computational social media

lecture 2: friending

part 3

daniel gatica-perez
announcements

teams have been created and mentors assigned

assignment #1 will be discussed today

paper #1 will be presented & discussed today

M. Burke, L. Adamic, K. Marciniak
Families on Facebook
ICWSM 2013
this lecture (continues the previous one)

introduction

a human-centric review of research on facebook

1. descriptive analysis of users
2. user motivations
3. user identity
4. the real-name web: privacy & information disclosure
3: identity presentation
how do people present themselves on facebook?
personality & facebook
what is a personality trait?

« stable individual differences in the reactivity of mental mechanisms designed to respond to particular classes of situations »

source: oxford university press
the big-five personality traits

“the Big-Five traits have been broadly accepted as a way of presenting all the major traits of a person at the highest level of abstraction” (Gosling, 2003)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Positive Characteristics</th>
<th>Negative Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>outgoing, enthusiastic</td>
<td>aloof, quiet</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>prone to stress &amp; worry</td>
<td>emotionally stable</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>organized, self-directed</td>
<td>spontaneous, careless</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>trusting, empathetic</td>
<td>uncooperative, hostile</td>
</tr>
<tr>
<td>Openness</td>
<td>creative, imaginative</td>
<td>practical, conventional</td>
</tr>
</tbody>
</table>

(Types & Christal, 1961; Norman, 1963; Goldberg, 1981; Costa & McCrae 1985; Digman, 1990)
the lens model: a basic model for interpersonal perception

Figure 2. The interpersonal circle. Adapted from “A Psychological Taxonomy of Trait-Descriptive Terms: The Interpersonal Domain” by J. S. Wiggins, Journal of Personality and Social Psychology, 1979, Vol. 37, p. 400. Copyright © 1979 by the American Psychological Association. Adapted by permission.

instruments to measure big-five traits: NEO FFI

**NEO FFI - Five-Factor Inventory**
(Costa & McCrae, 1992)
60 questions in total, 12 questions per trait
7-point Likert scale (from “strongly disagree” to strongly agree”)

**Conscientiousness items**
- I am always prepared
- I pay attention to details
- I get chores done right away
- I like order
- I follow a schedule

**Extraversion items**
- I am the life of the party
- I don't mind being the center of attention
- I feel comfortable around people
- I start conversations
- I talk to a lot of people at parties

**Agreeableness items**
- I am interested in people
- I sympathize with others' feelings
- I take time out for others
- I feel others' emotions
- I make people feel at ease

**Openness items**
- I have a rich vocabulary
- I have a vivid imagination
- I have excellent ideas
- I am quick to understand things
- I use difficult words

**Neuroticism items**
- I am easily disturbed
- I change my mood a lot
- I get upset easily
- I have frequent mood swings
- I worry about things

http://en.wikipedia.org/wiki/Big_Five_personality_traits
validity & reliability
## validity

<table>
<thead>
<tr>
<th>Construct Validity</th>
<th>Internal Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>“whether a test measures what it claims to measure” (Brown, 1996)</td>
<td>“whether the experimental procedures were performed correctly” (Salganik, 2018)</td>
</tr>
<tr>
<td>- social constructs: abstract ideas with no unique definition or measurement</td>
<td>- randomization</td>
</tr>
<tr>
<td>- operationalize: capture the construct with observable data (e.g. questionnaire)</td>
<td>- measurement of outcomes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statistical Conclusion Validity</th>
<th>External Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>“whether the statistical analysis of the experiment was done correctly” (Salganik, 2018)</td>
<td>“whether the results of the experiment can be generalized to other situations” (Salganik, 2018)</td>
</tr>
<tr>
<td>- select correct statistical tests</td>
<td>- different or larger populations</td>
</tr>
<tr>
<td>- measure effect sizes correctly</td>
<td>- lab vs. real-life</td>
</tr>
</tbody>
</table>

J. D. Brown, Testing in language programs. Prentice Hall, 1996  
reliability: extent to which measurements can be replicated
ratio: true variance / (true variance + error variance)

## Different Types of Reliability

<table>
<thead>
<tr>
<th>Types</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interrater reliability</td>
<td>It reflects the variation between 2 or more raters who measure the same group of subjects.</td>
</tr>
<tr>
<td>Test-retest reliability</td>
<td>It reflects the variation in measurements taken by an instrument on the same subject under the same conditions. It is generally indicative of reliability in situations when raters are not involved, such as self-report survey instrument.</td>
</tr>
<tr>
<td>Intrarater reliability</td>
<td>It reflects the variation of data measured by 1 rater across 2 or more trials.</td>
</tr>
</tbody>
</table>
interrater reliability illustrated

**interrater reliability:** variation between two or more raters who measure (or rate) the same group of targets

targets $i=1,\ldots,I$
raters $j=1,\ldots,J$

$y_i^j$: rating of target $i$ by rater $j$
ICC: intraclass correlation coefficient (Shrout & Fleiss ’79):

**ICC(1,1), ICC(1,k):** agreement when each target is rated by a different set of k raters randomly selected from a larger population of raters

**ICC(2,1), ICC(2,k):** agreement when a sample of k raters is randomly selected from a larger population, and each rater rates each target

**ICC(1,1), ICC(2,1):** expected reliability of a single rater’s ratings

**ICC(1,k), ICC(2,k):** reliability when ratings are aggregated over the k raters to obtain a mean rating

what is good ICC reliability?

<table>
<thead>
<tr>
<th>estimated ICC</th>
<th>low endpoint of 95% confidence interval of estimated ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.40</td>
<td>— poor</td>
</tr>
<tr>
<td>[0.40, 0.60)</td>
<td>— fair</td>
</tr>
<tr>
<td>[0.60, 0.75)</td>
<td>— good</td>
</tr>
<tr>
<td>[0.75, 1.00]</td>
<td>— excellent</td>
</tr>
</tbody>
</table>

< 0.50 — poor
[0.50, 0.75) — fair
[0.75, 0.90) — good
[0.90, 1.00] — excellent


T. Koo & M. Li. A guideline of selecting and reporting intraclass correlation coefficients for reliability research, JCM 15 (2), 2016
back to personality & facebook
facebook profiles

My Real Name

opinions  demographics  visual content

social behavior  social relations
do facebook profiles convey accurate impressions?
(Back et al., 2010)

NO
idealized virtual-identity hypothesis
+ people display idealized characteristics
+ impressions reflect ideal-self rather than real-self

YES
extended real-life hypothesis
+ people communicate real personality
+ impressions reflect real-self views

why?
+ accountability with friends
+ ideal-self difficult to control


Photo credit: Dunk @ flickr (cc) http://www.flickr.com/photos/dullhunk/3914761467/
measuring big-five traits: TIPI

TIPI - Ten-Item Personality Instrument (Gosling et al., 2003)
10 questions, 2 per trait
7-point Likert scale

Validity: correlation with BFI (44-item questionnaire):
[0.65 - 0.87]
N=1813

Test-retest reliability: correlation across sessions (6 weeks apart):
E(0.77); C(0.76), O(0.62)
mean 0.72
N=180

facebook profiles, actual personality, self-idealization

133 FB college students (USA)
103 StudiVZ users (DE)

**actual personality**
+ self-reports & four close friends
+ TIPI+NEO (US), BFI-10 (DE)

**ideal-self personality**
+ self-report: “describe yourself as you ideally would like to be”

**personality impressions**
+ 9-10 external annotators
+ TIPI (US), BFI-10 (DE)

**analysis**
1. reliability of impressions (ICC)
2. correlation analysis (r)
+ actual personality vs. impressions
+ self-ideal personality vs. impressions
results

Facebook profiles have a weak tendency to convey actual personality

<table>
<thead>
<tr>
<th>Observer rating</th>
<th>Impressions ICC (consensus)</th>
<th>Actual personality r (accuracy)</th>
<th>Ideal self r</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC(2,k)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td>.81***</td>
<td>.39***</td>
<td>.13</td>
</tr>
<tr>
<td>Average observer</td>
<td>.31***</td>
<td></td>
<td>.08*</td>
</tr>
<tr>
<td>Single observer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>.59***</td>
<td>.22**</td>
<td>.16</td>
</tr>
<tr>
<td>Average observer</td>
<td></td>
<td></td>
<td>.08*</td>
</tr>
<tr>
<td>Single observer</td>
<td>.13***</td>
<td>.11**</td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.77***</td>
<td>.27**</td>
<td>.05</td>
</tr>
<tr>
<td>Average observer</td>
<td></td>
<td></td>
<td>.03</td>
</tr>
<tr>
<td>Single observer</td>
<td>.27***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>.48***</td>
<td>.13</td>
<td>.12</td>
</tr>
<tr>
<td>Average observer</td>
<td></td>
<td></td>
<td>.04</td>
</tr>
<tr>
<td>Single observer</td>
<td>.09***</td>
<td>.06</td>
<td></td>
</tr>
<tr>
<td>Openness</td>
<td>.72***</td>
<td>.41***</td>
<td>.24**</td>
</tr>
<tr>
<td>Average observer</td>
<td></td>
<td></td>
<td>.24***</td>
</tr>
<tr>
<td>Single observer</td>
<td>.23***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* \( p_{rep} > .95 \)
** \( p_{rep} > .99 \)
*** \( p_{rep} > .999 \)

Limitations:
- + college students
- + small sample (N=236)

Welcome to the myPersonality Project Website

This wiki was established to share the data with researchers. Go HERE if you want to take a psychological test, or visit 🐟 Apply Magic Sauce, to predict your personality from your Facebook Likes.

News

Introduction

myPersonality was a popular Facebook application that allowed users to take real psychometric tests, and allowed us to record (with consent!) their psychological and Facebook profiles. Currently, our database contains more than 6,000,000 test results, together with more than 4,000,000 individual Facebook profiles. Our respondents come from various age groups, backgrounds, and cultures. They are highly motivated to answer honestly and carefully, as the only gratification that they receive for their participation is feedback on their results.

A wide variety of data is available to the registered collaborators, including:

- Psychometric tests' scores
- Records of users' Facebook profiles
- Item-level data

original webpage (no longer available)
http://mypersonality.org/wiki/doku.php
more information at:
https://sites.google.com/michalkosinski.com/mypersonality
Private traits and attributes are predictable from digital records of human behavior

Michal Kosinski\textsuperscript{a,1}, David Stillwell\textsuperscript{a}, and Thore Graepel\textsuperscript{b}

\textsuperscript{a}Free School Lane, The Psychometrics Centre, University of Cambridge, Cambridge CB2 3RQ United Kingdom; and \textsuperscript{b}Microsoft Research, Cambridge CB1 2FB, United Kingdom

Edited by Kenneth Wachter, University of California, Berkeley, CA, and approved February 12, 2013 (received for review October 29, 2012)
identity presentation in facebook
links between social media and personality psychology
personality is expressed and perceived online
the big-five model captures the major traits of a person
validity & reliability: key properties of a personality instrument
facebook profiles tend to reflect actual personality
personal traits can be disclosed through facebook activity
questions?

daniel.gatica-perez@epfl.ch