Analyzing Personality through Social Media Profile Picture Choice

Daniel Preoţiuc-Pietro
Leqi Liu, Zahra Riahi, Mohsen E. Moghaddam, Lyle Ungar
ICWSM 2016

Positive Psychology Center
University of Pennsylvania

22 March 2016

Personality Guess

Which personality trait are users with these real Twitter Profile pictures high in?

Personality

- **Openness to experience**:
  - (+) inventive/curious – consistent/cautious (-)
- **Conscientiousness**:
  - (+) efficient/organized – easy-going/careless (-)
- **Extraversion**:
  - (+) outgoing/energetic – solitary/reserved (-)
- **Agreableness**:
  - (+) friendly/compassionate – analytical/detached (-)
- **Neuroticism**:
  - (+) sensitive/nervous – secure/confident (-)

Personality Guess

Which personality trait are users with these real Twitter Profile pictures high in?

Personality Guess

Which personality trait are users with these real Twitter Profile pictures high in?

+ Extraversion

+ Conscientiousness

Personality Guess

Twitter profile pictures - an image the user considers representative for their online persona

Personality prediction from standard photos is a relatively well studied problem in psychology (*Penton-Voak et al. 2006, Naumann et al. 2009*)

Humans are good at predicting some personality traits from a single photo: extraversion

Research Questions

1. **Can we automatically predict personality from profile picture choice?**

2. **What are the distinctive features of profile photos for each personality trait?**

Research Questions

1. Can we automatically predict personality from profile picture choice?

Yes! *(Celli et al. 2014), (Al Moubayed et al. 2014)*

2. What are the distinctive features of profile photos for each personality trait?

Bag-of-Visual-Words or Deep learning are hardly interpretable

Use facial and attractiveness features

Two Datasets

TwitterText:

- 66,502 Twitter users
- text predicted personality
- self-reported gender
- text predicted age
- 104,500,740 tweets

TwitterSurvey:

- 434 Twitter users
- survey personality
- self-reported gender
- self-reported age
- used for validation (no statistical power)

Image Features - Color

- Image is Grayscale?
- RGB Spectrum
  - Red
  - Green
  - Blue
  - Average
- Brightness
- Contrast
- Saturation
- Hue
- Colorfulness
- Naturalness
- Sharpness
- Blur
- Color Emotions

Human judgements of the attractiveness of images are influenced by:

- color distributions *(Huang, Wang, and Wu 2006)*
- aesthetic principles related to color composition *(Datta et al. 2006)*

Image Features - Color

- Image is Grayscale?
- RGB Spectrum
  - Red
  - Green
  - Blue
  - Average
- Brightness
- Contrast
- Saturation
- Hue
- Colorfulness
- Naturalness
- Sharpness
- Blur
- Color Emotions

Black/White photos are more 'artistic'

Previous research showed that colors from images are related to psychologic traits *(Wexner 1954)*

- red – 'exciting-stimulating', 'protective-defending'
- green – 'calm-peaceful-serene'
- blue – 'secure-comfortable', 'calm-peaceful-serene'

Image Features - Color

- Image is Grayscale?
- RGB Spectrum
  - Red
  - Green
  - Blue
  - Average
- Brightness
- Contrast
- Saturation
- Hue
- Colorfulness
- Naturalness
- Sharpness
- Blur
- Color Emotions

High saturation indicates vividness and chromatic purity, which are more appealing to the human eye

Colourfulness = The difference against gray *(San Pedro and Siersdorfer 2009)*

Naturalness = The degree of correspondence between images and human perception *(Huang, Wang, and Wu 2006)*

Sharpness = Measures coarseness or the degree of detail contained in an image.

A proxy for the quality of the photographing gear and photographer *(Ke, Tang, and Jing 2006)*

Image Features - Color

- Image is Grayscale?
- RGB Spectrum
  - Red
  - Green
  - Blue
  - Average
- Brightness
- Contrast
- Saturation
- Hue
- Colorfulness
- Naturalness
- Sharpness
- Blur
- Color Emotions

Affective tone of colors *(Wei-ning, Ying-lin, and Sheng-ming 2006)*
Represented by 17 color histogram features

Correlations

Feature	Demographics		Personality Trait				
Color	Gender	Age	Ope	Con	Ext	Agr	Neu
Grayscale	-.050	-.014	.050	-.031	-.012		.014
Red	.026				-.041		
Green	-.021	.012			.021		.011
Blue	-.022				.045		
Average RGB	.030	.015		.025	.033	.019	
Brightness	.019			.030		.022	
Contrast			.014	.016		.017	-.011
Saturation	.046	.014	.013				
Hue		-.022	-.019	-.017	.024		.015
Colorfulness			-.014		.041	.030	-.034
Naturalness		.026	-.017	.014	-.028	.015	-.013
Sharpness	-.053		.028	-.026	.016	-.022	
Blur		.053	-.016	.036		.021	
Average Color Emotions	.020		-.020			.023	-.016

Pearson correlations between profile image and Big Five personality controlled for age and gender and with age and gender (coded as 1 – female, 0 – male) separately. Positive correlation is highlighted with green (paler green $p < .01$, deeper green $p < .001$, two-tailed t-test) and negative correlation with red (paler red $p < .01$, deeper red $p < .001$, two-tailed t-test).

Interpretation

Aesthetically pleasing images:

+ Brightness, **+** Contrast, **+** Sharpness, **+** Saturation, **-** Blur

Artistic images:

- Naturalness, **+** Grayscale

All correlated with **Ope** !

Interpretation

Correlated with **Ope** & **Neu**:

- Colorfulness, **-** Color Emotions

However, **Neu** no significant correlations with any of the aesthetically pleasing features.

Interpretation

Correlated with **Agr**:

Anti-correlated with all the aesthetically pleasing features.

Highest correlations with ‘Average Color Emotions’, ‘Colorfulness’ and ‘Brightness’

Ext shows similar, albeit lower correlations.

Image Features - Composition

- Rule of Thirds
- Edge Distribution
- Hue Count
- Visual Weight
- Static Lines
- Dynamic Lines

Edge Distribution = Spatial distribution of the high frequency edges of an image

In good quality photos, the edges are focused on the subject

The number of unique hues of a photo is another measure of simplicity

Good compositions have fewer objects, resulting in fewer distinct hues *(Ke, Tang, and Jing 2006)*.

Visual weight measures the clarity contrast between subject region and the whole image

The presence of lines in an image induces emotional effects *(Arnheim 2004)*

Correlations

Feature	**Demographics**		**Personality Trait**				
Image Composition	Gender	Age	Ope	Con	Ext	Agr	Neu
Average Rule of Thirds	.036	.052	-.029	-.022	.038	.036	-.036
Edge Distribution	-.038	.016	.046			-.051	.039
Hue Count		.026	-.016				
Visual Weight				-.017			
Static Lines	.056				.018	.019	
Dynamic Lines	.044		-.024			.033	

Pearson correlations between profile image and Big Five personality controlled for age and gender and with age and gender (coded as 1 – female, 0 – male) separately. Positive correlation is highlighted with green (paler green $p < .01$, deeper green $p < .001$, two-tailed t-test) and negative correlation with red (paler red $p < .01$, deeper red $p < .001$, two-tailed t-test).

Interpretation

Again, aesthetically pleasing features are **+** with **Ope** and **-** with **Agr**, and to a lesser extent **-** with **Ext**.

The number of dynamic lines (indicative of emotional content) is **-Ope** and **+Agr**.

Image Features - Type

- Default Image
- Is Not Face
- One Face
- Multiple Faces
- No. Faces

Detected using Face++ API

Correlations

Feature	**Demographics**		**Personality Trait**				
Image Type	Gender	Age	Ope	Con	Ext	Agr	Neu
Default Image			-.022		-.043	.015	-.023
Is Not Face	-.072	-.021	.061	-.121	-.108	-.070	.071
One Face	.054	.029	-.016	.102	.081	.046	-.057
Multiple Faces	.040	-.019	-.102	.043	.058	.053	-.032
No. Faces	.072		-.092	.106	.103	.078	-.067

Pearson correlations between profile image and Big Five personality controlled for age and gender and with age and gender (coded as 1 – female, 0 – male) separately. Positive correlation is highlighted with green (paler green $p < .01$, deeper green $p < .001$, two-tailed t-test) and negative correlation with red (paler red $p < .01$, deeper red $p < .001$, two-tailed t-test).

Interpretation

Two ‘clusters’:

1) **Ope** & **Neu**

Not default picture and preferably no face

Multiple faces strongest **- Ope**

No face strongest **- Neu**

2) **Con** & **Ext** & **Agr**

One or more faces: all **+**.

Con strongest correlated with single face and strongest anti-correlated with no face.

Image Features - Demographics

- Age
- Gender
- Race
  - Asian
  - Black
  - White

Detected using Face++ API

Correlations

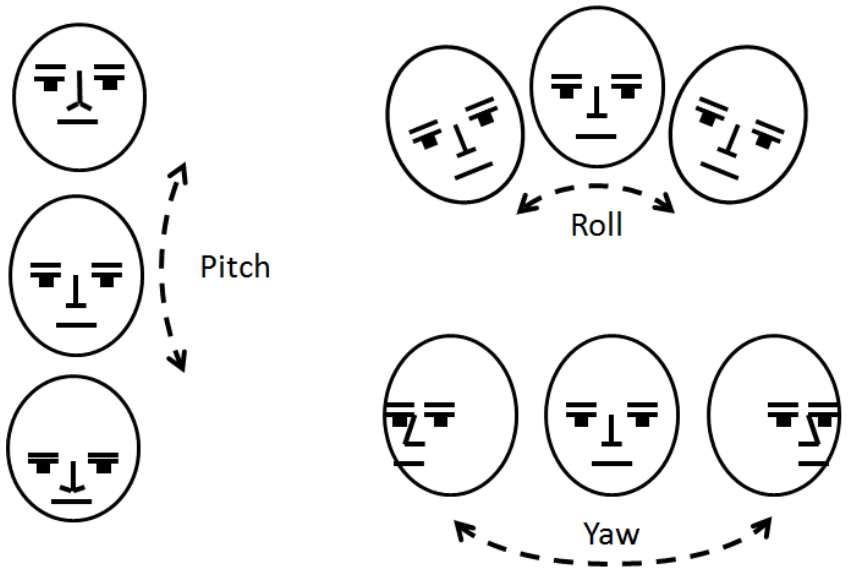
Feature	**Demographics**		**Personality Trait**				
Image Demographics	Gender	Age	Ope	Con	Ext	Agr	Neu
Age	-.310	.306	.050	.105	-.036		
Gender	.795	-.041			.035	.034	
Asian	.064	-.150	-.072	-.042			
Black	-.034	-.061	.047	.050	.085	-.055	-.096
White	-.033	.169	.031		-.066	.026	.071

Pearson correlations between profile image and Big Five personality controlled for age and gender and with age and gender (coded as 1 – female, 0 – male) separately. Positive correlation is highlighted with green (paler green $p < .01$, deeper green $p < .001$, two-tailed t-test) and negative correlation with red (paler red $p < .01$, deeper red $p < .001$, two-tailed t-test).

Image Features - Facial Presentation

- No Glasses
- Reading Glasses
- Sunglasses
- Pitch Angle
- Roll Angle
- Yaw Angle
- Face Ratio

Detected using Face++ API



Yaw – Usually predictive of selfies

Correlations

Feature	**Demographics**		**Personality Trait**				
Facial Presentation	Gender	Age	Ope	Con	Ext	Agr	Neu
No Glasses	.145	-.036		.027	.085	.026	-.065
Reading Glasses	-.141	.054	.020		-.099	-.017	.071
Sunglasses	-.034	-.020	-.017	-.028		-.019	
Pitch Angle	-.043						
Roll Angle	.017						
Yaw Angle							
Face Ratio	.034	.036	.038	-.039	-.097	-.039	.057

Pearson correlations between profile image and Big Five personality controlled for age and gender and with age and gender (coded as 1 – female, 0 – male) separately. Positive correlation is highlighted with green (paler green $p < .01$, deeper green $p < .001$, two-tailed t-test) and negative correlation with red (paler red $p < .01$, deeper red $p < .001$, two-tailed t-test).

Interpretation

Reading Glasses **+ Neu** and **- Ext, Agr**

Sunglasses **- Con**

Face ratio **+ Ope, Neu** and **- Con, Ext, Agr**

Combined with previous findings, **Ope & Neu** prefer no faces in picture, but when a face is present, this occupies a larger part of the photo.

Image Features - Facial Expression

- Smiling
- Anger
- Disgust
- Fear
- Joy
- Sadness
- Surprise
- Left Eye Openness
- Right Eye Openness
- Attention
- Expressiveness
- Neutral Expression4
- Positive Mood
- Negative Mood
- Valence

Smile detected using Face++ API

All other features detected using EmoVu

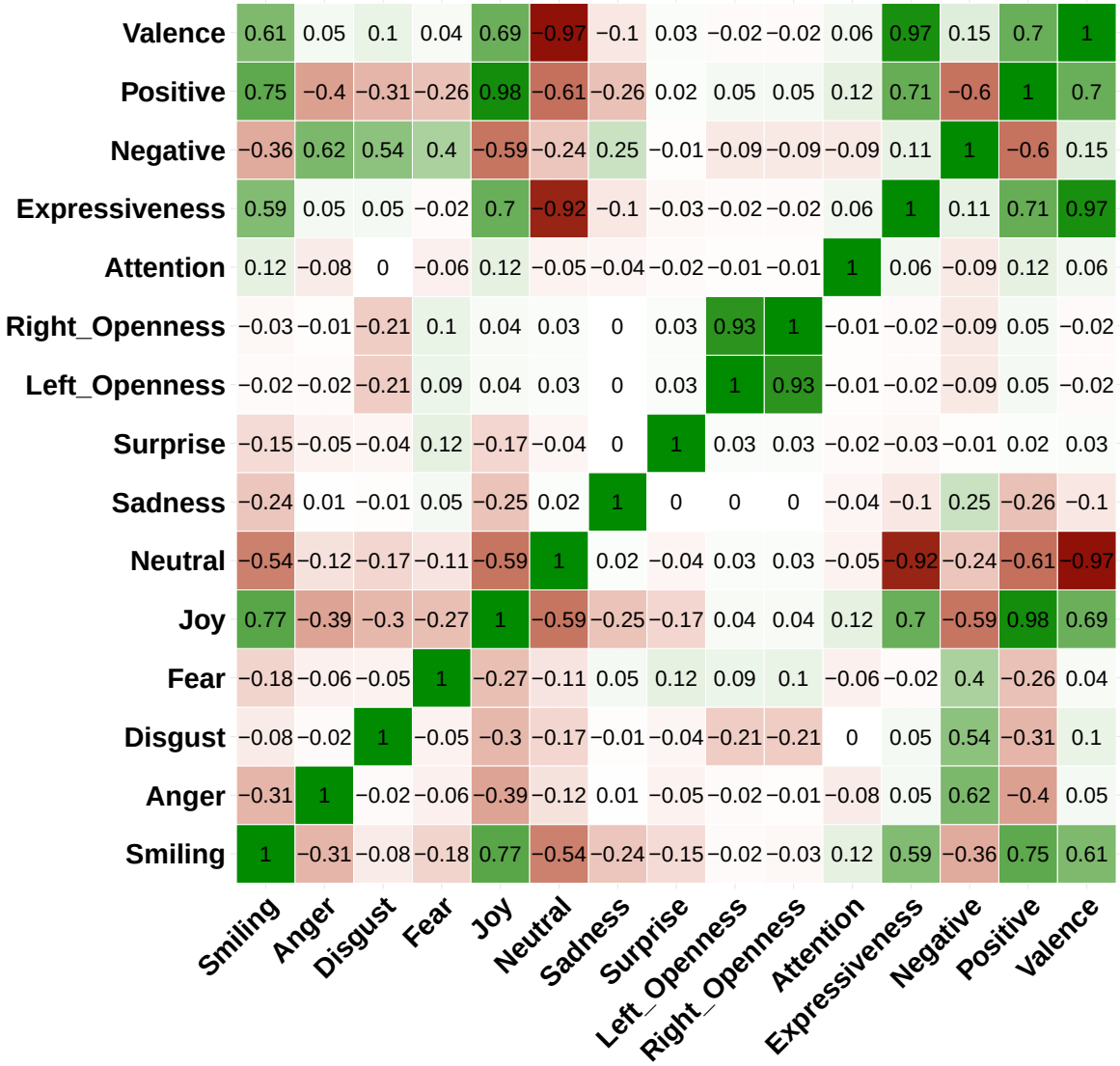
Expressiveness is the highest emotion value

Negative mood is the maximum value of the negative emotions (anger, disgust, fear, sadness)

Positive mood is the maximum value of the positive emotions (joy, surprise)

Valence is the average of positive and negative mood

Facial Expression Intercorrelation



Correlations

Feature	Demographics		Personality Trait				
Facial Expressions	Gender	Age	Ope	Con	Ext	Agr	Neu
Smiling	.229	.141	-.089	.190	.050	.148	-.104
Anger	-.108	-.019	.037	-.080	-.042	-.055	.056
Disgust	-.142	.048					
Fear		-.017	.018	-.029		-.043	.018
Joy	.191	.119	-.093	.180	.061	.140	-.107
Sadness	-.122	-.032	.023	-.051		-.034	.026
Surprise	.038	-.064		-.041		-.031	
Left Eye Openness	.093			.025			
Right Eye Openness	.091			.027			
Attention	-.055	.061	-.047	.049	.018	.040	-.048
Expressiveness	.101	.123	-.072	.140	.054	.106	-.089
Neutral	-.064	-.133	.068	-.128	-.047	-.093	.081
Positive Mood	.198	.111	-.093	.175	.065	.137	-.107
Negative Mood	-.164		.043	-.079	-.029	-.067	.044
Valence	.101	.132	-.075	.140	.053	.105	-.090

Pearson correlations between profile image and Big Five personality controlled for age and gender and with age and gender (coded as 1 – female, 0 – male) separately. Positive correlation is highlighted with green (paler green $p < .01$, deeper green $p < .001$, two-tailed t-test) and negative correlation with red (paler red $p < .01$, deeper red $p < .001$, two-tailed t-test).

Interpretation

Again, two ‘clusters’:

1) **Ope & Neu**

2) **Con & Ext & Agr**

Interpretation

1) **Ope & Neu**

- smiling

Emotions: **+** Anger, **+** Fear, **+** Sadness, **-** Joy

- Positive mood, **+** Negative mood, **-** Valence

- Attention, **-** Expressiveness, **+** Neutral

Interpretation

2) **Con & Ext & Agr**

Almost exact opposite of previous cluster.

Exceptions:

- Surprise cf. **0** Surprise

Eye Openness **+ Con**

Intriguingly, **Con** highest in all 'positive' emotions.

Ext lowest in all 'positive' emotions.

Overview - Openness

- artistic photos
- aesthetically pleasing
- low in color emotions
- less faces, especially more than one
- expressing more negative facial emotions
- less expressive, more neutral

Overview - Neuroticism

- neither artistic or not
- neither aesthetically pleasing or not
- low in color emotions
- less faces
- expressing strongest negative facial emotions
- less expressive, more neutral

Overview - Conscientiousness

- neither artistic or not
- neither aesthetically pleasing or not
- no relation with color emotions
- strongest preference for a single face
- expressing strongest positive facial emotions
- most expressive

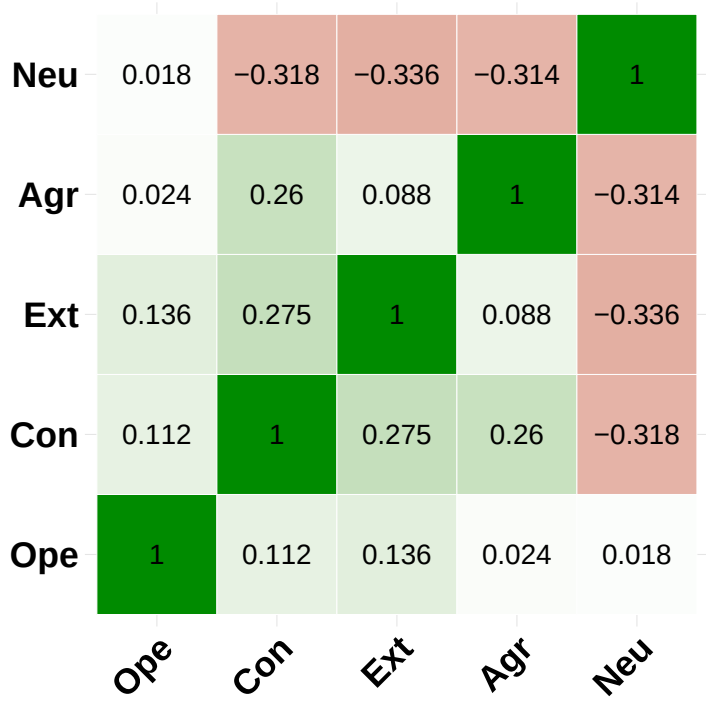
Overview - Agreableness

- photos are not artistic
- photos are not aesthetically pleasing
- most positive color emotions
- prefers faces
- expressing positive facial emotions

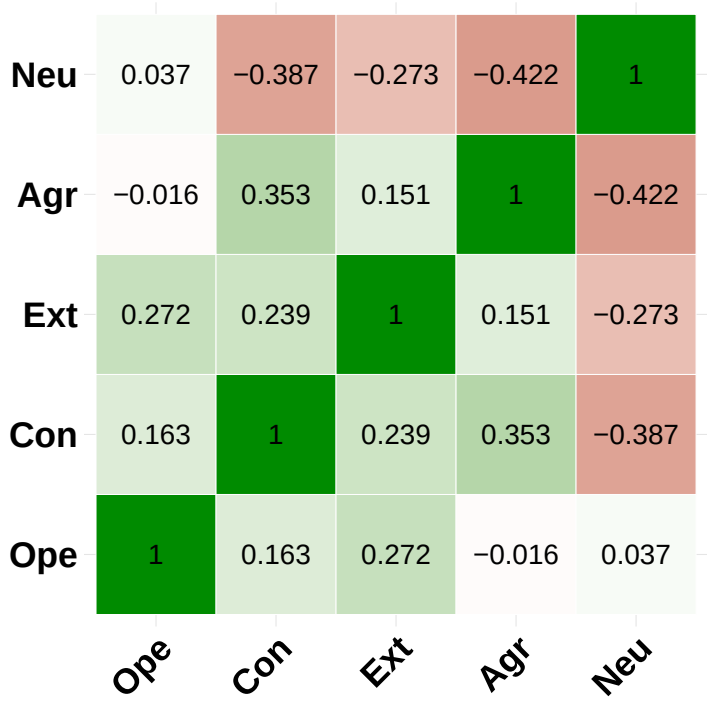
Overview - Extraversion

- photos are not artistic, but less than Agr
- photos are not aesthetically pleasing, but less than Agr
- positive color emotions
- prefers faces, especially multiple faces
- expressing positive facial emotions, less than Agr

Dataset Validation

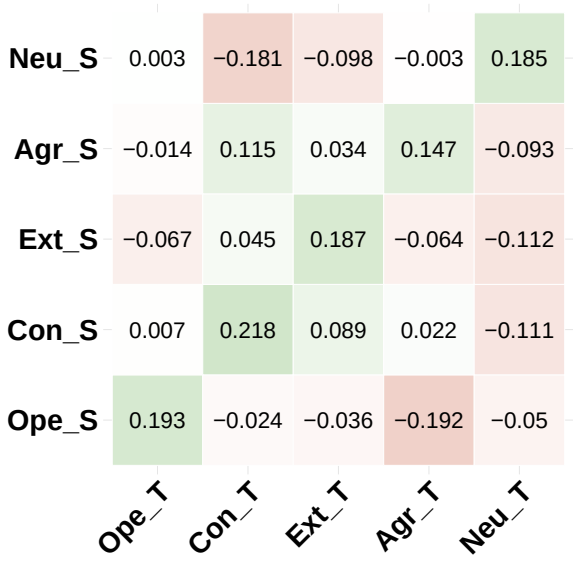


TwitterSurvey Big 5 intercorrelations



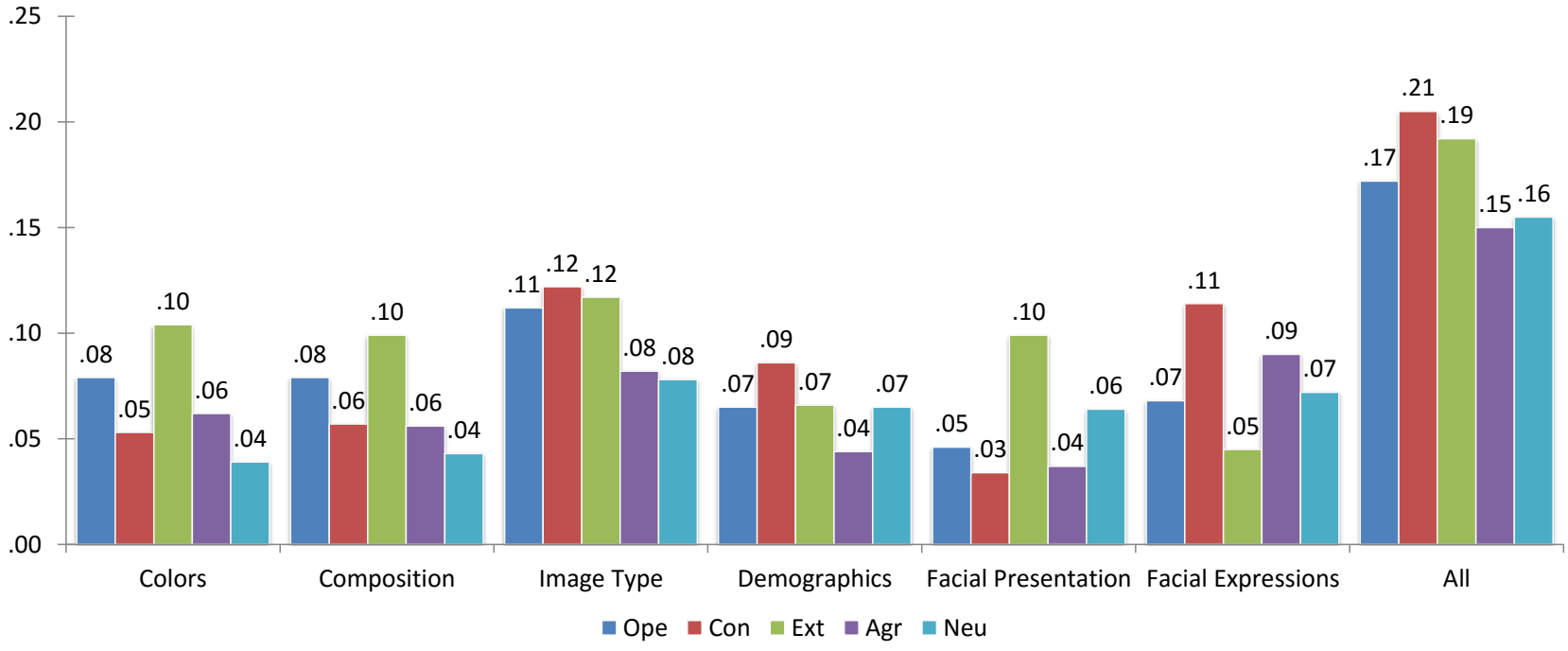
TwitterText Big 5 intercorrelations

Dataset Validation



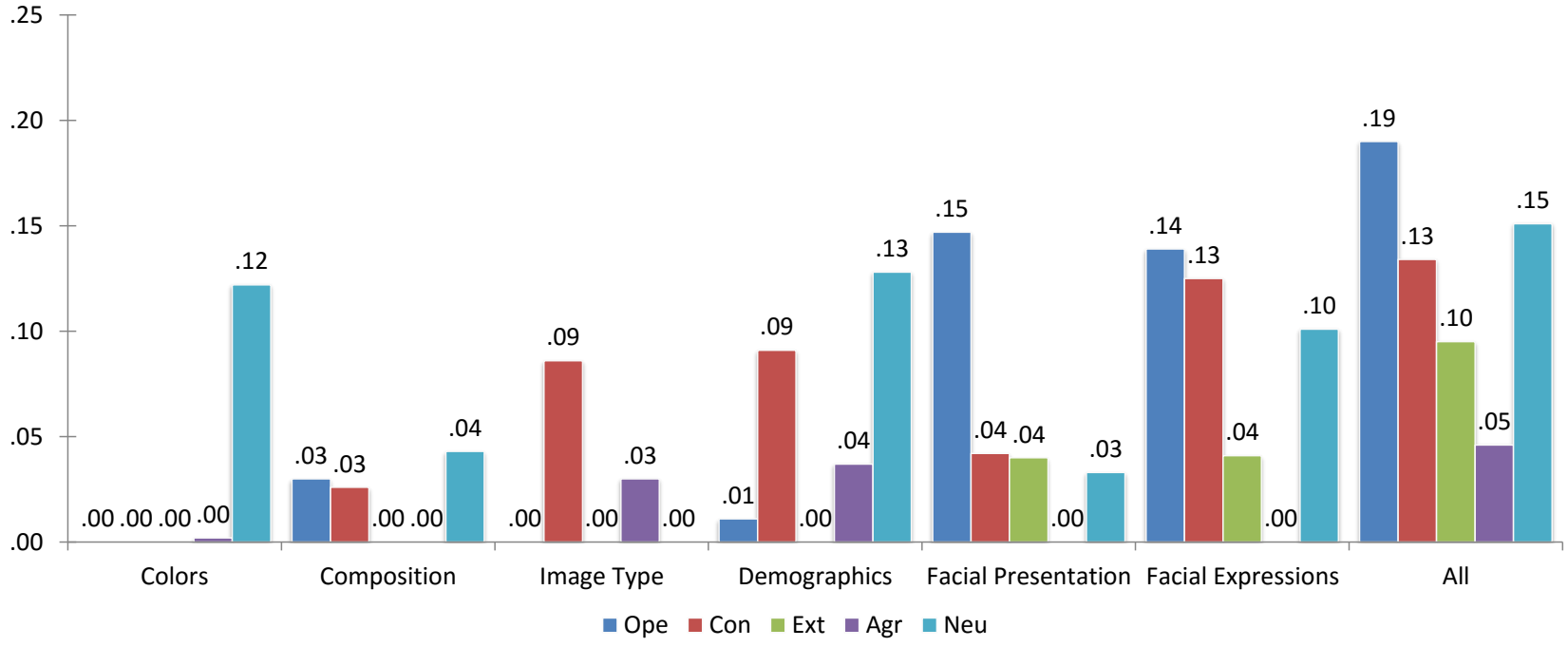
Survey personality & Text-predicted personality correlations between on the TwitterSurvey dataset.

Predictive Performance



TwitterText data set. Predictive performance using Linear Regression, measured in Pearson correlation over 10-fold cross-validation. All correlations are significant ($p < .05$, two-tailed t-test).

Predictive Performance



TwitterSurvey data set. Predictive performance using Linear Regression, measured in Pearson correlation over 10-fold cross-validation. All correlations > 0.95 are significant ($p < .05$, two-tailed t-test).

Take Aways

- Profile picture choice is influenced by personality
- Interpretable computer vision features hold significant prediction accuracy across all personality traits
- Text predicted personality is a good stand-in for survey personality that offers orders of magnitude statistical power

Thank You!

Thank you!

Questions?