Promises and Challenges of Big Data in Personality

Emorie D. Beck

Washington University in St. Louis

Questions	Aggregation
Description	Between-Person
Prediction	Within-Person
Explanation	Idiographic

Questions

Description

Prediction

Explanation

Aggregation

Between-Person

Within-Person

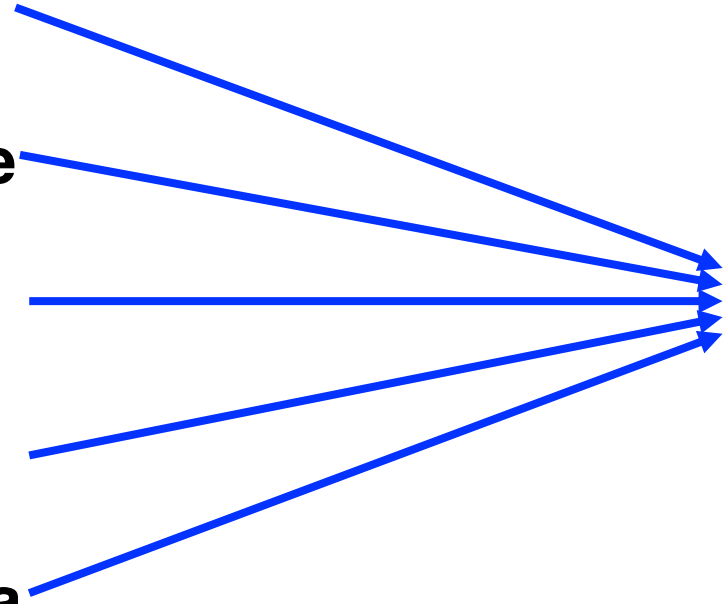
Idiographic

Prediction

Between-Person

Predicting Personality from:

- Facebook Likes
- Twitter Language
- Email
- Profile Photos
- Smartphone Data
- etc.



The Big Five

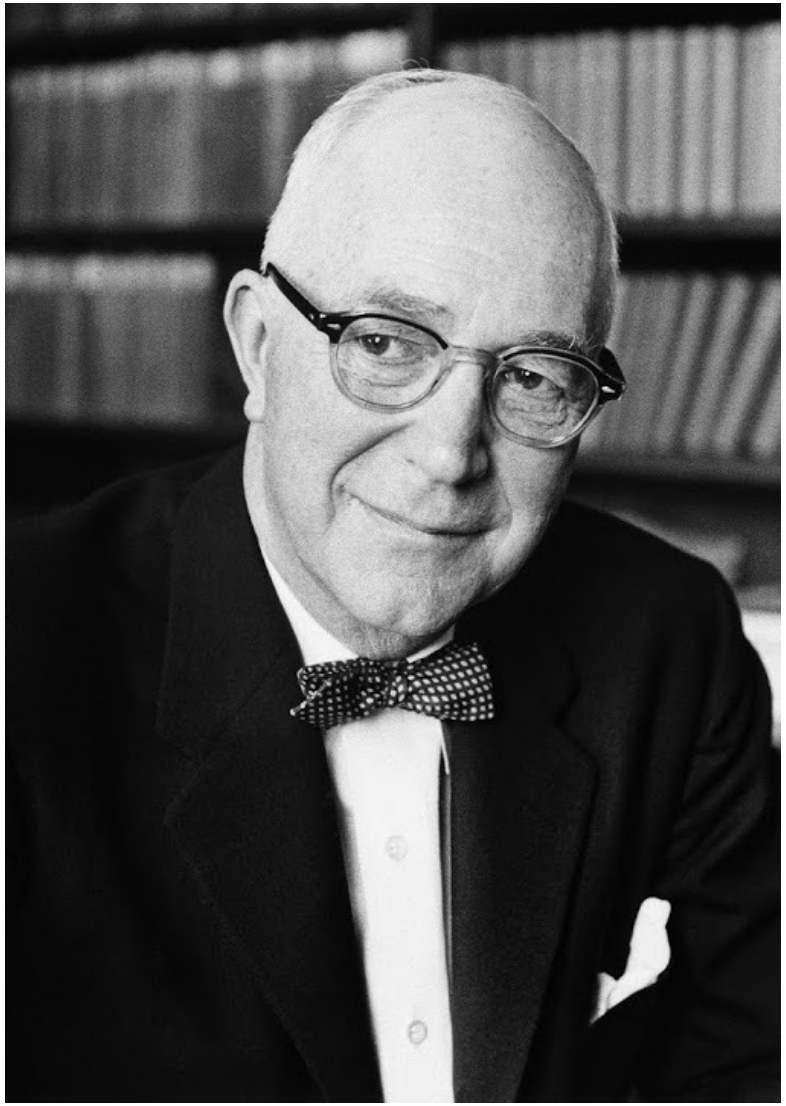
E A C N O

Sources of Big Data

- ❑ Text
- ❑ Clicks / Engagement
- ❑ Web History
- ❑ Likes
- ❑ Photos
- ❑ Videos
- ❑ Audio
- ❑ Smartphone Usage
- ❑ Location Data
- ❑ Social Media Connections
- ❑ Etc.

Machine Learning Tools

- ✓ Classification
- ✓ Regularization
- ✓ Cross-Validation (k-fold and loo)
- ✓ Random Forest
- ✓ Neural Networks (and deep learning)
- ✓ Support Vector Machine (SVM)
- ✓ Ensemble Methods
- ✓ Clustering (k-means, nearest neighbor)



Galloping Empiricism

Galloping empiricism, which is our present occupational disease, dashes forth like a headless horseman. It has no rational objective; uses no rational method other than mathematical; reaches no rational conclusion. It lets the discordant data sing for themselves.

Heuristic Realism

By contrast heuristic realism says "While we are willing to rest our case for traits on empirical evidence, the area we carve out for study should be rationally conceived, tested by rational methods, and the findings should be rationally interpreted.

Allport, 1968, p. 50

Galloping Empiricism

- ☒ A priori atheoretical
- ☒ Poorly defined
- ☒ Little contribution to theory building
- ☒ Difficult to interpret

Heuristic Realism

- ☑ A priori theoretical
- ☑ Carefully and rationally defined
- ☑ Incremental theory building
- ☑ Built for interpretability

Questions	Aggregation
Description	Between-Person
Prediction	Within-Person
Explanation	Idiographic

Questions

Description

Prediction

Explanation

Aggregation

Between-Person

Within-Person

Idiographic

Questions	Aggregation
Description	Between-Person
Prediction	Within-Person
Explanation	Idiographic

Description

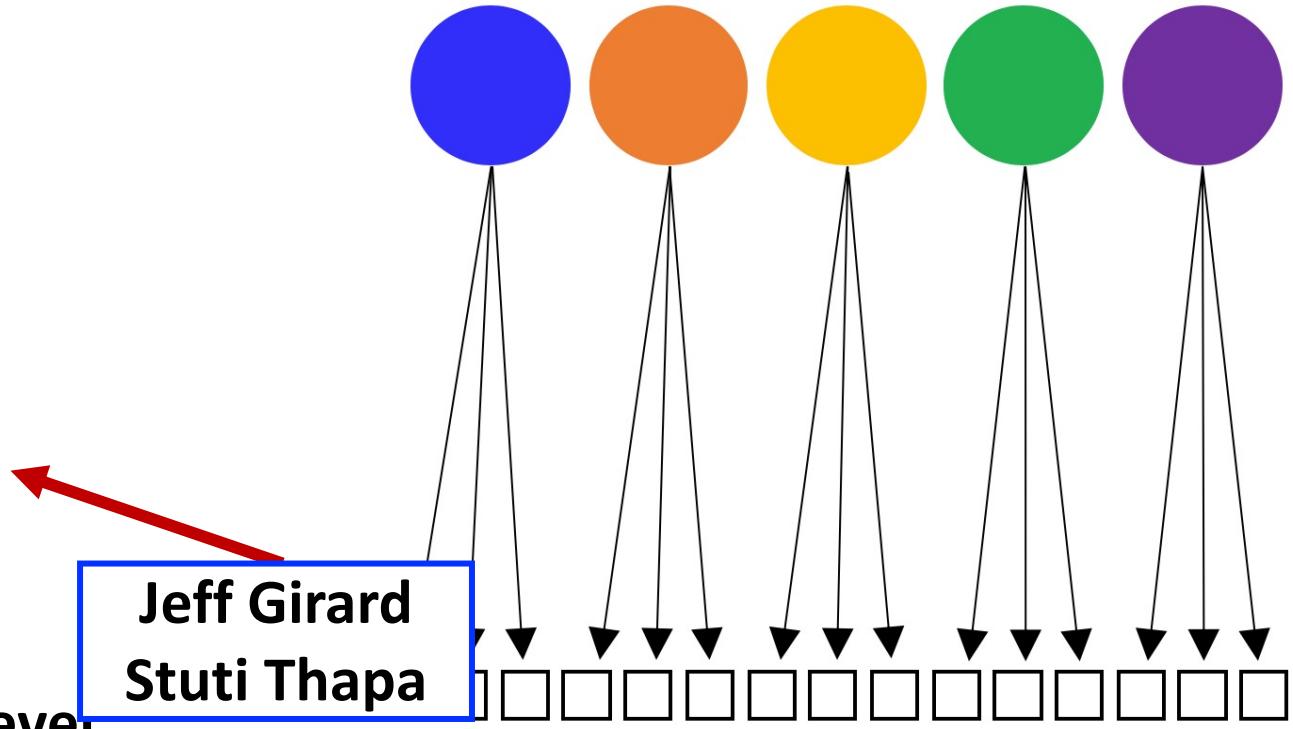
Between-Person

Structure / Constructs: correlations, data reduction (EFA, PCA)

Descriptives: mean, standard deviation

Differences: cross-cultural, age differences, gender differences, etc.

Change: longitudinal mean-level, rank-order, and structural change



Jeff Girard
Stuti Thapa

Description

Within-Person

Structure / Constructs: correlations, data reduction (EFA, PCA)

Descriptives: mean, standard deviation, instability

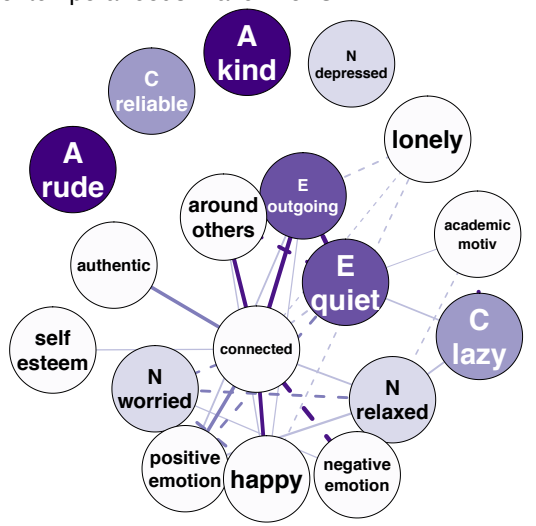
Stuti Thapa

Differences: Cross-situational differences

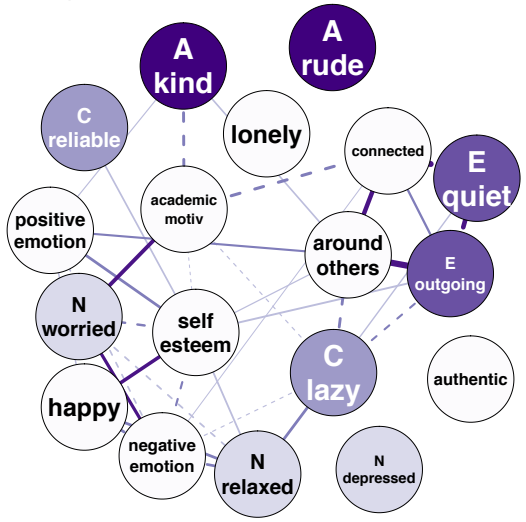
Alex Danvers

Change: longitudinal mean-level change, structural change, and ipsative consistency

Contemporaneous Wave 1 for S1



Contemporaneous Wave 1 for S2



Prediction

Within-Person

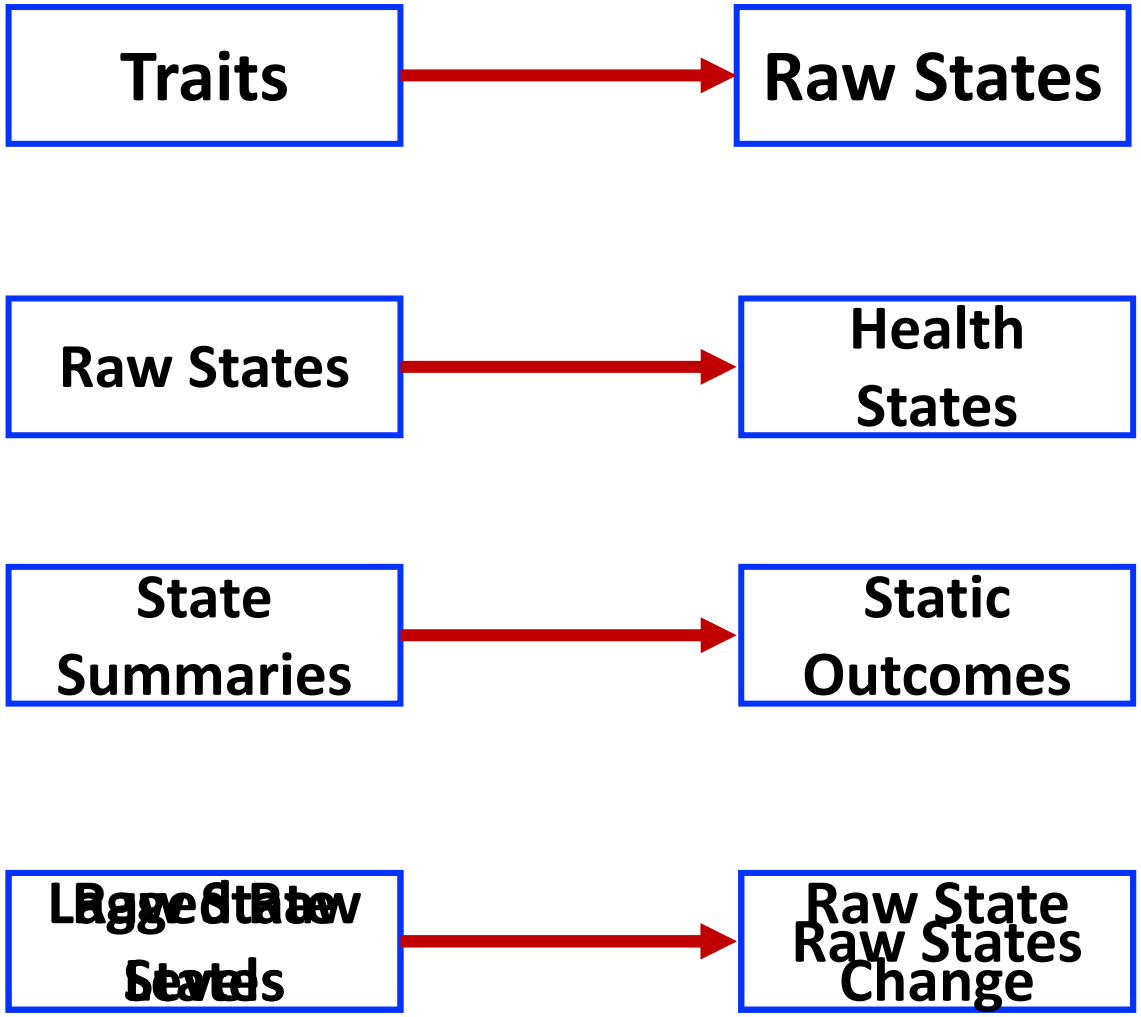
Predicting personality states: Level, moving average, etc.

Predicting Outcomes Within-Person: State-level outcomes, e.g. health, emotions, etc.

Predicting Between-Person Outcomes: Within-person descriptives → between person outcomes

Predicting Change: Lagged correlations, differential equations

Alex Danvers



Sources of Big Data

- ❑ Text
- ❑ Clicks / Engagement
- ❑ Web History
- ❑ Likes
- ❑ Photos
- ❑ Videos
- ❑ Audio
- ❑ Smartphone Usage
- ❑ Location Data
- ❑ Social Media Connections
- ❑ Etc.

Machine Learning Tools

- ✓ Classification
- ✓ Regularization
- ✓ Cross-Validation (k-fold and loo)
- ✓ Random Forest
- ✓ Neural Networks (and deep learning)
- ✓ Support Vector Machine (SVM)
- ✓ Ensemble Methods
- ✓ Clustering (k-means, nearest neighbor)

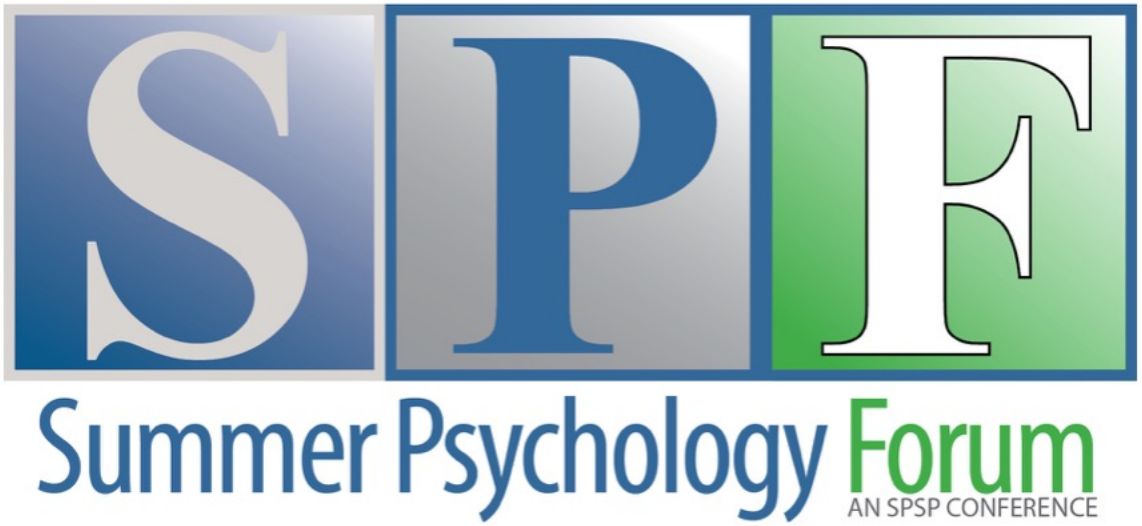
Galloping Empiricism

- ☒ A priori atheoretical
- ☒ Poorly defined
- ☒ Little contribution to theory building
- ☒ Difficult to interpret

Heuristic Realism

- ☑ A priori theoretical
- ☑ Carefully and rationally defined
- ☑ Incremental theory building
- ☑ Built for interpretability

Questions	Aggregation
Description	Between-Person
Prediction	Within-Person
Explanation	Idiographic



Modeling and Analyzing the Dynamics of Motivation, Affect, and Social Interaction

July 8 - 10, 2020

The Marriott City Center Minneapolis
Minneapolis, MN

Use code **20SPF25**
for $25 off registration by March 16

Speakers and Workshop Leaders:

Robin Vallacher
Niall Bolger
Jonas Haselbek
Aaron Fisher
Tom Rodebaugh
Jeff Girard
Aaron Likens
Patrick Doyle
Alex Danvers
Brian Eiler
Steve Read
Emorie Beck