Multimodal Analysis and Prediction of Latent User Dimensions

Laura Wendlandt and Rada Mihalcea

University of Michigan
{wenlaura,mihalcea}@umich.edu

Ryan L. Boyd and James W. Pennebaker

The University of Texas at Austin
{ryanboyd,pennebaker}@utexas.edu

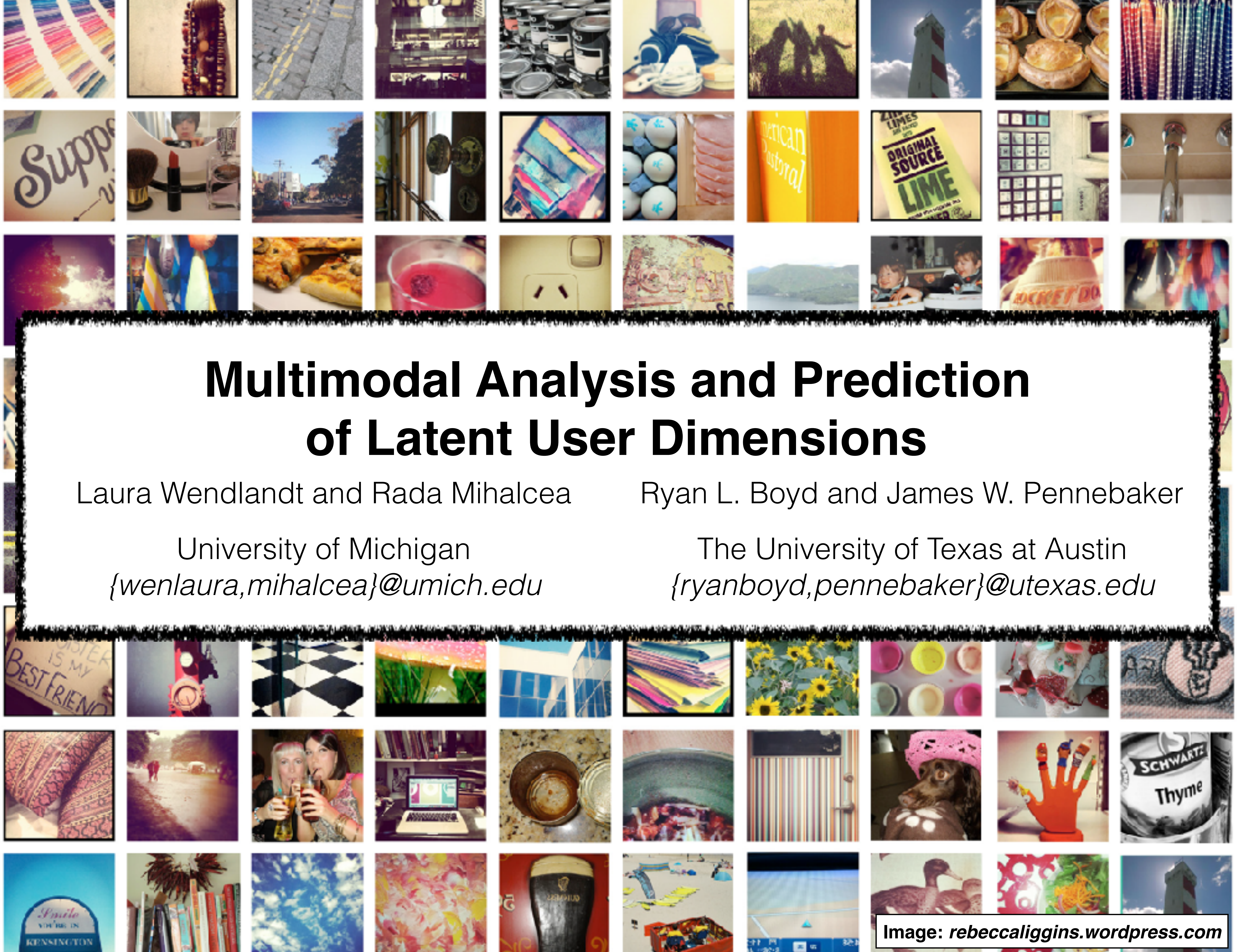
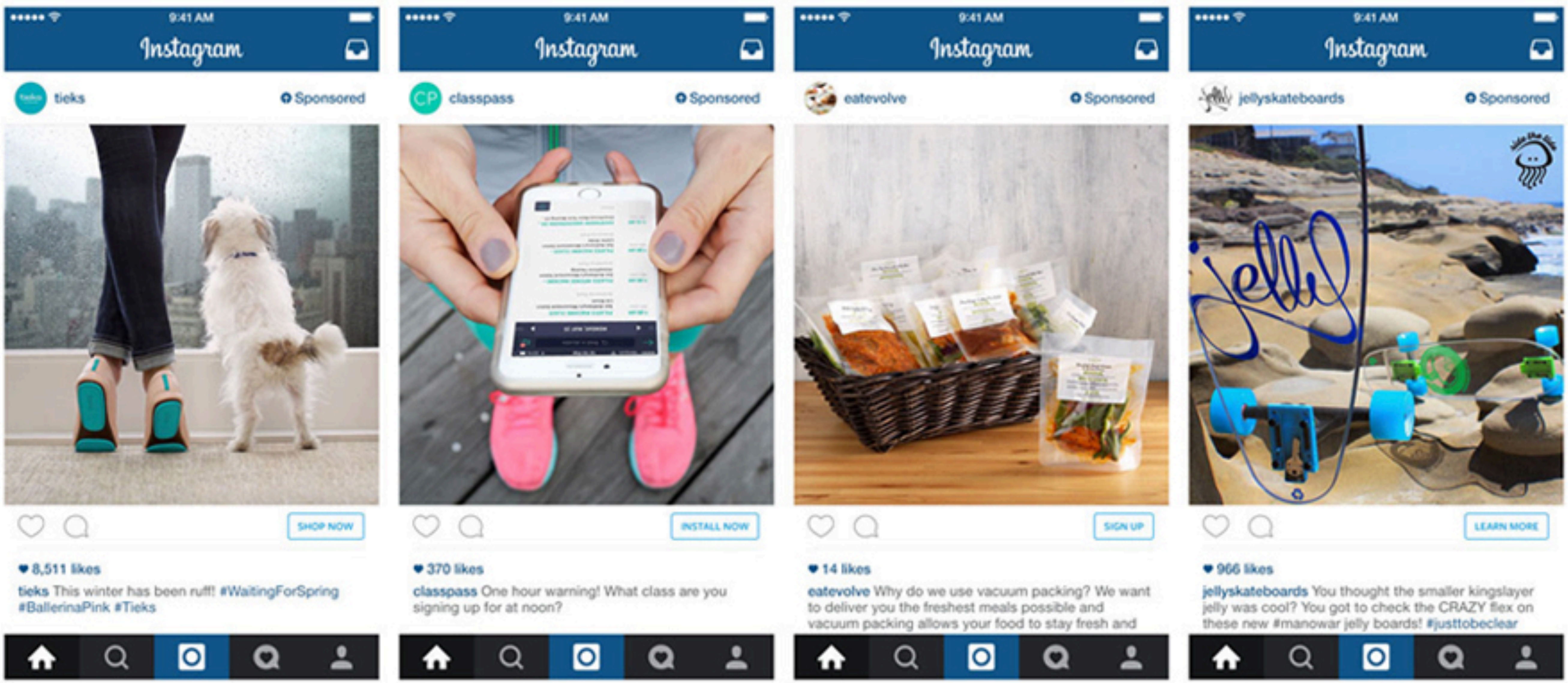


Image: ***rebeccaliggins.wordpress.com***

1.8 ***Billion*** Images / Day!



Conclusions

1. Correlational techniques provide **interpretable psychological insight** into personality and gender.

2. For the task of personality prediction, **multimodal models** outperform both visual features and textual features in isolation, using a relatively small dataset.

Outline

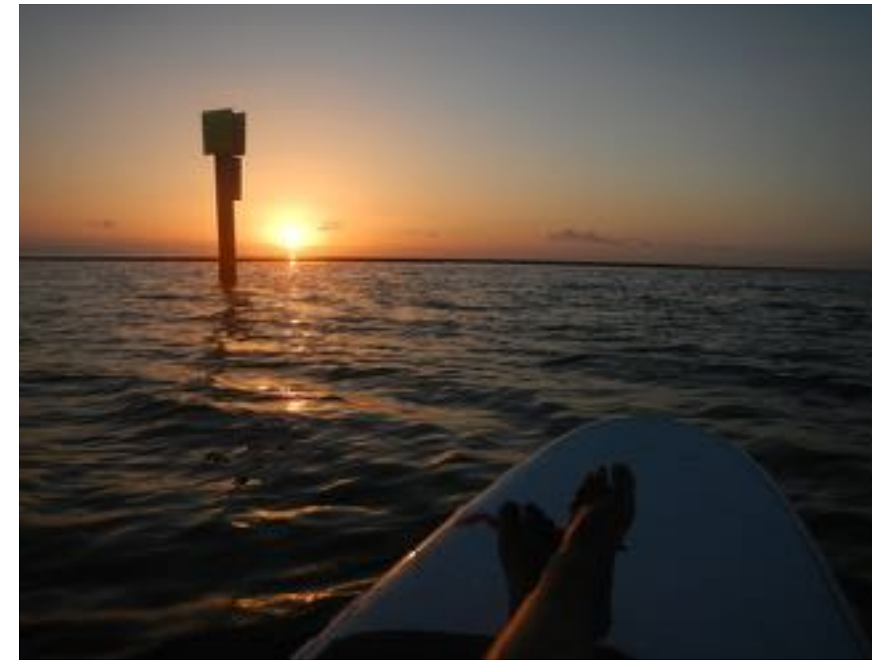
1. **Dataset**
2. Features
3. Correlational Analysis
4. Multimodal Prediction

Dataset

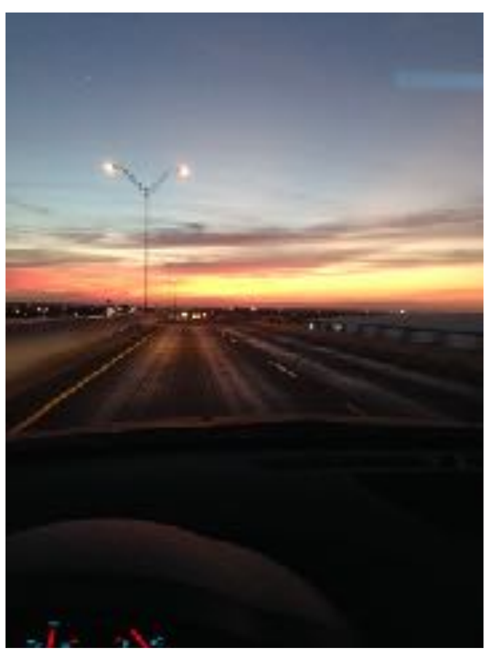
- Sam Gosling & James Pennebaker, UT Austin
- Fall 2015 introductory undergraduate psych class
- Students from all majors
- Images, captions, gender, & personality
- 1,353 students

The real me is right behind you.

Gotta find something to do when I have nothing to say.



I’d rather be on the water.



I crossed this bridge almost every day for 18 years and never got tired of it.

The littlest things are always so pretty (and harder to capture).

Big 5 Personality Traits

Openness
Artistic
Curious
Original

Conscientiousness
Efficient
Organized
Thorough

Extraversion
Assertive
Enthusiastic
Outgoing

Agreeableness
Appreciative
Sympathetic
Trusting

Neuroticism
Anxious
Unstable
Worrying

McCrae and John 1992

Outline

1. Dataset
2. **Features**
3. Correlational Analysis
4. Multimodal Prediction

Feature Extraction

- Want meaningful and interpretable features with some connection to the user

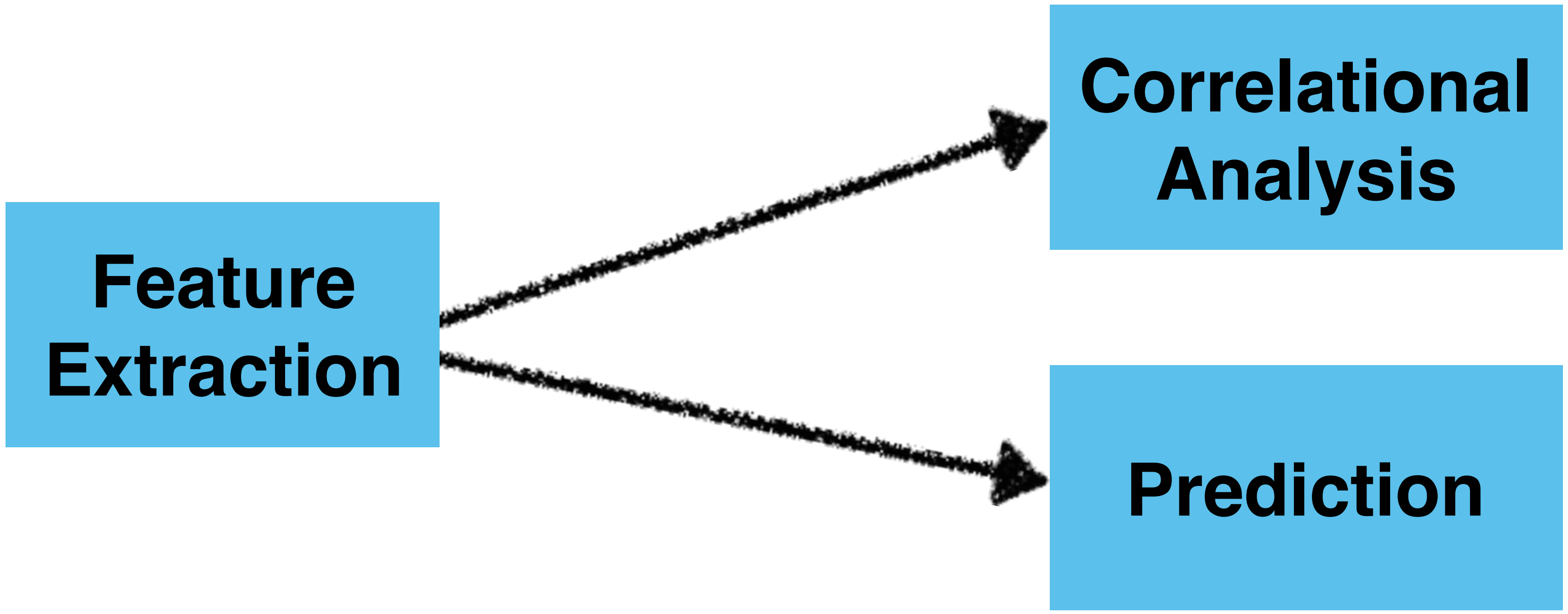
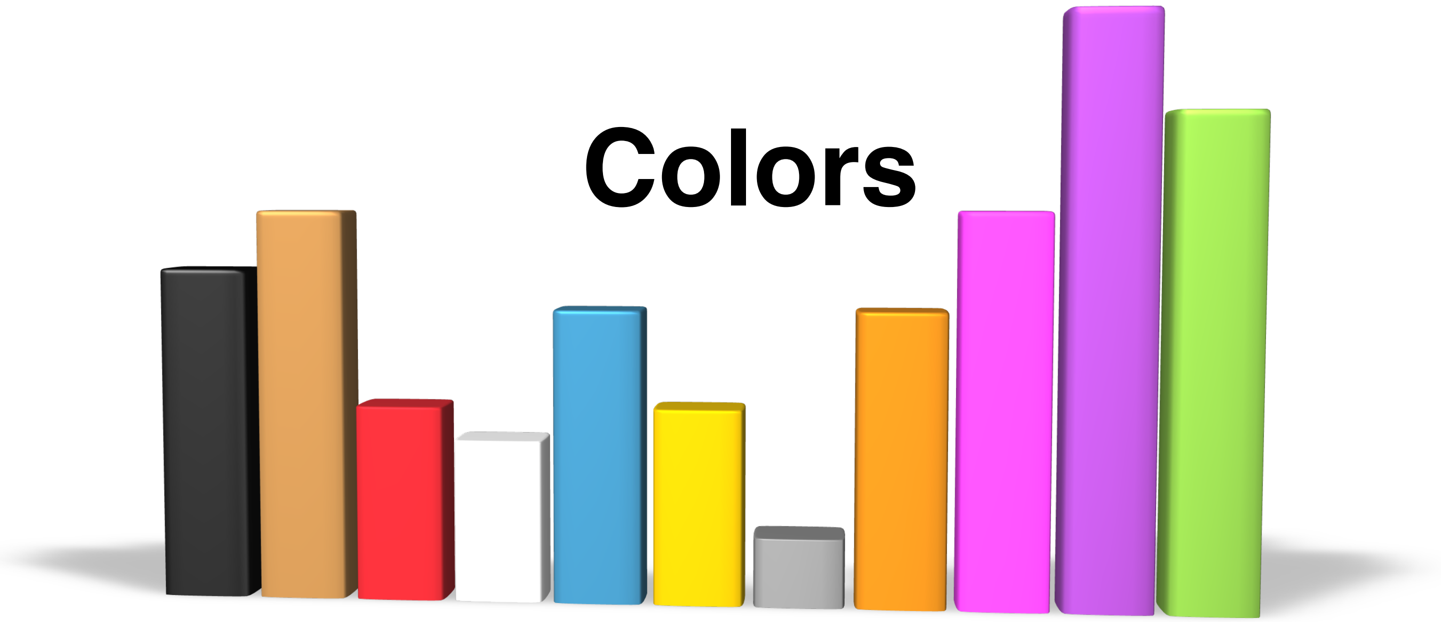


Image Attributes

Colors



Scenes

Objects

Faces

Caption Attributes

Named Entities

Readability

word2vec

LIWC

Anxiety
worried
fearful
nervous

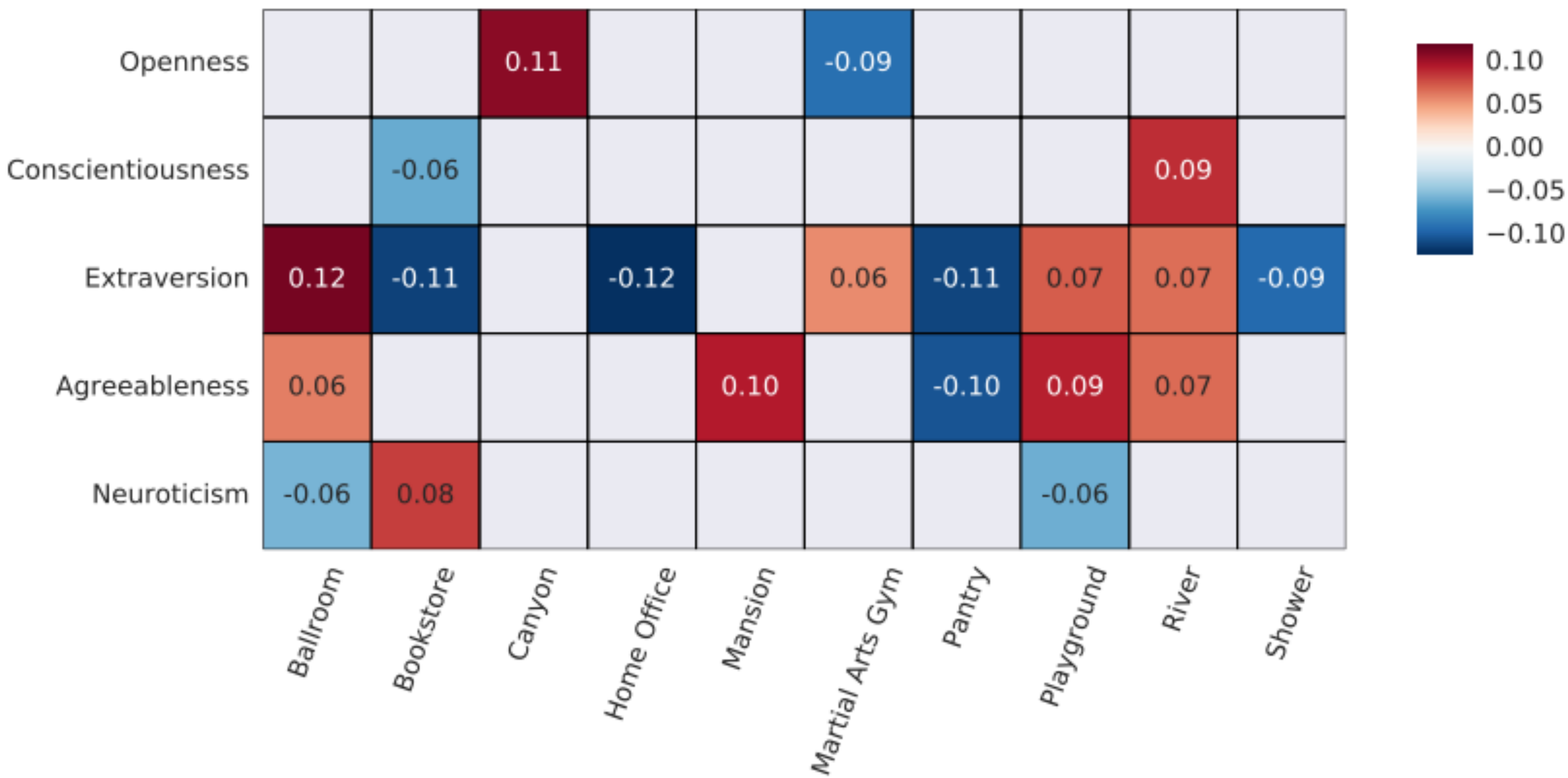
Causation
because
effect
hence

Images: dpti.sa.gov.au, www.tckpublishing.com, Mikolov et al. 2013

Outline

1. Dataset
2. Features
3. **Correlational Analysis**
4. Multimodal Prediction

Selected Correlations: Scenes



Selected Gender Effects: Scenes

Female	*Effect Size*
Beauty Salon	0.347
Ice Cream Parlor	0.340
Slum	0.286
Herb Garden	0.224
Art Studio	0.221

Male	*Effect Size*
Office	0.290
Football Stadium	0.267
Baseball Stadium	0.222
Gas Station	0.222
Music Studio	0.222

Images: herb.co

Outline

1. Dataset
2. Features
3. Correlational Analysis
4. **Multimodal Prediction**

Classification Task

- **Data division**:
  High segment > $\mu + 0.5\sigma$
  Low segment < $\mu - 0.5\sigma$

- **Random forest**: 500 trees
  (10-fold cross validation across individuals)

- **Baseline**: Most common training class

- **Comparison**: Mairesse et al. 2007

Image-Enhanced Unigrams (IEUs)

Color: yellow, orange

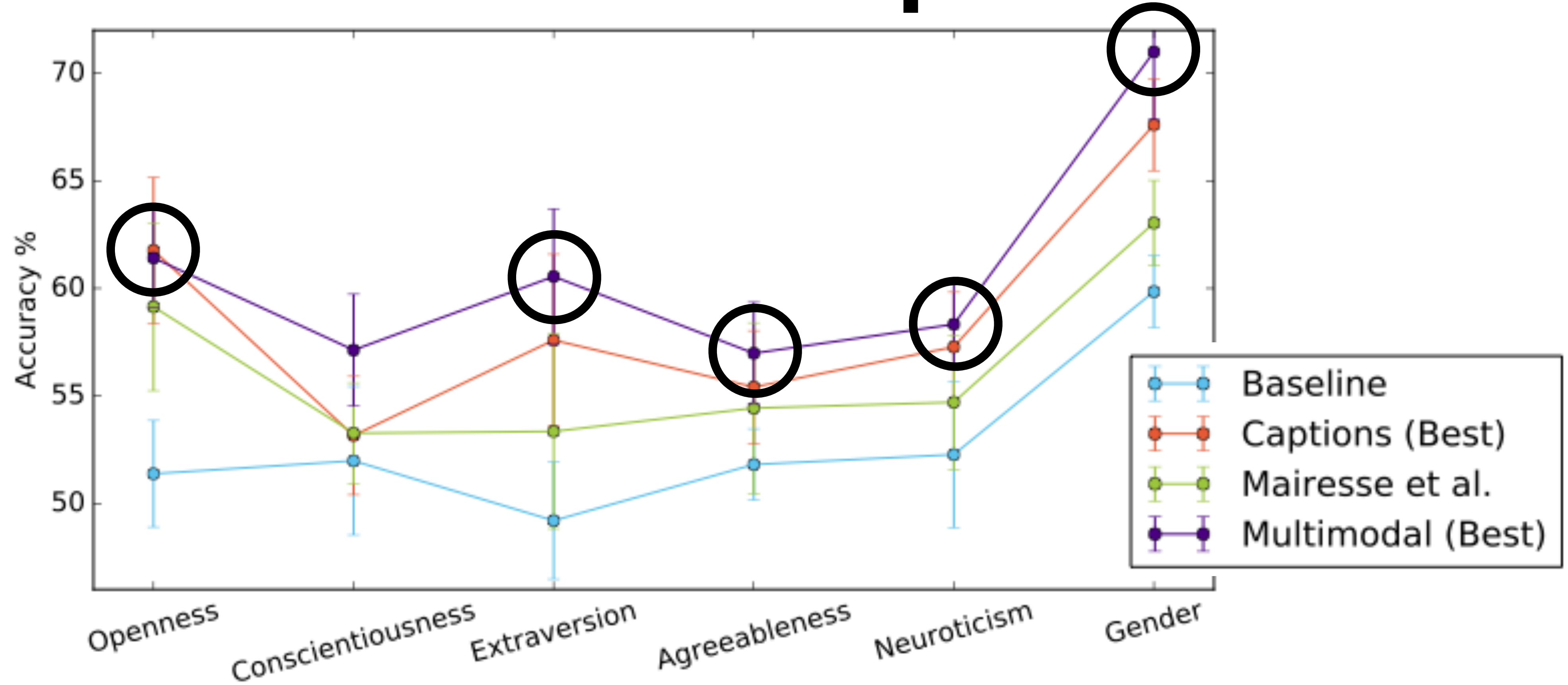
Scene: kitchen

Objects: plate, fork, salad, table

Best Multimodal Model

1. Extract IEUs from each individual images
2. Combine IEUs with caption unigrams
3. Get *word2vec* embedding for each unigram
4. Average all embeddings to get feature vector

Classification Comparison



	O	*C*	*E*	*A*	*N*	*Gender*
Baseline	51.4±2.5	52±3.4	49.2±2.7	51.8±1.7	52.3±3.4	59.8±1.7
Captions (Best)	**61.7±3.4**	53.2±2.8	**57.6±4.0**	55.4±2.6	57.3±2.5	**67.6±2.1**
Mairesse et al.	**59.1±3.9**	53.3±2.3	53.3±4.5	54.4±4.0	54.7±3.1	63±2.0
Multimodal (Best)	**61.4±2.3**	57.1±2.6	**60.5±3.2**	**57±2.4**	**58.3±2.1**	**71±3.2**
Rel. Error Reduc.	5.6%	8.1%	15.4%	5.7%	7.9%	21.6%

Conclusions

1. Correlational techniques provide **interpretable psychological insight** into personality and gender.

2. **Multimodal models** outperform both visual features and textual features in isolation, using a relatively small dataset.

Acknowledgements

- Chris Pittman *(data collection help)*, Shibamouli Lahiri *(readability code)*, Steve Wilson *(implementation of Mairesse et al.)*

- *Funding*: National Science Foundation (#1344257), the John Templeton Foundation (#48503), the Michigan Institute for Data Science