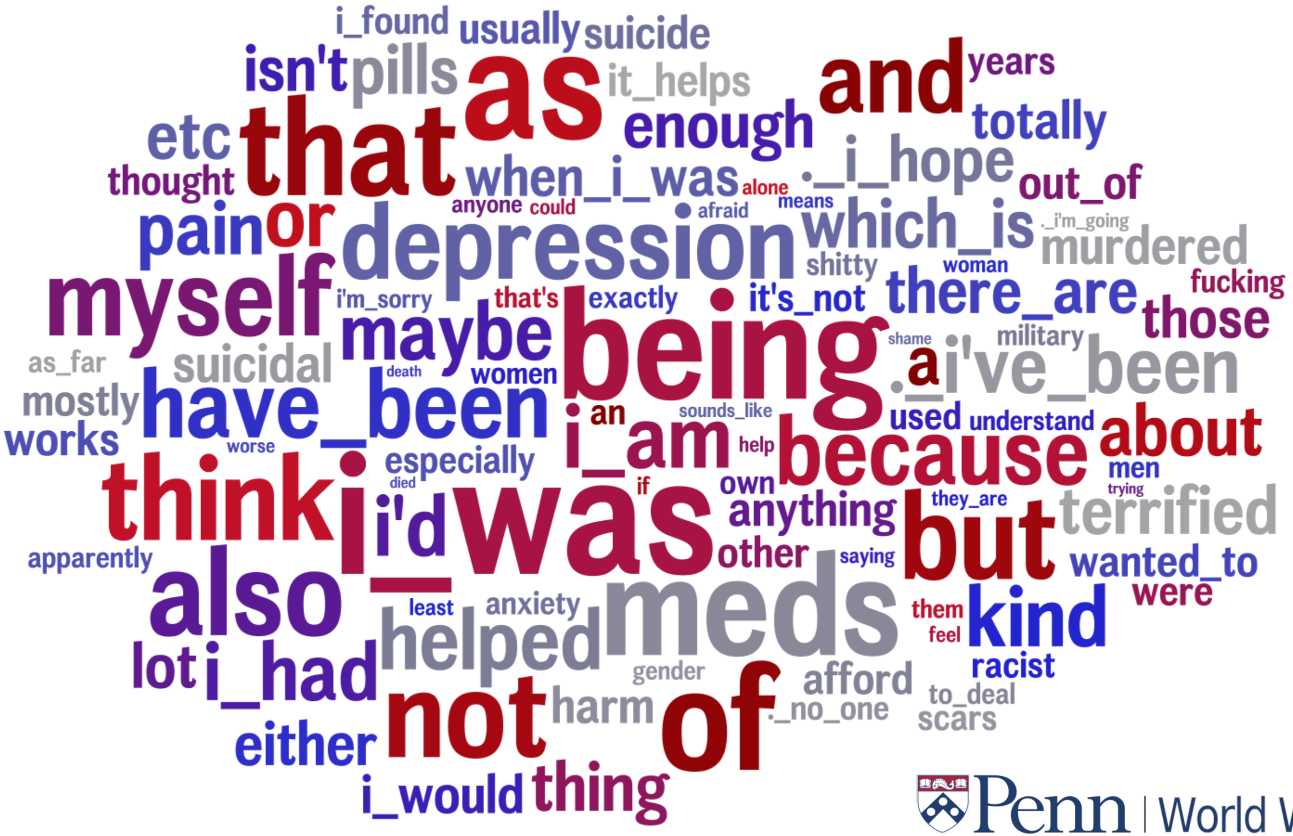


Penn | World Well-Being Project

1-3 grams: Depressed vs. Controls

Penn | World Well-Being Project

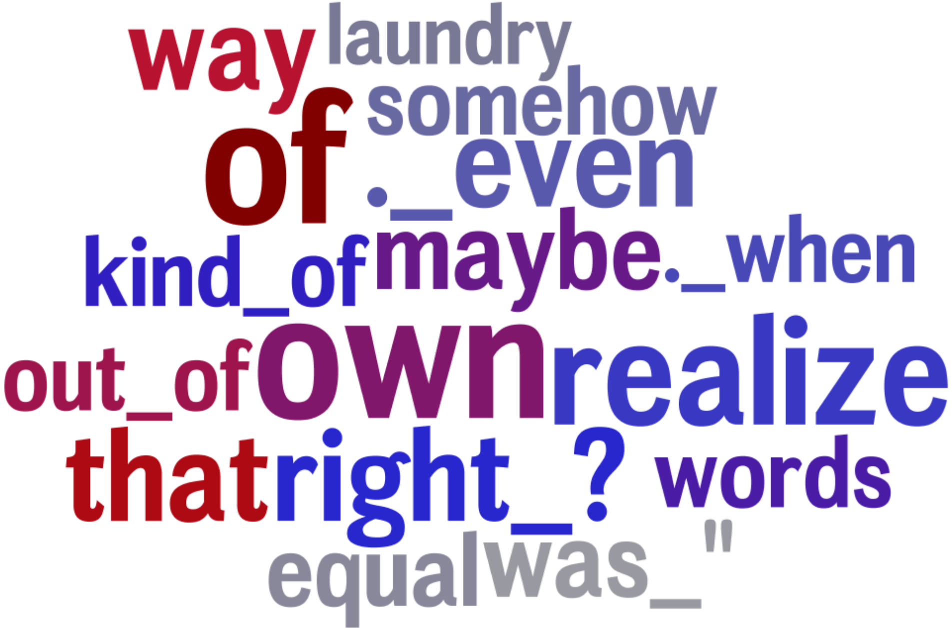
1-3 grams: PTSD vs. Controls



Penn | World Well-Being Project

1-3 grams: Depressed vs. PTSD

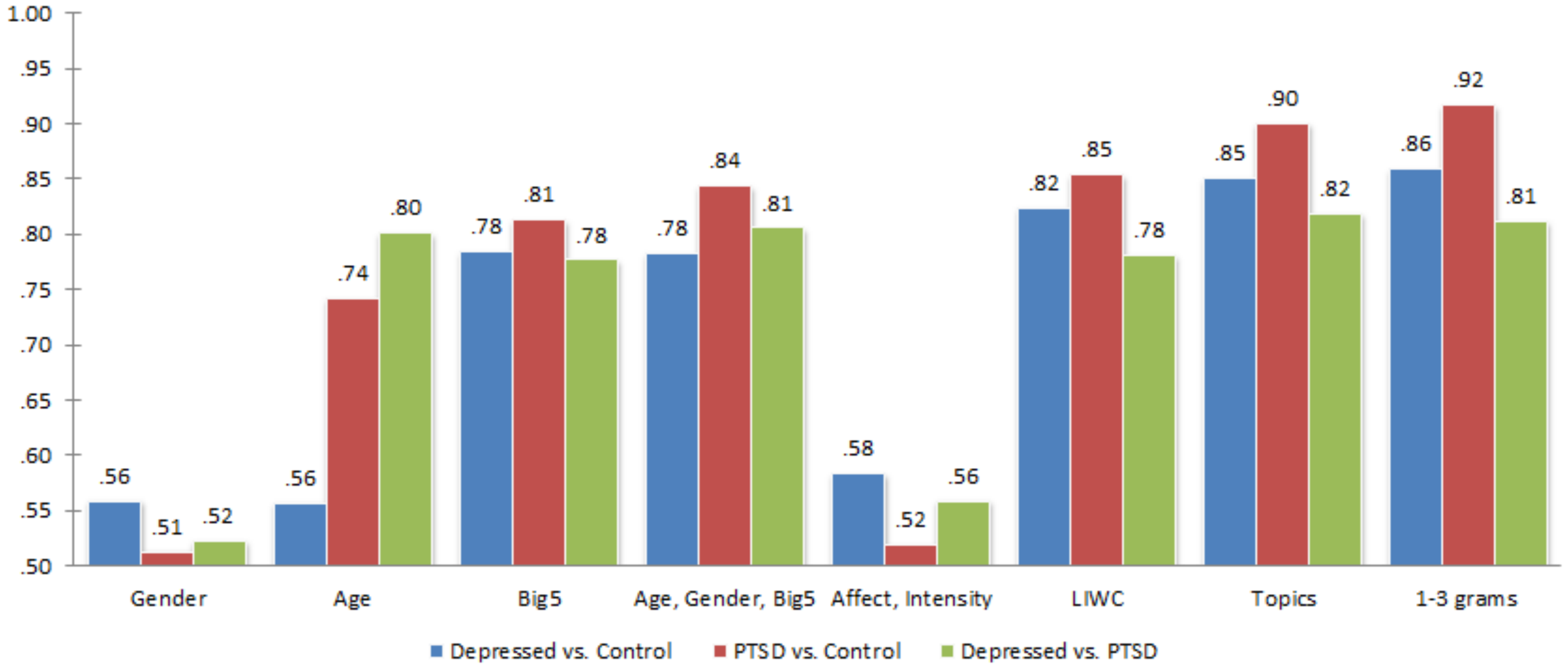
depression



Almost nothing left when controlling for age and gender

Penn | World Well-Being Project

1-3 grams



Penn | World Well-Being Project

Better Predictions

- use metadata features

  # friends, #statuses

- use different word clusters

  Brown clustering, NPMI Spectral clustering, Word2Vec/GloVe embeddings

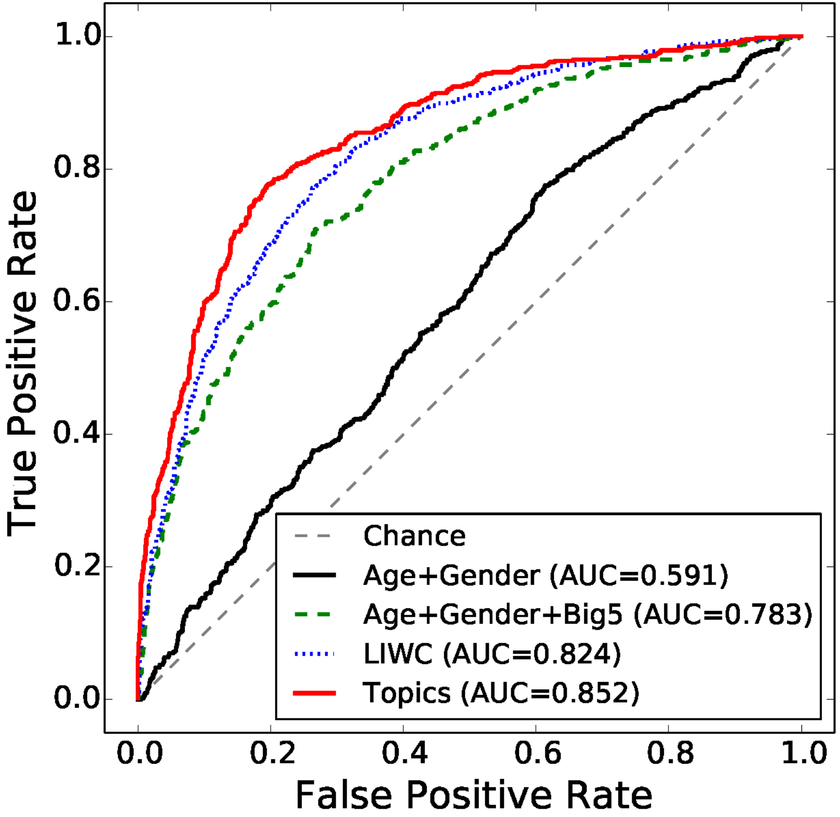
- linear ensemble of logistic regression classifiers

ROC Curve

Depressed

vs.

Controls



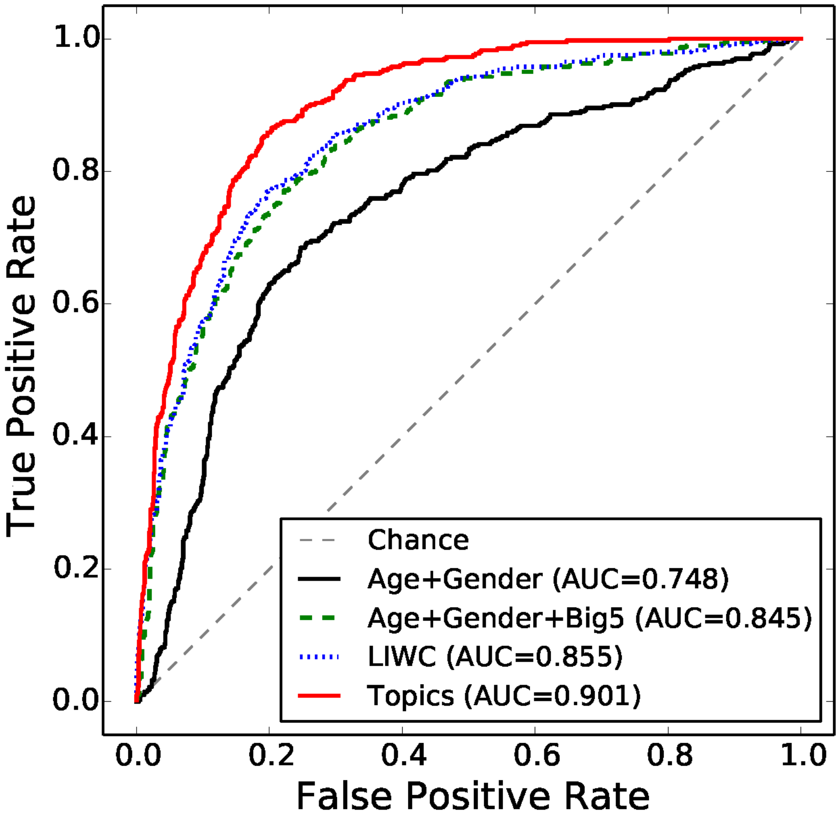
Penn | World Well-Being Project

ROC Curve

PTSD

vs.

Controls



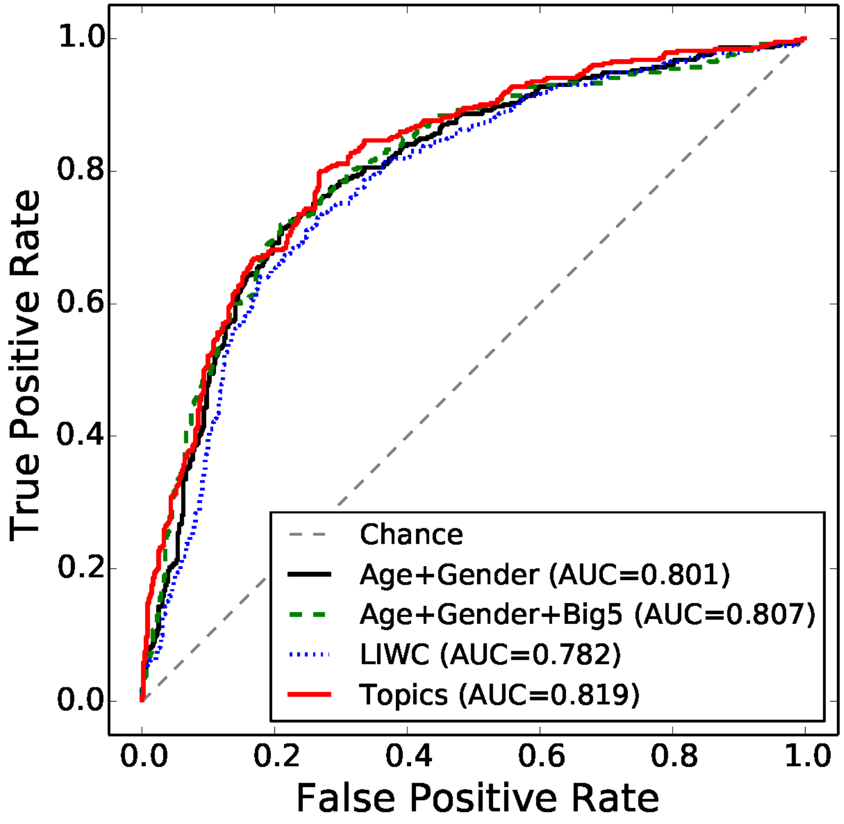
Penn | World Well-Being Project

ROC Curve

Depressed

vs.

PTSD



Penn | World Well-Being Project

Take Home

- Control the analysis for age & gender
- Personality plays an important role in mental illnesses
- Language use of depressed/PTSD users uncovers interesting patterns

Thank you!

Penn | World Well-Being Project

wwbp.org