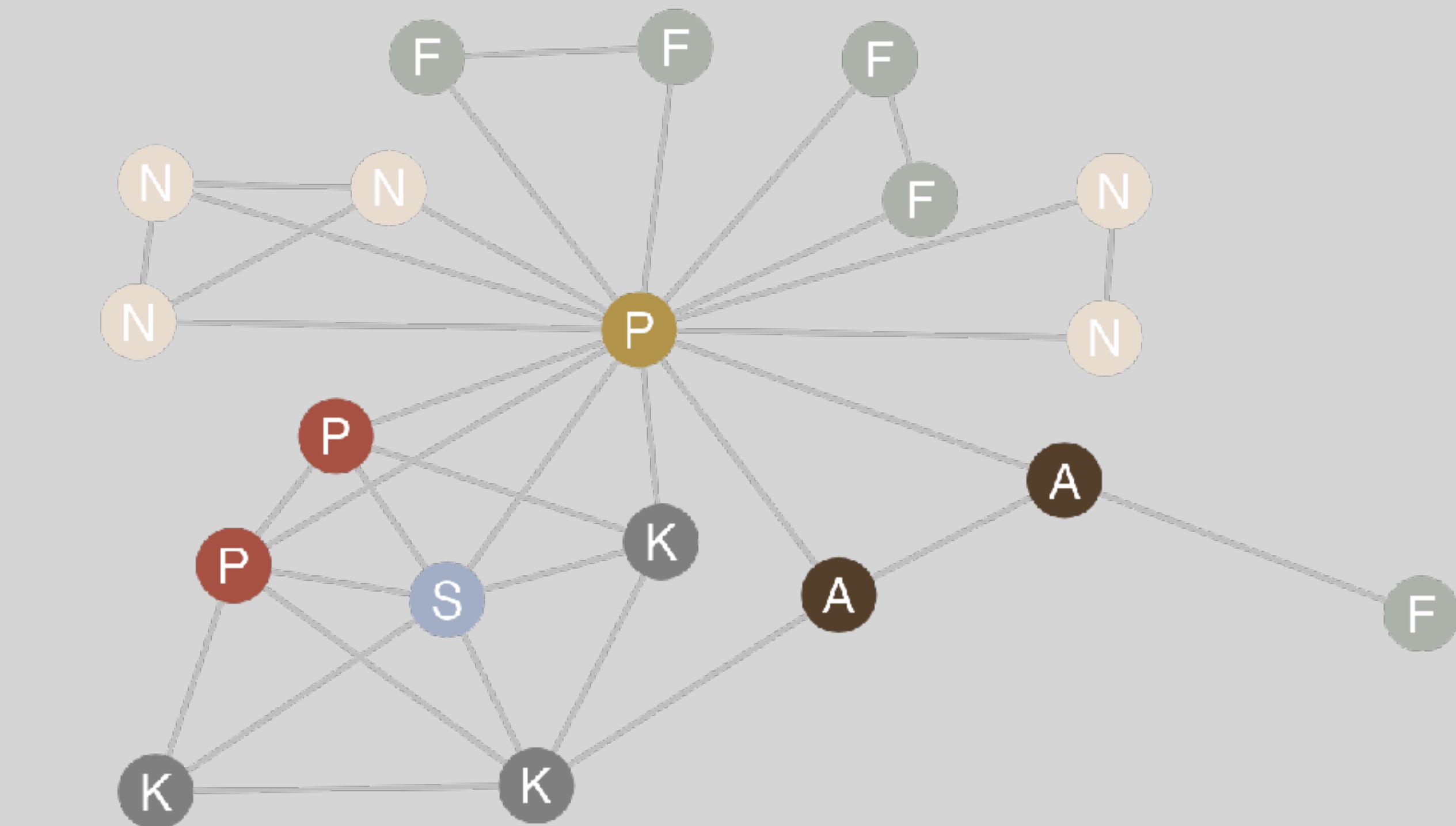
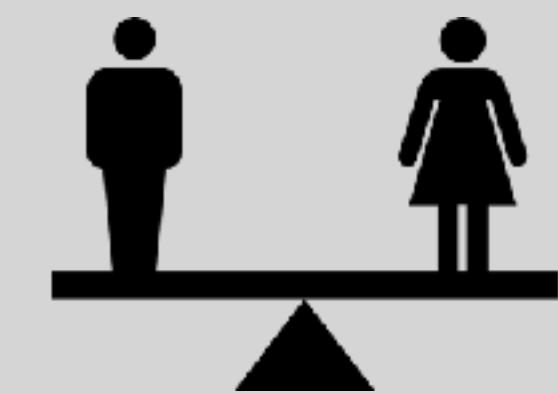
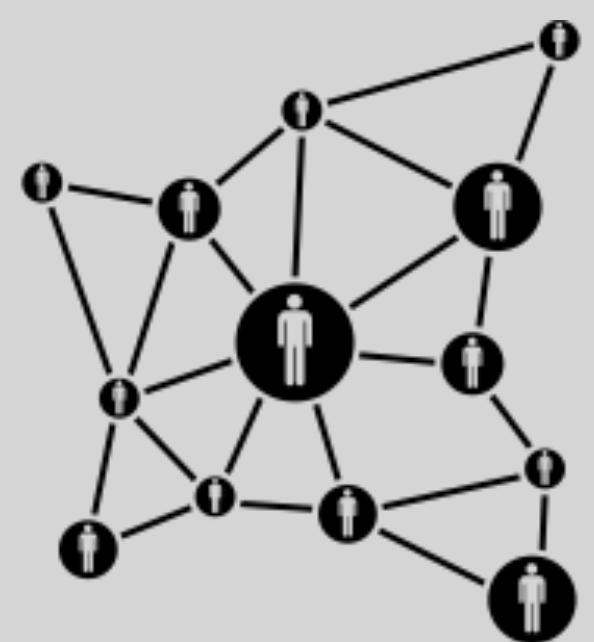


# Predicting fertility outcomes with networks





variables  
explain  
little

Fewer  
births  
because of  
study and  
flexwork?



“ total effect on fertility ...  
rather small

# incomparable results

Downloaded from <http://rsta.royalsocietypublishing.org>



1

1

10

10

10

10

10

1

1

University D  
70-70346-004

#### **Summary of Population**



# surprising patterns

# non-replicable results

Population Review

Volume 48, Number 2, 2021

Type: Article, pp. 346–364

## Explaining the Association of Education and Occupation with Childlessness: The Role of Desires and Expectations in Remote Childless

Antonella Renna, Gert J. J. van den Berg, Armando C. Mills  
 Authors' affiliations: Department of Public Administration and Sociology, Erasmus University Rotterdam, Burgemeester Oudhuisweg 50, 3060 DR, Rotterdam, The Netherlands; and Department of Sociology and CES, U.S. University of Groningen, Postbus 800, 9700 JE, Groningen, The Netherlands (Verma); U.S. University of Groningen, Gronenborghaven 19, 9742 EG, Groningen, The Netherlands (Bijlsma); Department of Epidemiology, University Medical Center Groningen, Postbus 30.000, 9700 RB, Groningen, The Netherlands (Mills); Department of Sociology and Radford College, University of Oxford, 42 Park End Street, Oxford OX1 1LB, UK (Mills)

Corresponding author: Antonella Renna, Konstanz, Germany; email: [antonella@uni-konstanz.de](mailto:antonella@uni-konstanz.de)

### Abstract

Although there are well-established relationships between women's higher education, women's labor participation (LFP), and occupation as one tool and childlessness as the other hand, the underlying reasons and the role that childlessness desire and expectations play remain unclear. We use the National Longitudinal Survey of Youth in the United States (NLSY72) women and apply both logistic regression models to examine the role of childlessness desire early in life, and multilevel models for repeated measures to examine the role of childlessness expectations throughout the life course. We find that higher educational attainment and LFP are positively associated with childlessness. We do not find, however, that higher educated and working women more often desire or expect to remain childless. In contrast, we find that among women who already remain childless, those women who work full-time and have higher status occupations have higher expectations to have children throughout their life course. These results suggest that education and occupation produce consequences, resulting in the propensity of childbearing childless women to remain childless. So desire and expectations, when many working women remain childless despite the desire and expectation to become a mother, are findings from the importance of "willingness to remain." It facilitates highlighting the importance of increasing public awareness regarding the domestic boundary with age.

### Keywords

Childlessness, education, occupation, fertility desires, longitudinal research

**Authors' contributions:** ARM receives funding from the EU grants H2020-ES-2014-2015-2016 and H2020-ES-2017-2018 and The Leiden-Tata, Larchmont Center for Demographic Research. AR was supported by the Netherlands Organization for Scientific Research (NWO) grant number 452.05.030.

© 2021 Institute of Demography Wien

346

## Childless Expectations and Childless Life Course

Anna Rydell and S. Philip Morgan, University of North Carolina at Chapel Hill

Using a nationally representative panel of the 1972 National Longitudinal Survey, we construct life-lines characterizing women's fertility behavior. One-quarter of women in the NLSY expect no children but only 14.8 percent of childless women follow two predominant life course patterns of childbearing and the subsequent adoption of a life stage or (2) introduce about parenthood delayed through later expectations across various ages. We also find women become a mother after considering childlessness research on childlessness and childbearing preferences. It is problematic to assign expected and unexpected reproductive experiences of childless women. In addition, childless women strongly consider permanent life stage rather than temporary life stage. These childless expectancies indicate a strong effect of aging on the increasingly selective group of childless women. Women's childlessness is primarily associated with later with reports of a childless expectation. We thus argue that reproductive and situational contexts on childless women predominantly through reported childbearing preferences.

## Introduction

Permanent and temporary childlessness in industrialized (Dye 2013; Rawland 2007; Sobek 2017). In the short run, one in seven women aged 40–44 have not almost a half of women aged 15–29 were childless

(Contraceptive Use in 16–49-Year-Old Women in 2016). We acknowledge support provided by Carolina Population Center, University of North Carolina at Chapel Hill, which receives funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P30-HD036290) and (T15-HD041576). Direct all correspondence to Anna Rydell, Department of Sociology, University of North Carolina at Chapel Hill, 117 Hanes Hall, Chapel Hill, NC 27599, USA; email: [arydell@email.unc.edu](mailto:arydell@email.unc.edu).

The authors thank reviewers for comments and useful insights. This paper was presented at the Population Association Annual Meeting in 2016. We acknowledge the support provided by Carolina Population Center, University of North Carolina at Chapel Hill, which receives funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P30-HD036290) and (T15-HD041576). Direct all correspondence to Anna Rydell, Department of Sociology, University of North Carolina at Chapel Hill, 117 Hanes Hall, Chapel Hill, NC 27599, USA; email: [arydell@email.unc.edu](mailto:arydell@email.unc.edu).

Received: 10/07/2019; revised: 07/01/2020; accepted: 07/01/2020  
 © Springer Nature Switzerland AG 2021

# Replication Crisis

PSYCHOLOGY

## Estimating the reproducibility of psychological science

Open Science Collaboration\*

ROYAL SOCIETY  
OPEN SCIENCE

[rsos.royalsocietypublishing.org](http://rsos.royalsocietypublishing.org)

Research



CrossMark  
click for updates

Cite this article: Smaldino PE, McElreath R.  
2016 The natural selection of bad science.

### The natural selection of bad science

Paul E. Smaldino<sup>1</sup> and Richard McElreath<sup>2</sup>

<sup>1</sup>Cognitive and Information Sciences, University of California, Merced, CA 95343, USA

<sup>2</sup>Department of Human Behavior, Ecology, and Culture, Max Planck Institute for  
Evolutionary Anthropology, Leipzig, Germany

PES, 0000-0002-7133-5620; RME, 0000



General Article

### False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Joseph P. Simmons<sup>1</sup>, Leif D. Nelson<sup>2</sup>, and Uri Simonsohn<sup>1</sup>

<sup>1</sup>The Wharton School, University of Pennsylvania, and <sup>2</sup>Haas School of Business, University of California, Berkeley

**aps**  
ASSOCIATION FOR  
PSYCHOLOGICAL SCIENCE

Psychological Science  
22(11) 1359–1366  
© The Author(s) 2011  
Reprints and permission:  
[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)  
DOI: 10.1177/0956797611417632  
<http://pss.sagepub.com>  
SAGE

# Crisis in Family Sociology



## Reasons unlikely

- ✓ Strong methods
- ✓ Strong focus on representative data
- ✓ Less measurement error
- ✓ Open data
- ✓ Large N
- ✓ Often descriptive



## Reasons not unlikely

- ✗ Non-experimental
- ✗ Correlational, but little causal inference
- ✗ Large N, yet star gazing
- ✗ Controlling at will
- ✗ Long reign linearity

# Overcoming the Crisis



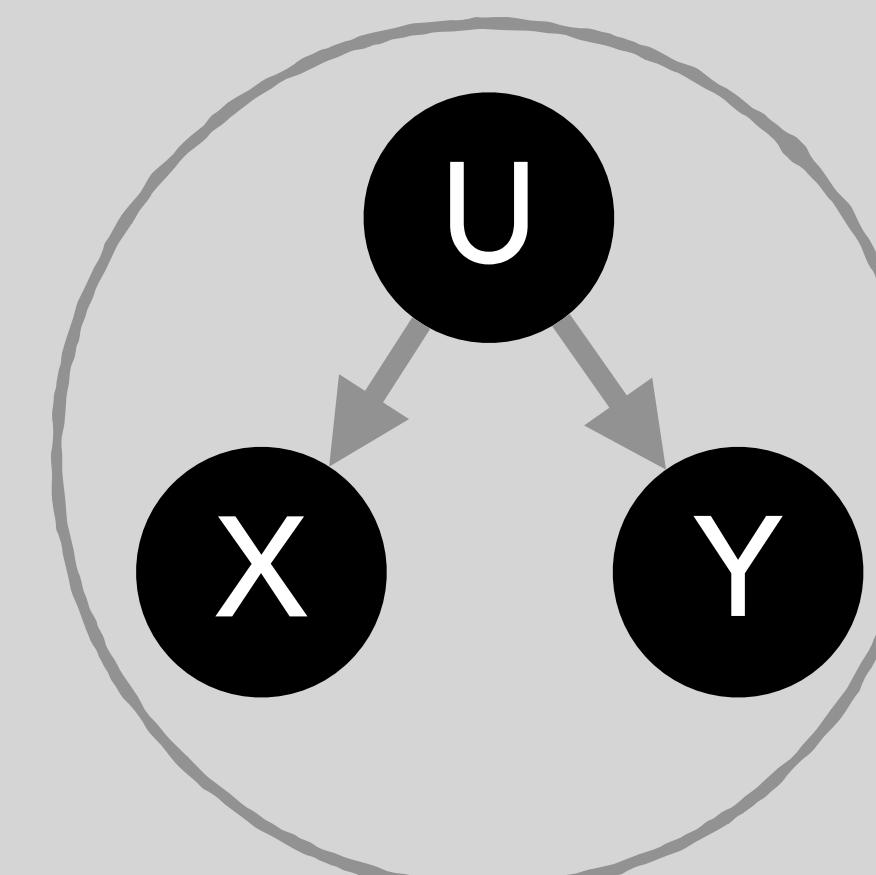
theory



measurement



incentives



causal inference



prediction

a shift towards **prediction**  
leads to a more reliable  
and useful social science

out-of-sample predictive ability:



clear measure of  
effect size



# out-of-sample predictive ability

- ✓ is easy(ier) to understand
- ✓ can be compared across analytical techniques
- ✓ can be compared across models
- ✓ is less gameable

European Sociological Review, VOLUME 24 | NUMBER 1 | 2018 63–82  
DOI: 10.1111/essr.12065, available online at [www.wileyonlinelibrary.com](http://www.wileyonlinelibrary.com)  
online publication 9 March 2018

## Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It

Carina Mood

Logistic regression estimates do not behave like linear regression estimates in an important respect. They are affected by omitted variables unrelated to the independent variables in the model. This has been largely unnoticed by sociologists. I present how to interpret log-odds ratios or odds ratios as effect measures, the degree of unobserved heterogeneity in the model, and log-odds ratios or odds ratios for similar models across groups or across models with different independent variables in these problems and possible ways of overcoming them.

### Introduction

The use of logistic regression is routine in the social sciences when studying outcomes that are naturally or necessarily represented by binary variables. Examples are many in stratification research (education/transitions, promotion), demographic research (divorce, childbirth, non-leaving), social medicine (diagnosis, mortality), research into social exclusion (unemployment, benefit take up), and research about political behavior (voting, participation in collective action). When fitting a dichotomous dependent variable, sociologists almost automatically turn to logistic regression, and this practice is generally recommended in textbooks in quantitative methodology. However, our common ways of interpreting results from logistic regression have some important problems.<sup>1</sup>

The problems stem from unobservables, or the fact that we can seldom include in a model all variables that affect an outcome. Unobserved heterogeneity is

© The Author 2018. Published by Oxford University Press. All rights reserved.  
<https://doi.org/10.1093/essr/essr12065>



Annual Review of Sociology

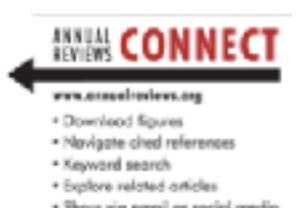
Interpreting and Understanding Logits, Probits, and Other Nonlinear Probability Models

Richard Breen,<sup>1</sup> Kristian Bernt Karlson,<sup>2</sup> and Anders Holm<sup>3</sup>

<sup>1</sup>Faculty of College and Department of Sociology, University of Oxford, OX1 1NP, United Kingdom; email: richard.breen@ox.ac.uk

<sup>2</sup>Department of Sociology, University of Copenhagen, DK-1333 Copenhagen, Denmark

<sup>3</sup>Department of Sociology, University of Western Ontario, London, Ontario N6A 3G2, Canada



### Keywords

logit, probit, KLIM method, F-standardization, marginal effects, linear probability model, mediation

### Abstract

Methods textbooks in sociology and other social sciences routinely recommend the use of the logit/probit model when an outcome variable is binary, an ordered logit or ordered probit when it is ordinal, and a multinomial logit when it has more than two categories. But these methodological guidelines take little or no account of a body of work that, over the past 10 years, has pointed to problematic aspects of these nonlinear probability models and, particularly, to difficulties in interpreting their parameters. In this review, we draw on that literature to explain the problems, show how they manifest themselves in research, discuss the strengths and weaknesses of alternatives that have been suggested, and point to lines of further analysis.

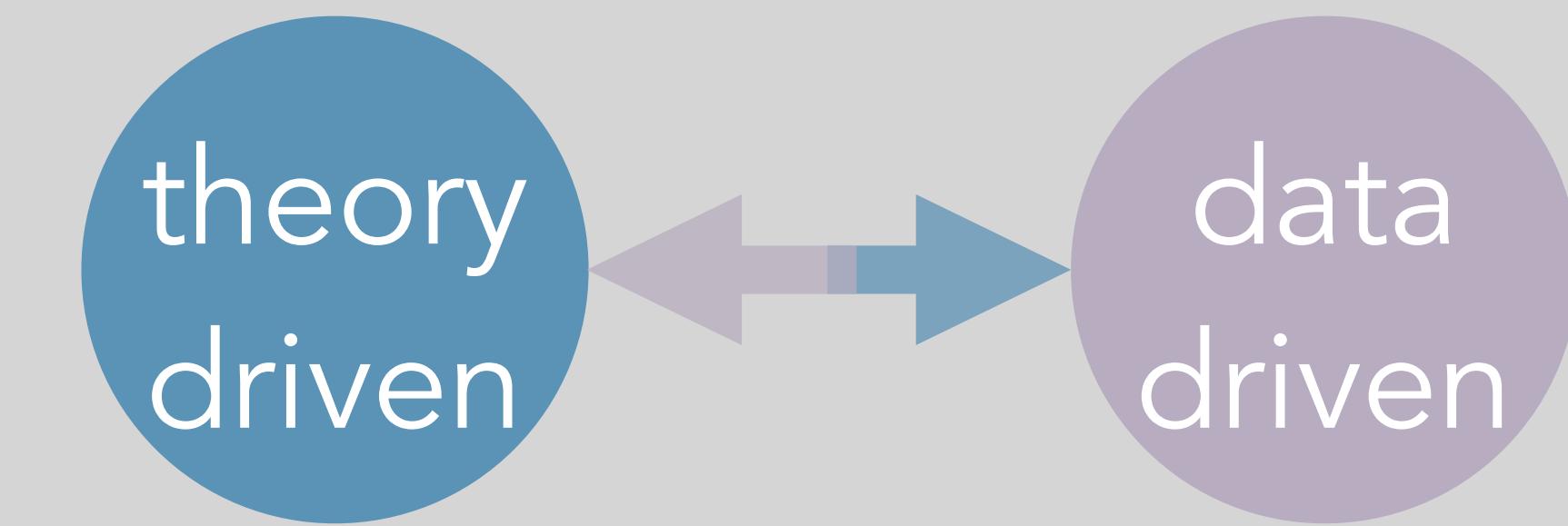


a shift towards prediction  
leads to a more reliable  
and useful social science

out-of-sample predictive ability:



clear measure of  
effect size



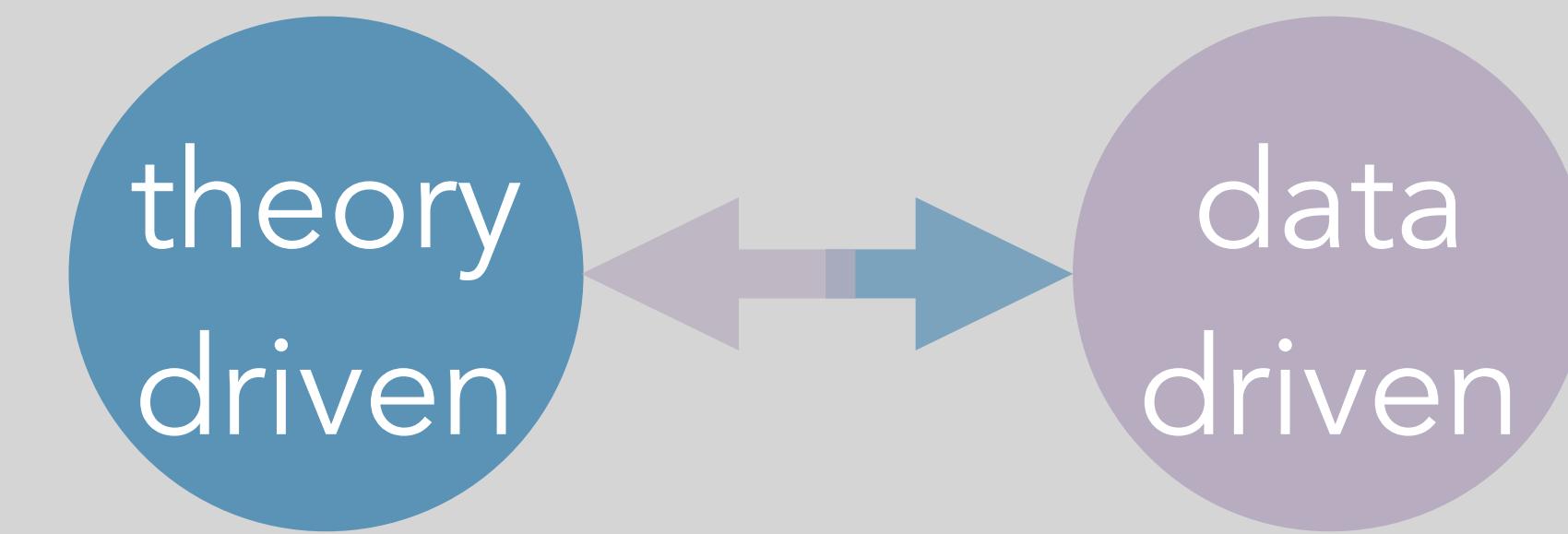
facilitates dialogue  
theory- and data-driven models

a shift towards prediction  
leads to a more reliable  
and useful social science

## out-of-sample predictive ability:



clear measure of  
effect size



facilitates dialogue  
theory- and data-driven models



measure of distance  
theory and practice



out-of-sample predictive ability  
is a measure of how useful  
our theory is in the real world

Articles

## The perils of policy by p-value: Predicting civil conflicts

Michael D Ward

*Department of Political Science, Duke University*

Brian D Greenhill

*Department of Political Science, University of Washington*

Kristin M Bakke

*Department of Political Science, University College London*

Journal of  
*Peace*  
Research

Journal of Peace Research  
47(4) 363–375  
© The Author(s) 2010  
Reprints and permission:  
[sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)  
DOI: 10.1177/0022343309356491  
[jpr.sagepub.com](http://jpr.sagepub.com)

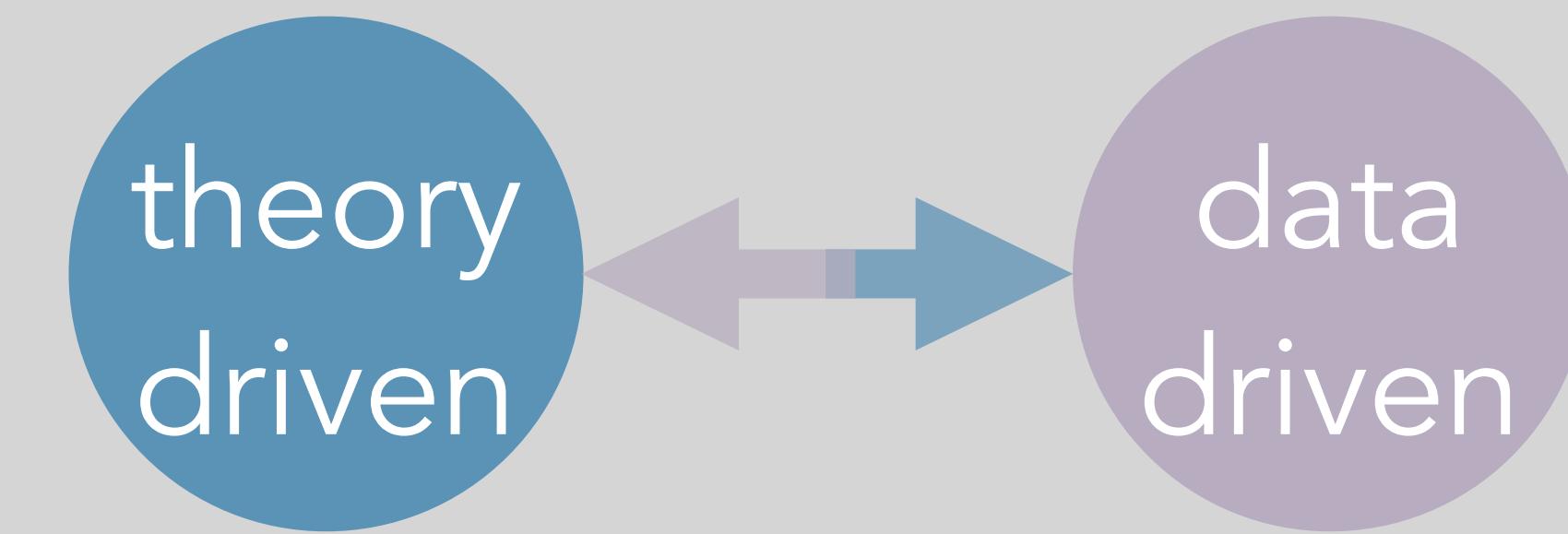
SAGE

a shift towards prediction  
leads to a more reliable  
and useful social science

## out-of-sample predictive ability:



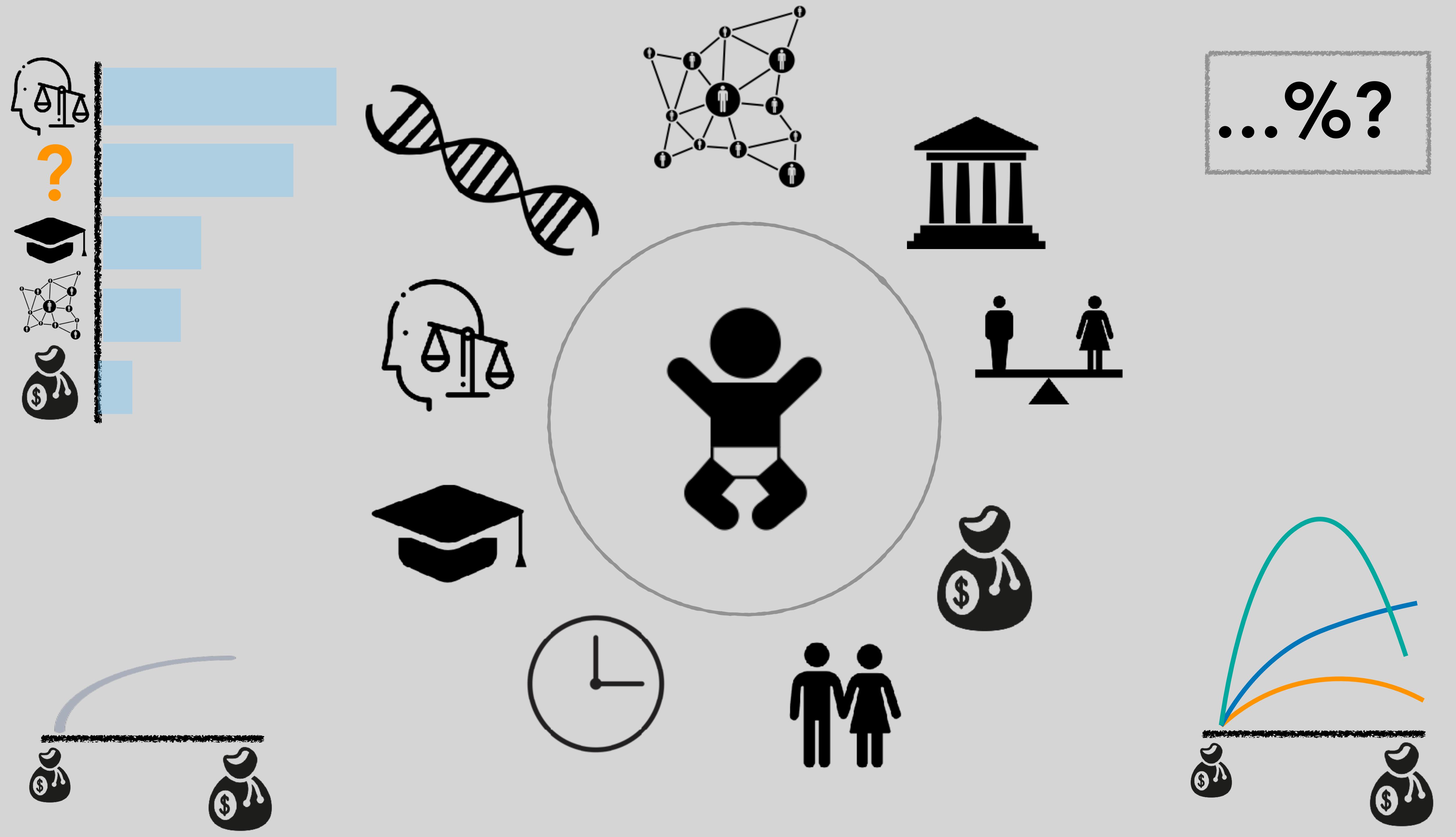
clear measure of  
effect size

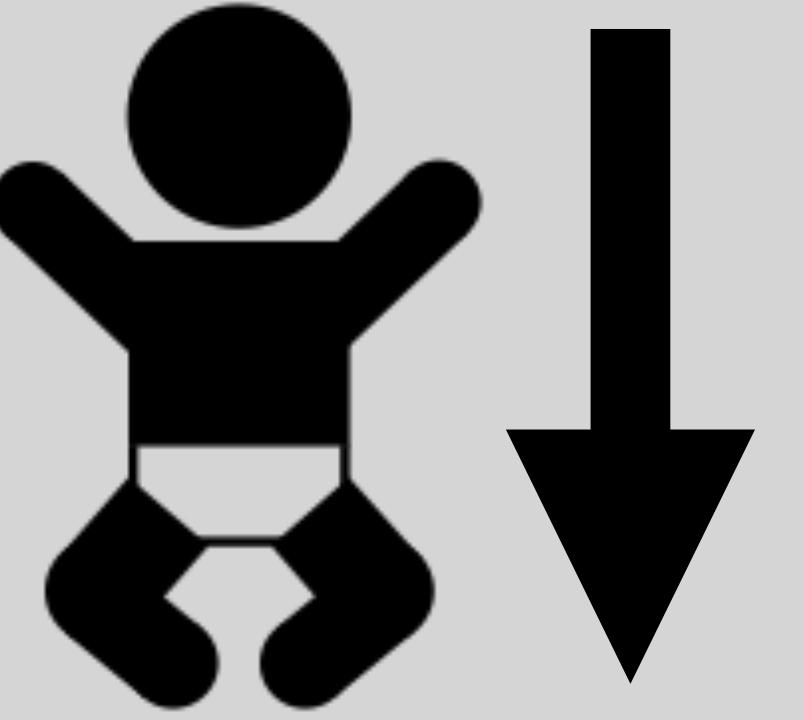
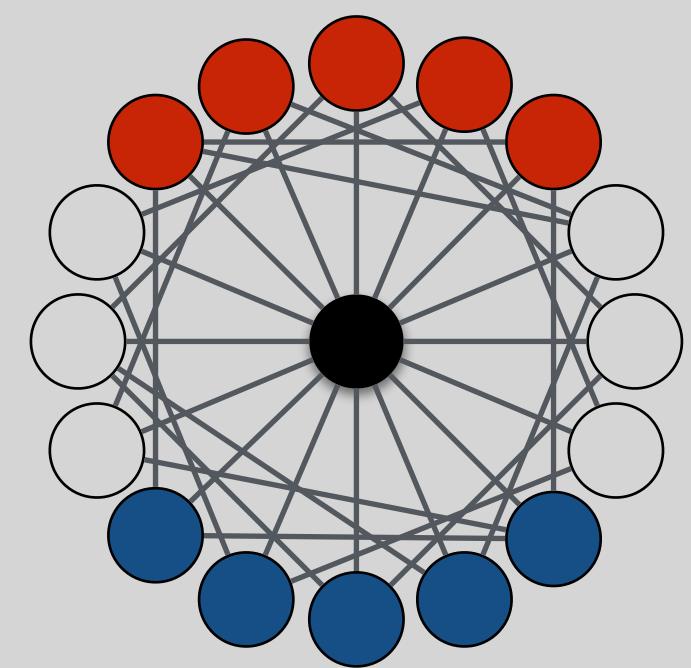


facilitates dialogue  
theory- and data-driven models



measure of distance  
theory and practice





Pre  
Fer  
1|0

Predicting Fertility  
data challenge



# Predicting Fertility data challenge

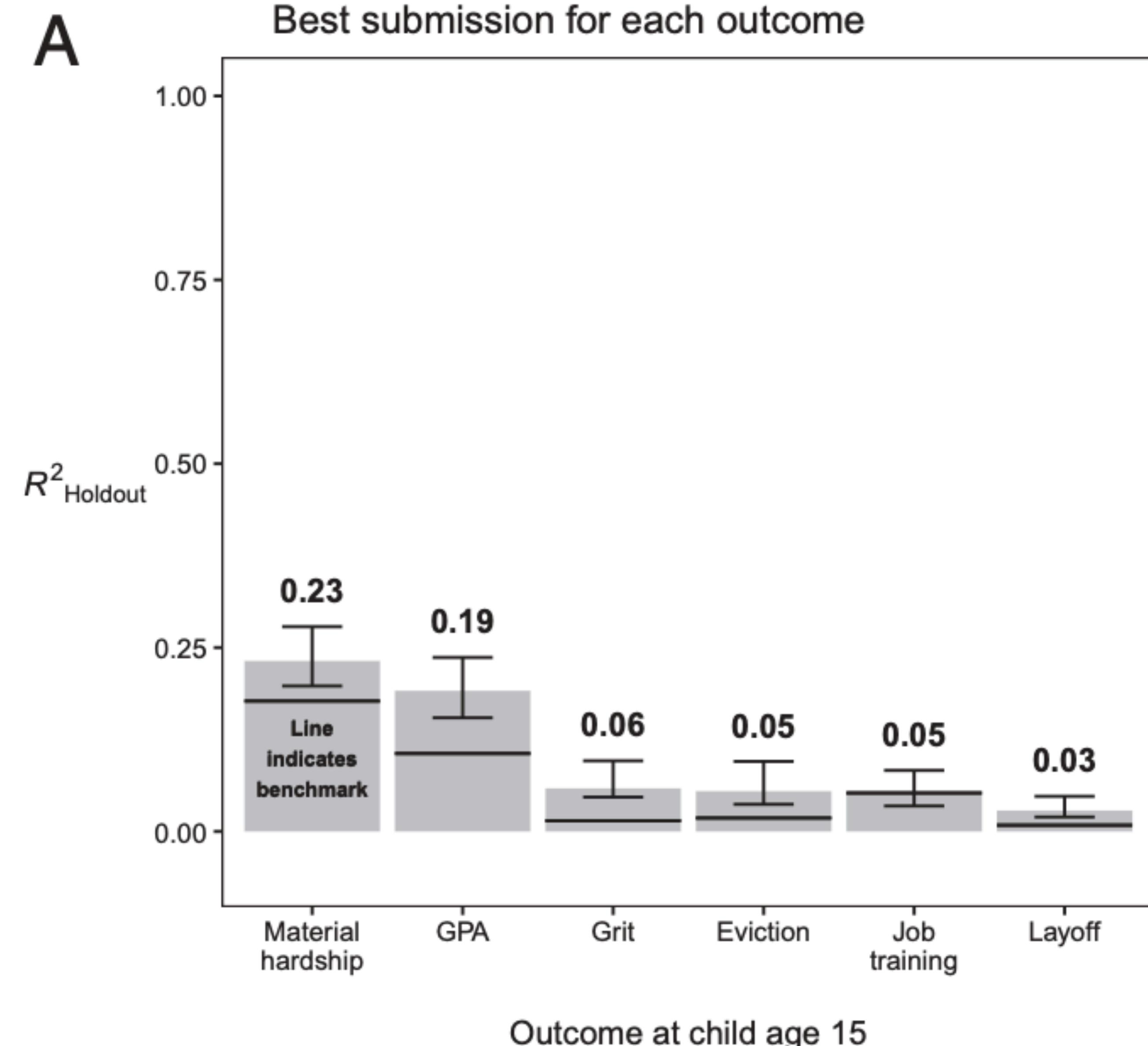
theory- and data-driven teams  
engage in common task  
using common data  
and common metric

# Low predictability

## Measuring the predictability of life outcomes with a scientific mass collaboration

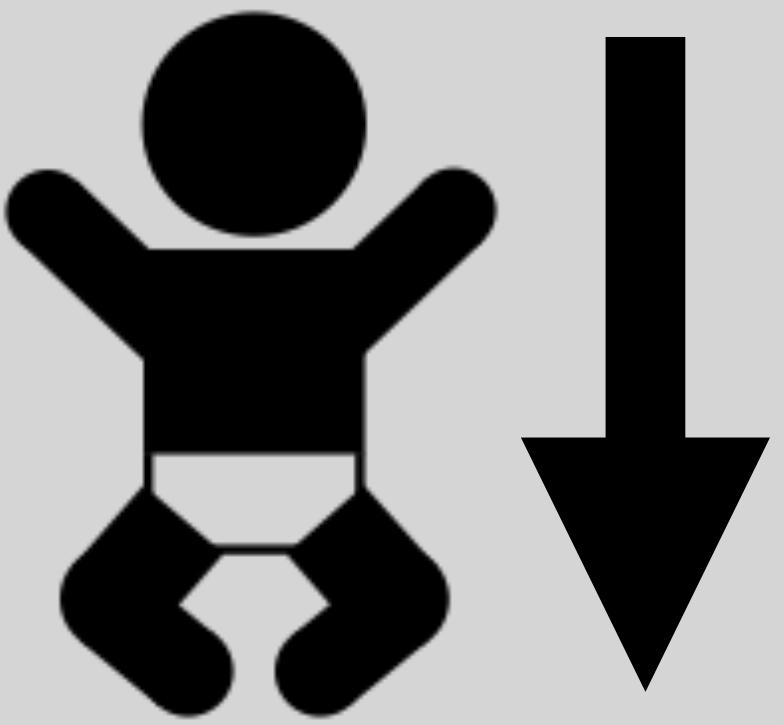
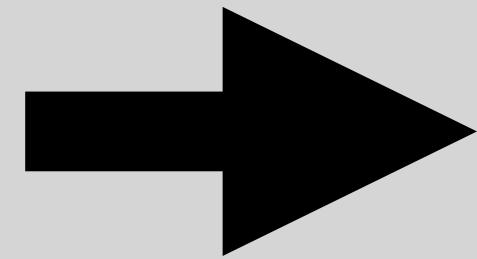
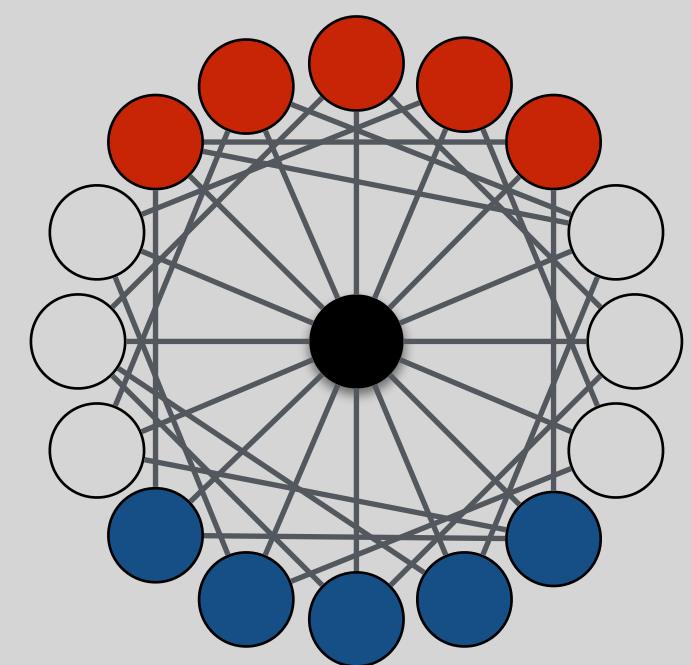
Matthew J. Salganik<sup>a,1</sup>, Ian Lundberg<sup>a</sup>, Alexander T. Kindel<sup>a</sup>, Caitlin E. Ahearn<sup>b</sup>, Khaled Al-Ghoneim<sup>c</sup>, Abdullah Almaatouq<sup>d,e</sup>, Drew M. Altschul<sup>f</sup>, Jennie E. Brand<sup>b,g</sup>, Nicole Bohme Carnegie<sup>h</sup>, Ryan James Compton<sup>i</sup>, Debanjan Datta<sup>j</sup>, Thomas Davidson<sup>k</sup>, Anna Filippova<sup>l</sup>, Connor Gilroy<sup>m</sup>, Brian J. Goode<sup>n</sup>, Eaman Jahani<sup>o</sup>, Ridhi Kashyap<sup>p,q,r</sup>, Antje Kirchner<sup>s</sup>, Stephen McKay<sup>t</sup>, Allison C. Morgan<sup>u</sup>, Alex Pentland<sup>v</sup>, Kivan Polimis<sup>w</sup>, Louis Raes<sup>x</sup>, Daniel E. Rigobon<sup>x</sup>, Claudia V. Roberts<sup>y</sup>, Diana M. Stanescu<sup>z</sup>, Yoshihiko Suhara<sup>t</sup>, Adaner Usmani<sup>aa</sup>, Erik H. Wang<sup>x</sup>, Muna Adem<sup>bb</sup>, Abdulla Alhajri<sup>cc</sup>, Bedoor AlShebli<sup>dd</sup>, Redwane Amin<sup>ee</sup>, Ryan B. Amos<sup>y</sup>, Lisa P. Argyle<sup>ff</sup>, Livia Baer-Bositis<sup>gg</sup>, Moritz Büchi<sup>hh</sup>, Bo-Ryehn Chung<sup>ii</sup>, William Eggert<sup>ll</sup>, Gregory Faletto<sup>kk</sup>, Zhilin Fan<sup>ll</sup>, Jeremy Freese<sup>gg</sup>, Tejomay Gadgil<sup>mm</sup>, Josh Gagné<sup>gg</sup>, Yue Gao<sup>nn</sup>, Andrew Halpern-Manners<sup>bb</sup>, Sonia P. Hashim<sup>y</sup>, Sonia Hausen<sup>gg</sup>, Guanhua He<sup>oo</sup>, Kimberly Higuera<sup>gg</sup>, Bernie Hogan<sup>pp</sup>, Ilana M. Horwitz<sup>qq</sup>, Lisa M. Hummel<sup>gg</sup>, Naman Jain<sup>x</sup>, Kun Jin<sup>rr</sup>, David Jurgens<sup>ss</sup>, Patrick Kaminski<sup>bb,tt</sup>, Areg Karapetyan<sup>uu,ww</sup>, E. H. Kim<sup>gg</sup>, Ben Leizman<sup>y</sup>, Naijia Liu<sup>z</sup>, Malte Möser<sup>y</sup>, Andrew E. Mack<sup>x</sup>, Mayank Mahajan<sup>y</sup>, Noah Mandell<sup>ww</sup>, Helge Marahrens<sup>bb</sup>, Diana Mercado-Garcia<sup>aa</sup>, Viola Mocz<sup>xx</sup>, Katarina Mueller-Gastell<sup>gg</sup>, Ahmed Musse<sup>yy</sup>, Qiankun Niu<sup>ee</sup>, William Nowak<sup>zz</sup>, Hamidreza Omidvar<sup>aa</sup>, Andrew Or<sup>y</sup>, Karen Ouyang<sup>y</sup>, Katy M. Pinto<sup>bbb</sup>, Ethan Porter<sup>cc</sup>, Kristin E. Porter<sup>dd</sup>, Crystal Qian<sup>y</sup>, Tamkinat Rauf<sup>gg</sup>, Anahit Sargsyan<sup>ee</sup>, Thomas Schaffner<sup>y</sup>, Landon Schnabel<sup>gg</sup>, Bryan Schonfeld<sup>z</sup>, Ben Sender<sup>ff</sup>, Jonathan D. Tang<sup>y</sup>, Emma Tsurkov<sup>gg</sup>, Austin van Loon<sup>gg</sup>, Onur Varol<sup>gg,hh,ii</sup>, Xiafei Wang<sup>ll</sup>, Zhi Wang<sup>hh,jj</sup>, Julia Wang<sup>y</sup>, Flora Wang<sup>ff</sup>, Samantha Weissman<sup>y</sup>, Kirstie Whitaker<sup>kk,ii</sup>, Maria K. Wolters<sup>mm</sup>, Wei Lee Woon<sup>nn</sup>, James Wu<sup>ooo</sup>, Catherine Wu<sup>y</sup>, Kengran Yang<sup>aa</sup>, Jingwen Yin<sup>ll</sup>, Bingyu Zhao<sup>ppp</sup>, Chenyun Zhu<sup>ll</sup>, Jeanne Brooks-Gunn<sup>qq,rr</sup>, Barbara E. Engelhardt<sup>yy,ii</sup>, Moritz Hardt<sup>ss</sup>, Dean Knox<sup>x</sup>, Karen Levy<sup>tt</sup>, Arvind Narayanan<sup>y</sup>, Brandon M. Stewart<sup>x</sup>, Duncan J. Watts<sup>uu,vv,ww</sup>, and Sara McLanahan<sup>a,1</sup>

**data challenge:**  
predicting life outcomes  
based on ~6000 variables  
by 160 teams  
both theory- & data-driven

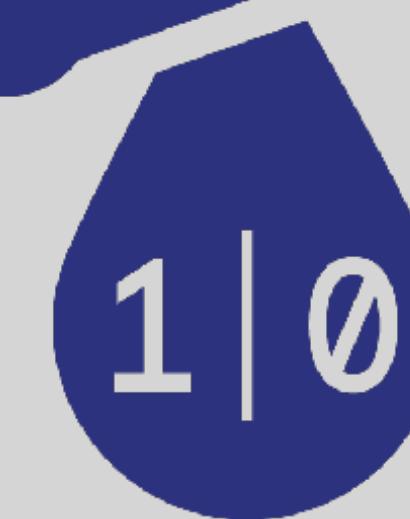
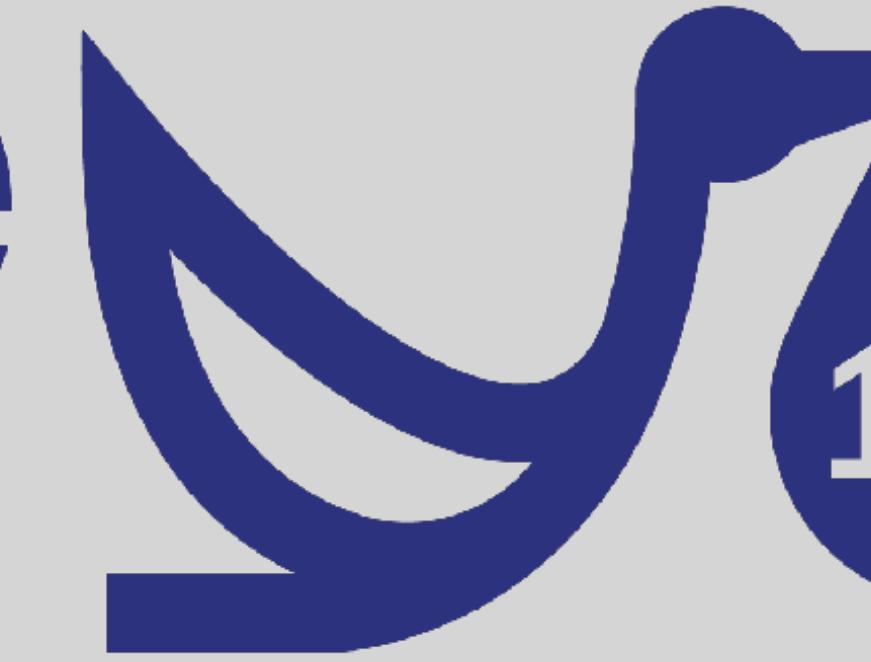




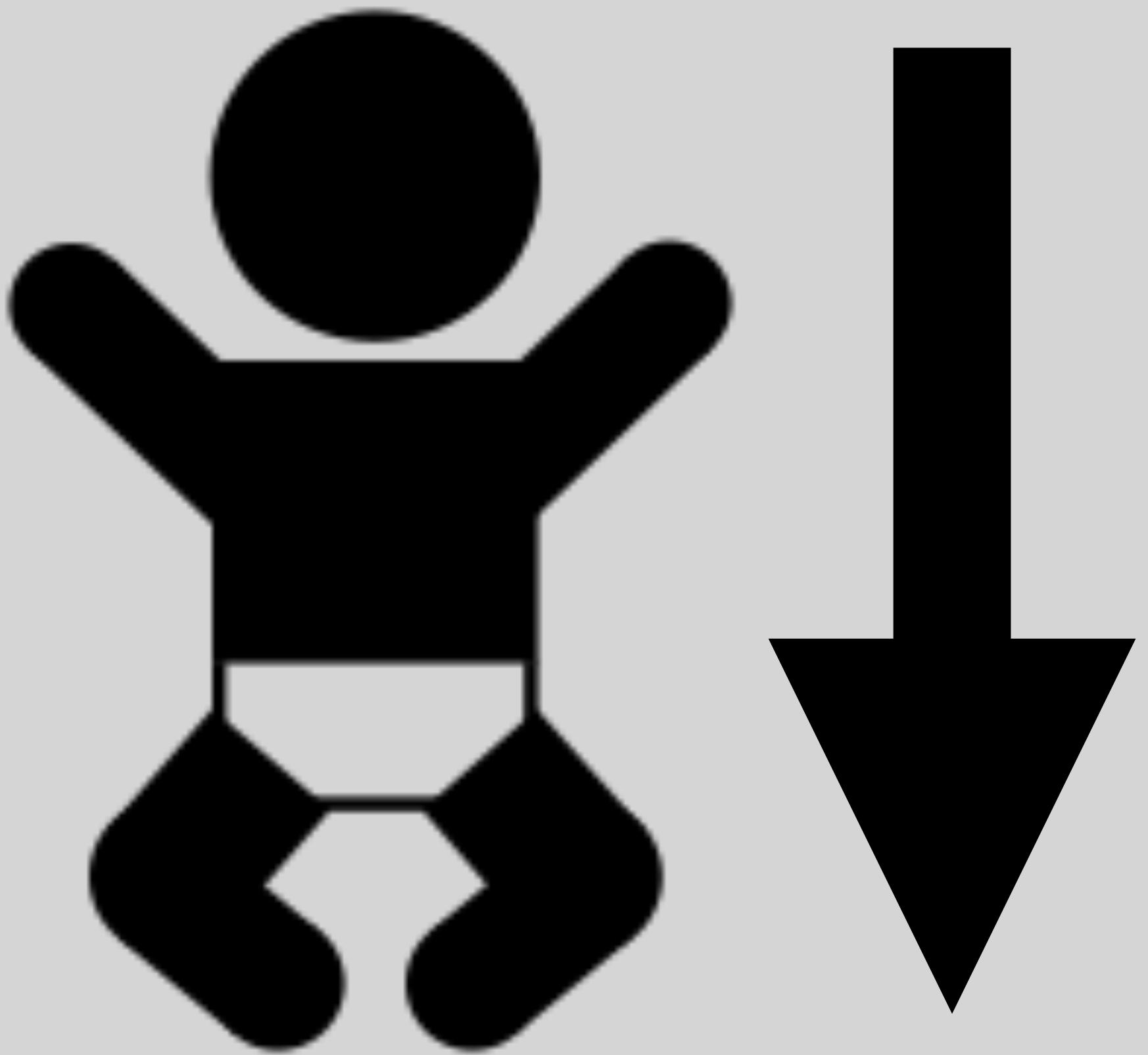
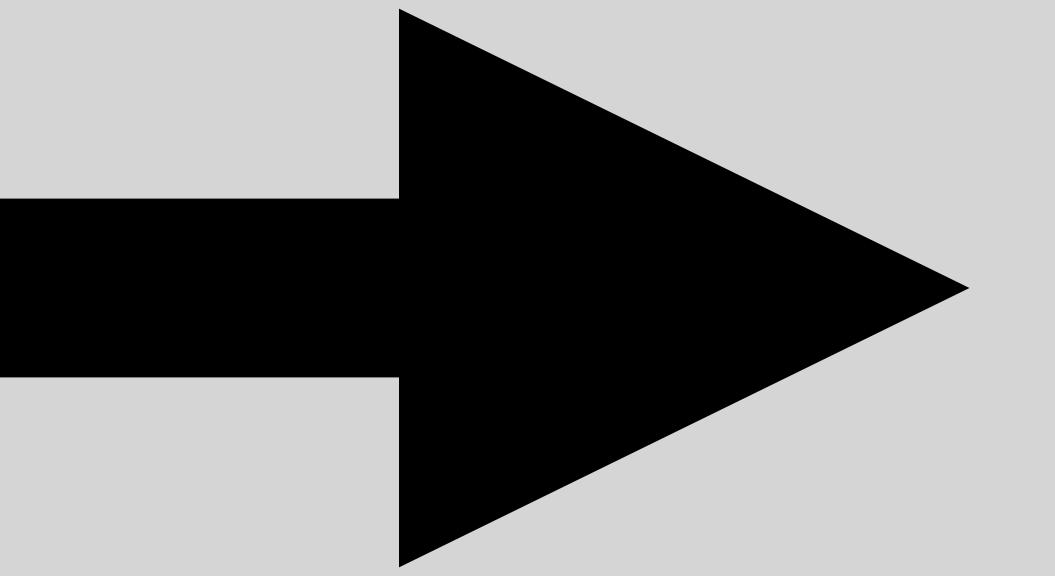
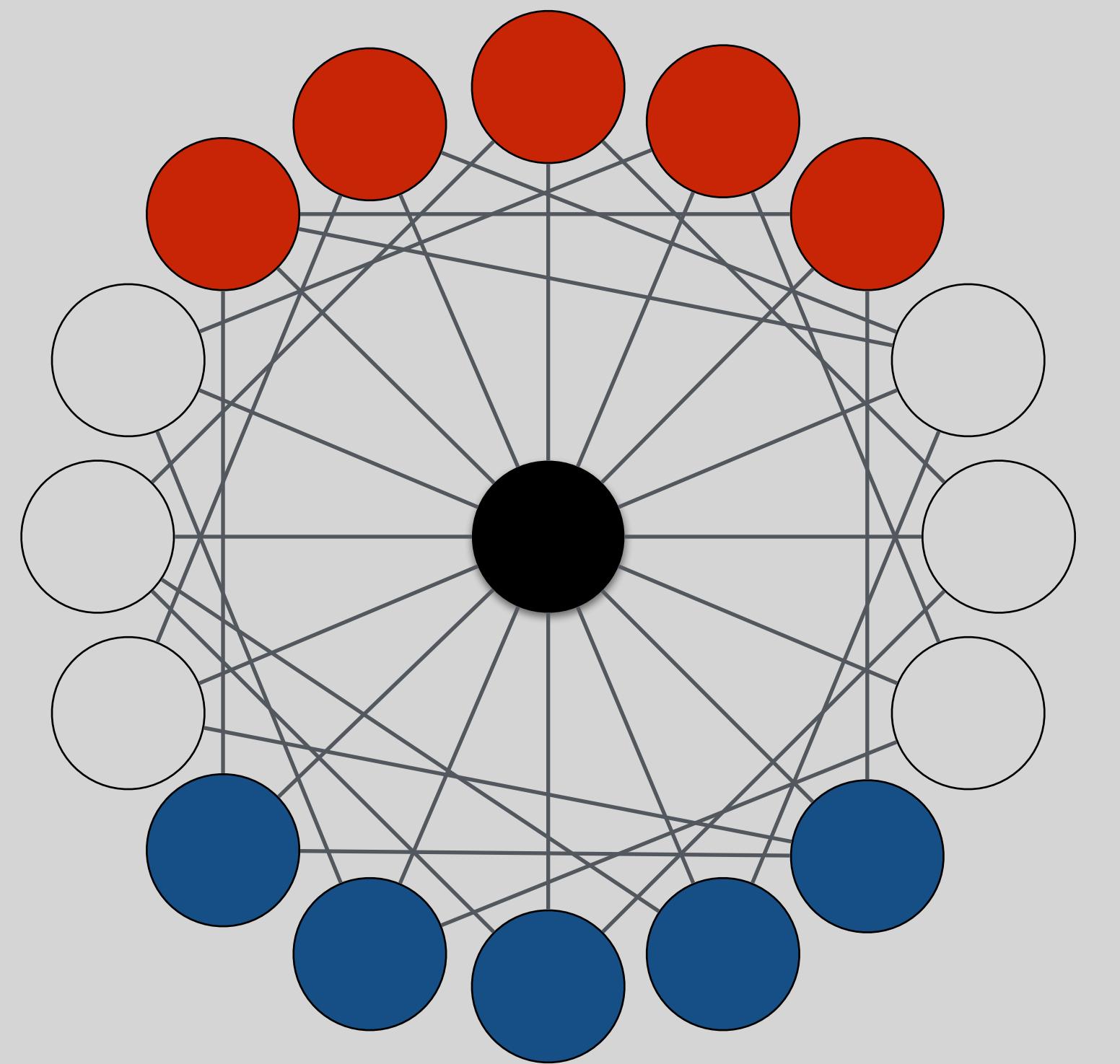
Social scientists studying the life course must find a way to reconcile a widespread belief that understanding has been generated by these data—as demonstrated by more than 750 published journal articles using the Fragile Families data with the fact that the very same data could not yield accurate predictions of these important outcomes.

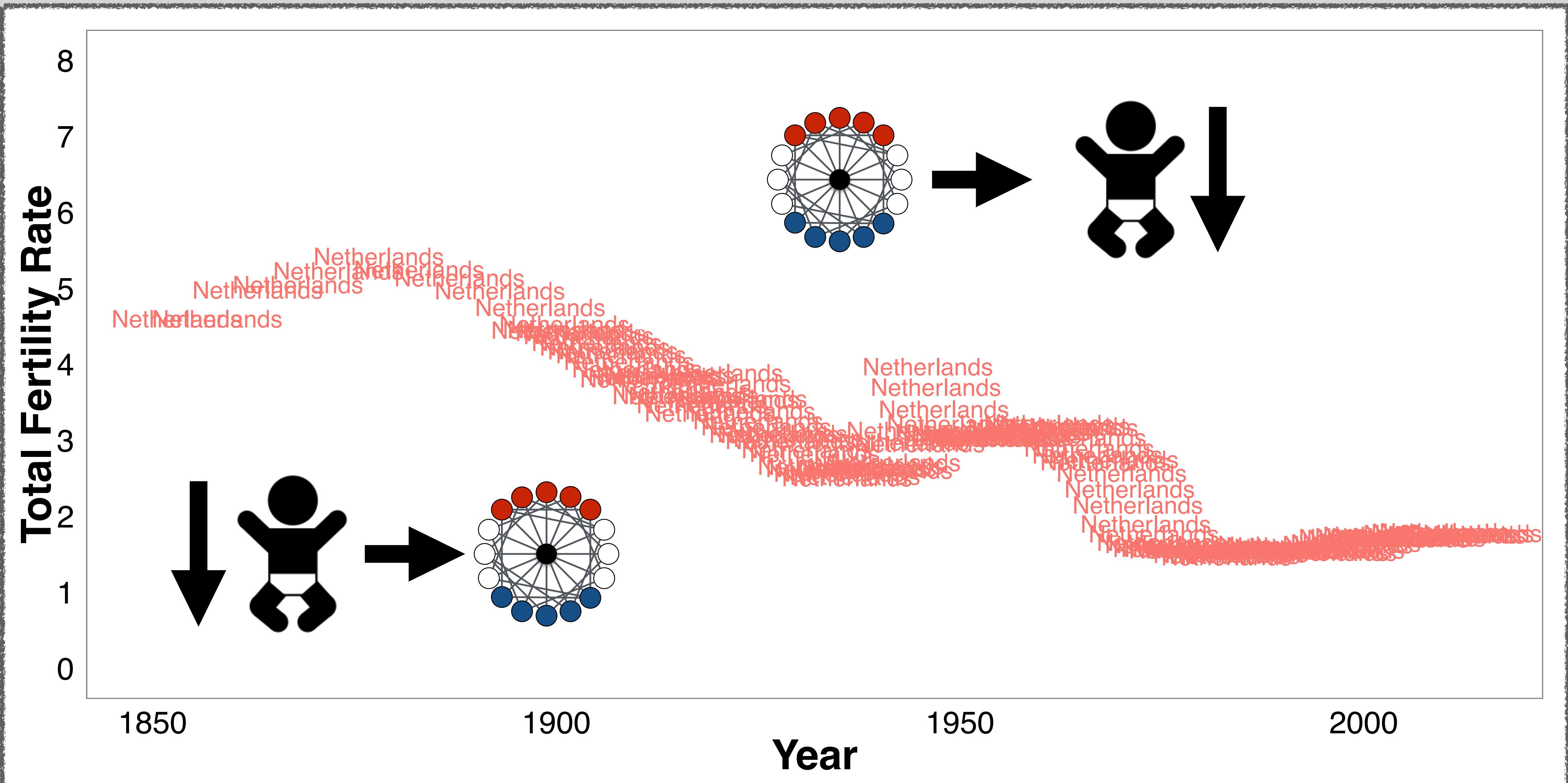


# Pre Fer



## Predicting Fertility data challenge



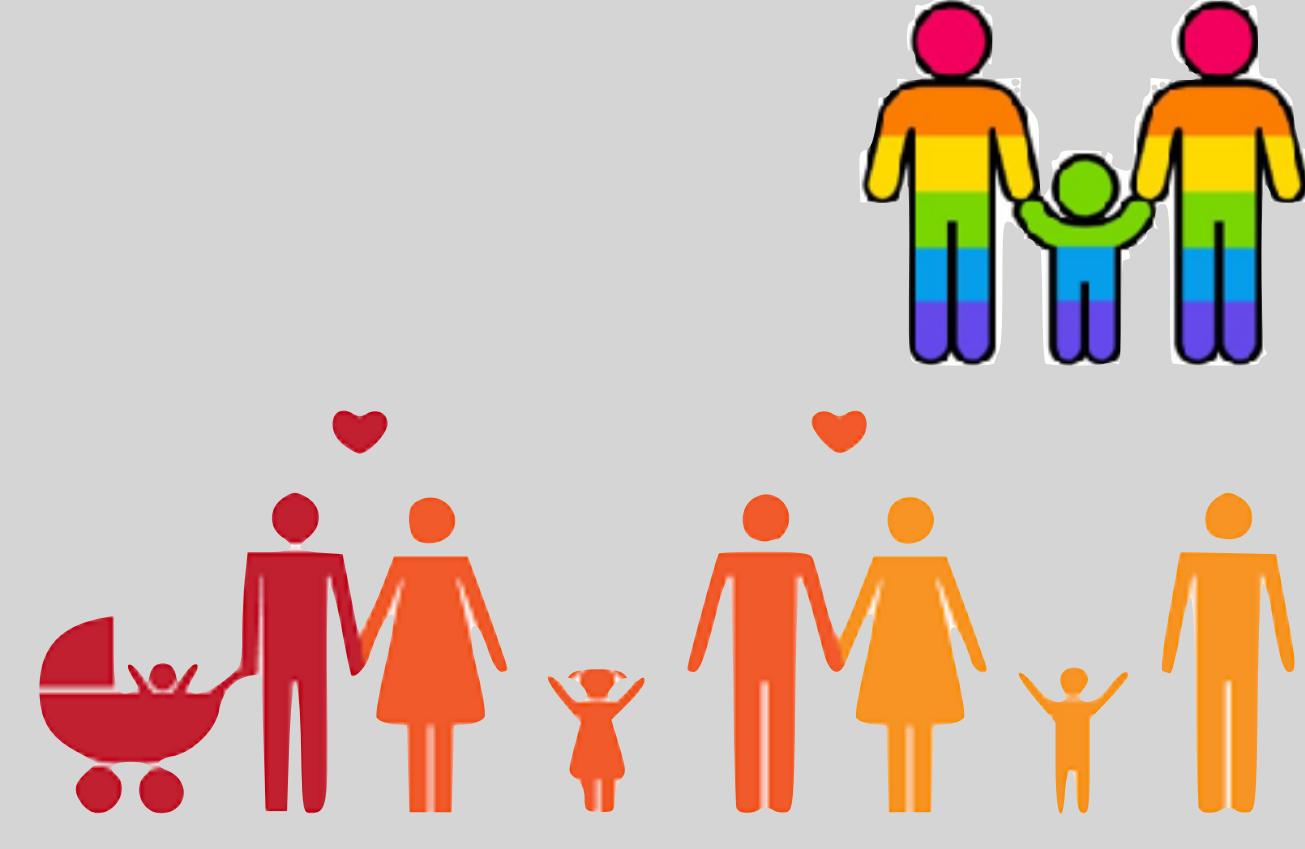




FROM  
EXTENDED  
KIN  
NETWORKS



TO  
IMPORTANCE  
OF NUCLEAR  
FAMILY



TO  
A DIVERSITY  
OF  
FAMILIES

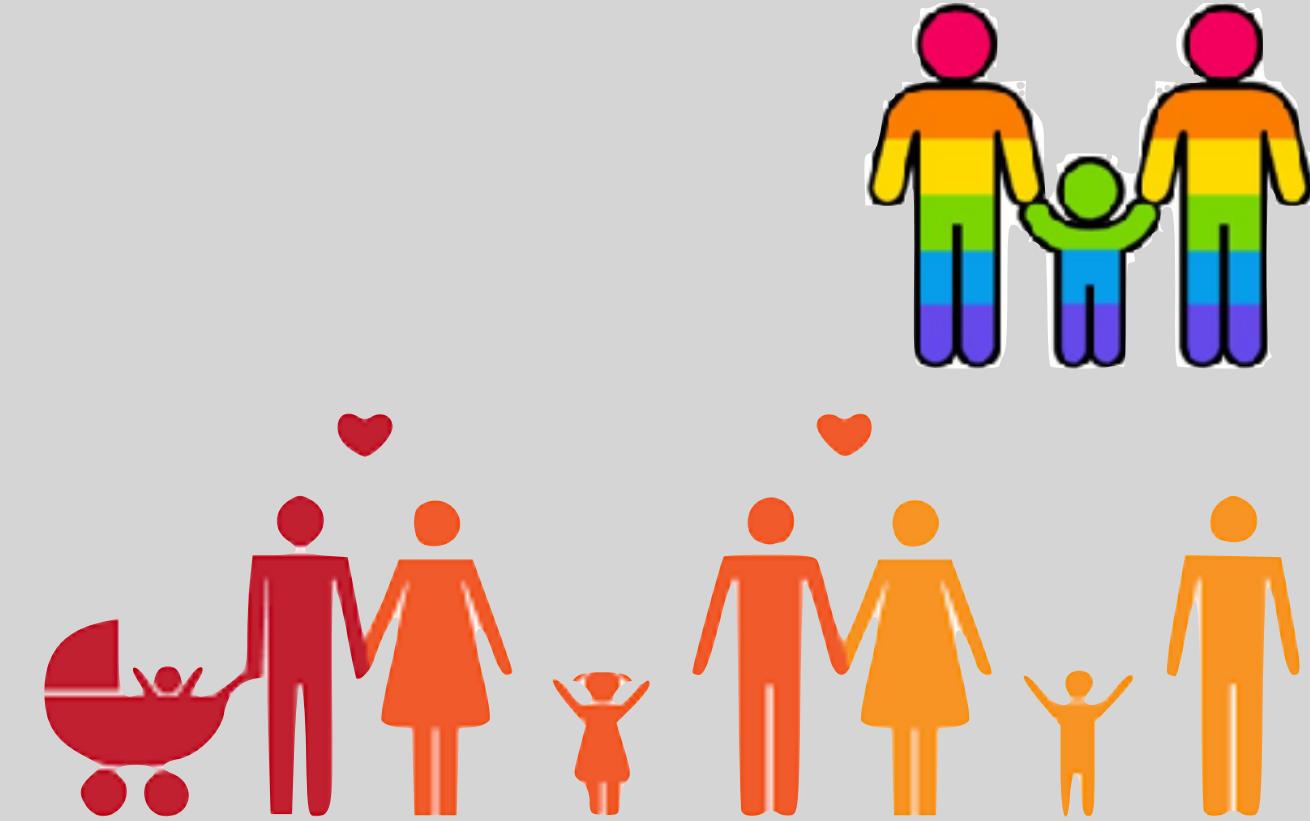


FROM  
EXTENDED  
KIN  
NETWORKS



TO  
IMPORTANCE  
OF NUCLEAR  
FAMILY

worries about social cohesion and  
the general demise of civilisation



TO  
A DIVERSITY  
OF  
FAMILIES



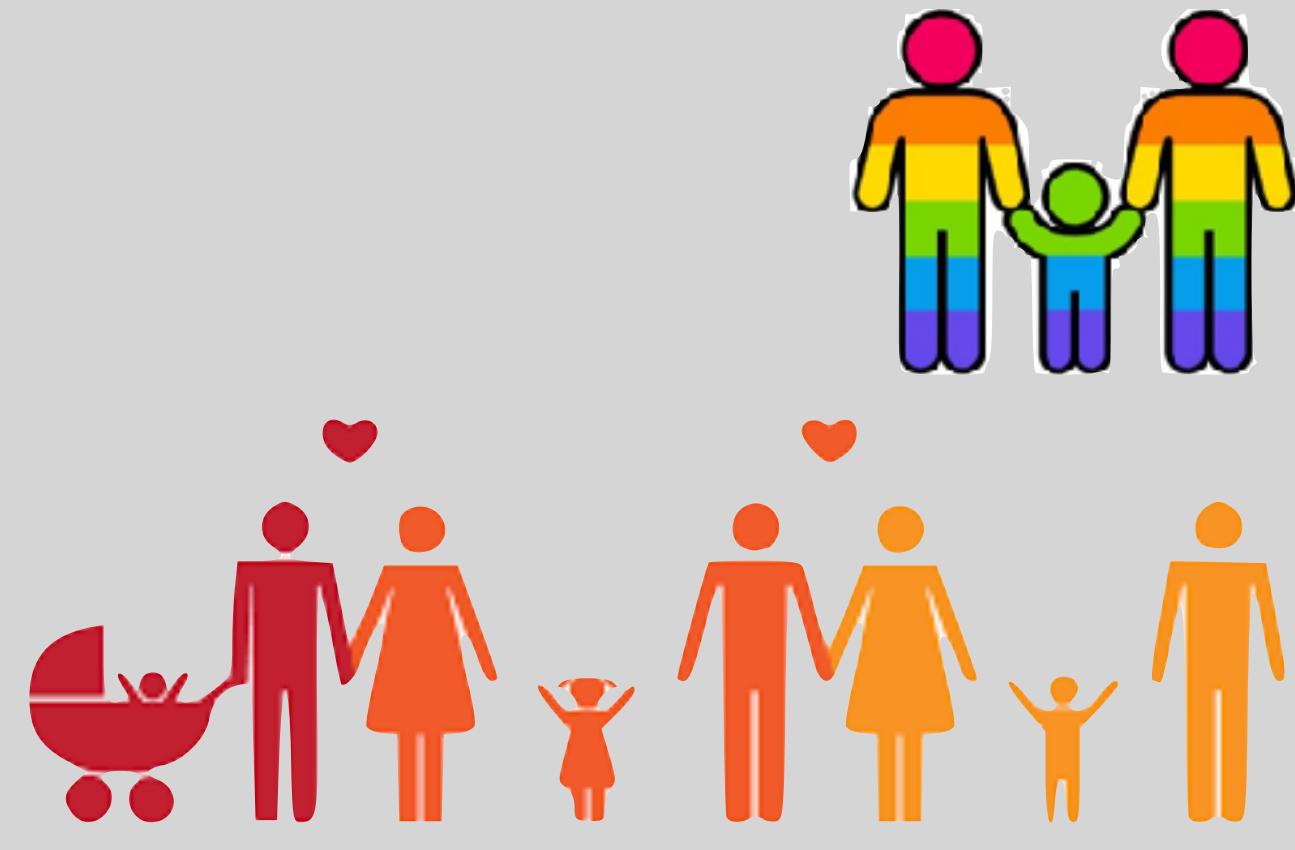
FROM  
EXTENDED  
KIN  
NETWORKS

CUSTOM  
AND  
LAWS

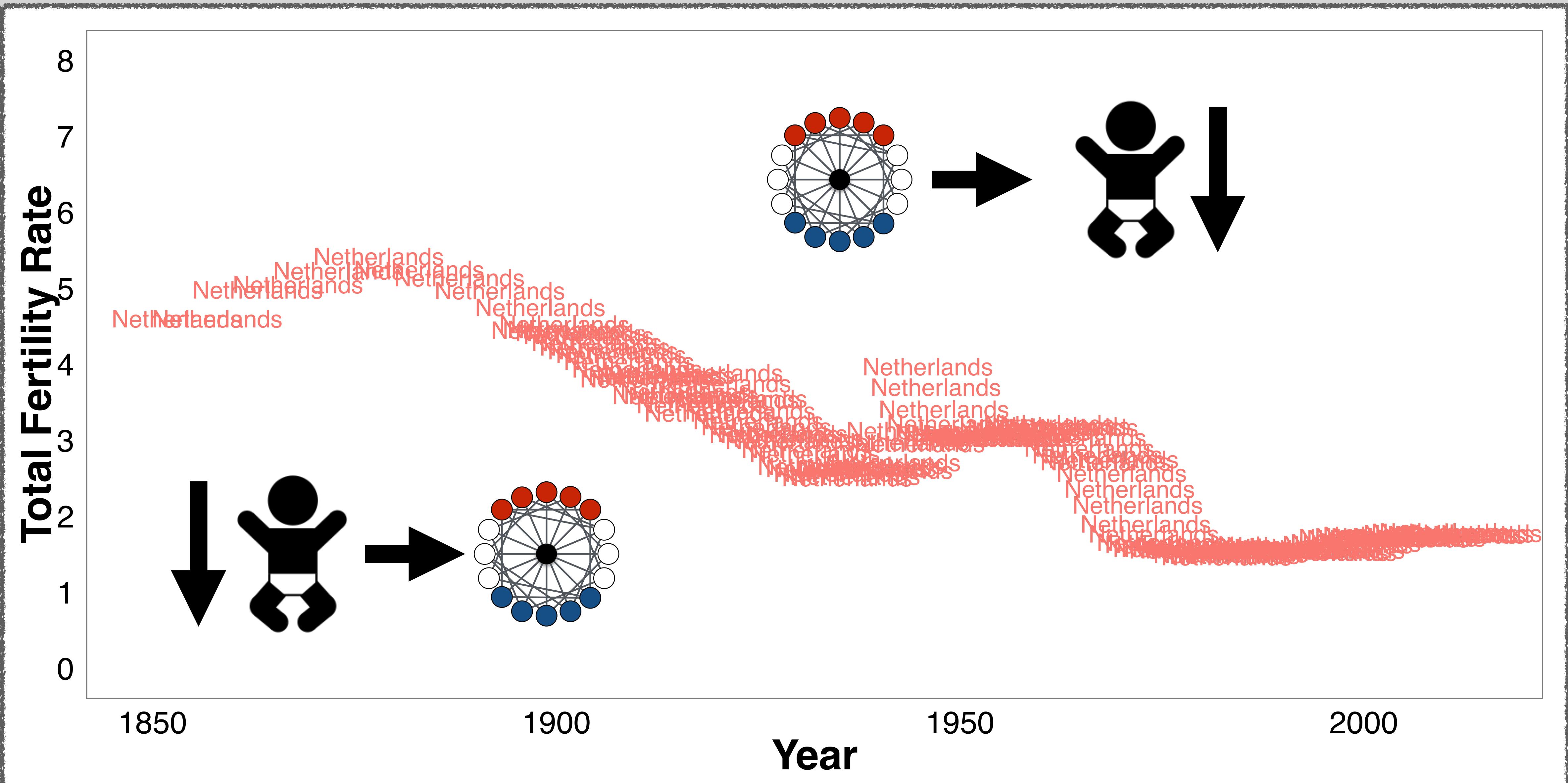


TO  
IMPORTANCE  
OF NUCLEAR  
FAMILY

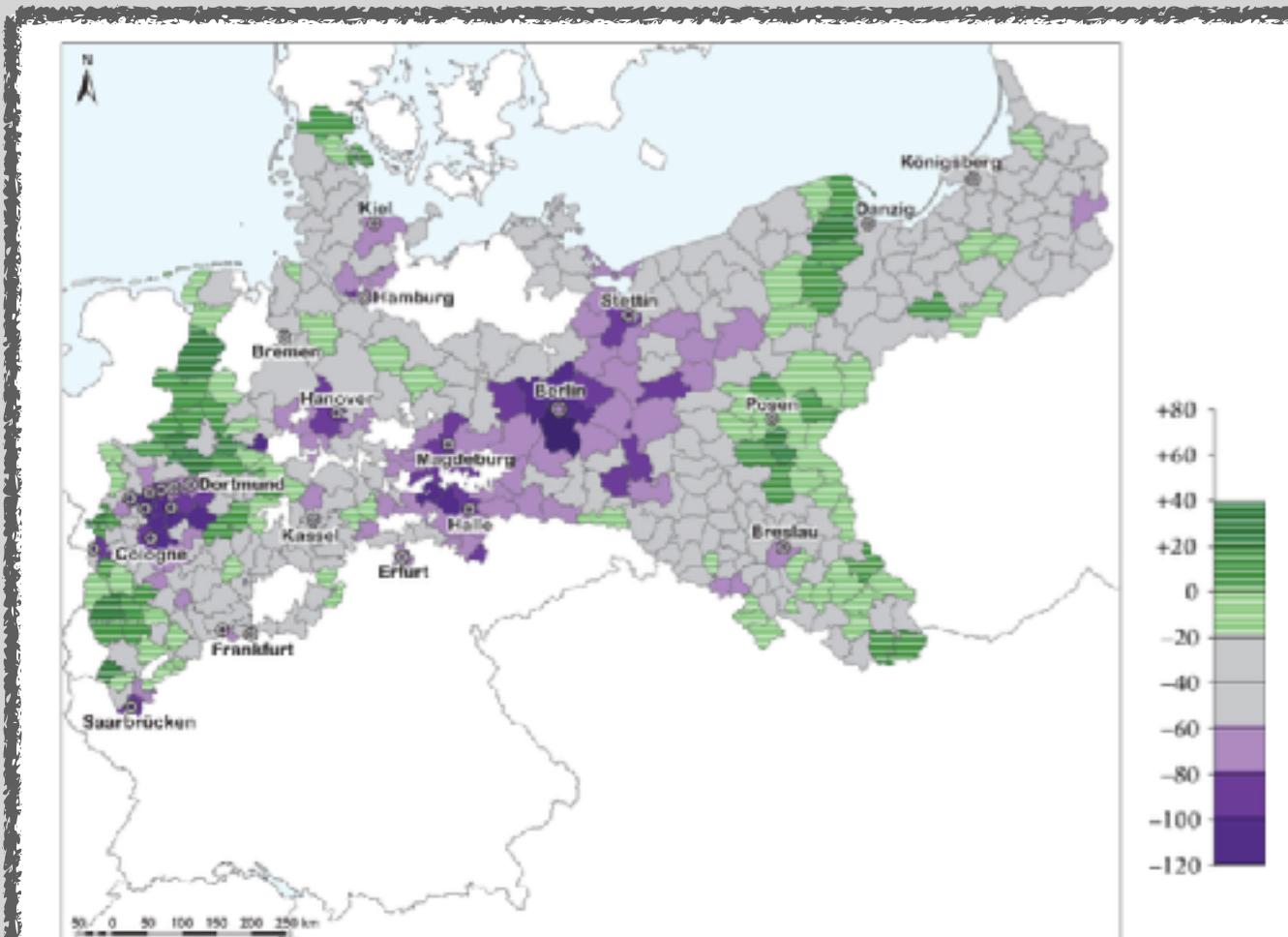
*love  
and  
affection*



TO  
A DIVERSITY  
OF  
FAMILIES



historical  
data



**Spatial Analysis of the  
Causes of Fertility Decline  
in Prussia**

JOSHUA R. GOLDSTEIN  
SEBASTIAN KLÜSENER



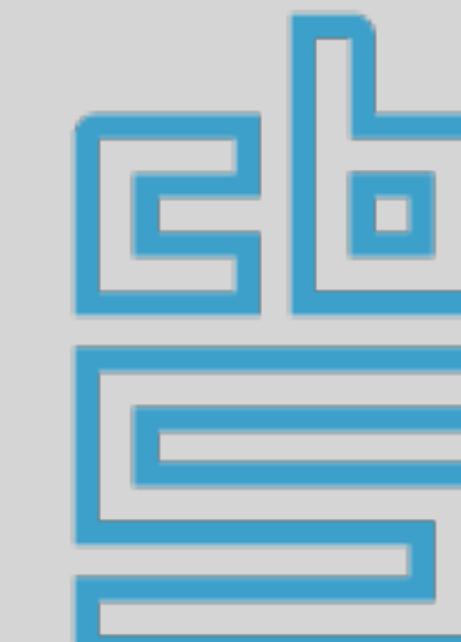
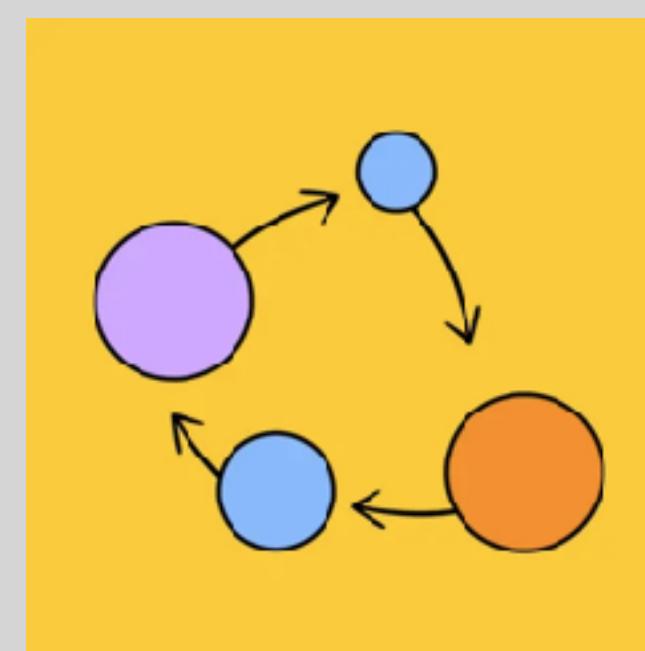
convenience  
samples

**Does Fertility Behavior  
Spread among Friends?**

Nicoletta Balbo<sup>a</sup> and Nicola Barban<sup>b</sup>

**Family, Firms, and Fertility: A Study of Social  
Interaction Effects**

Zafer Buyukkececi<sup>1</sup> · Thomas Leopold<sup>2</sup> · Ruben van Gaalen<sup>3</sup> ·  
Henriette Engelhardt<sup>4</sup>

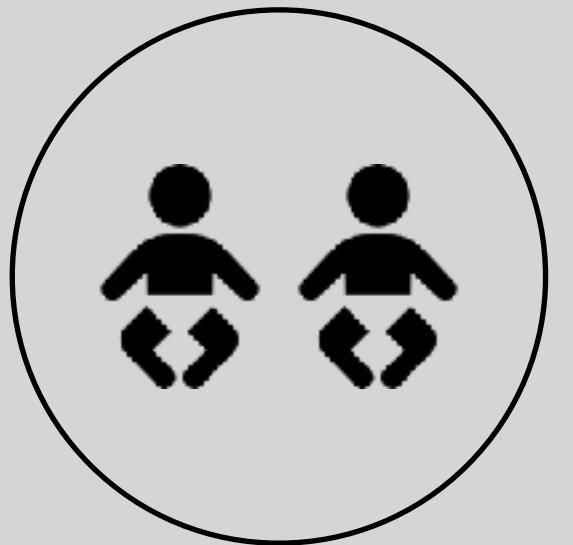


causal  
design

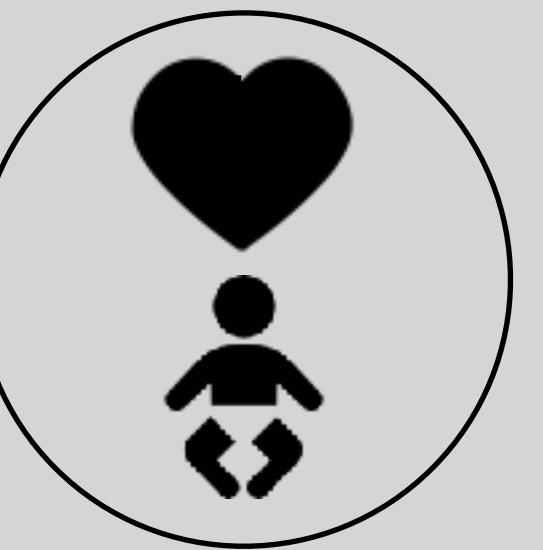
*social learning  
social contagion  
social pressure  
social support*

qualitative  
studies

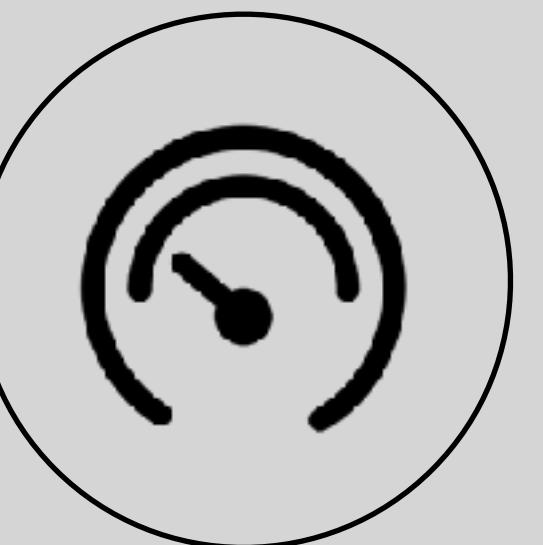
social learning



social contagion

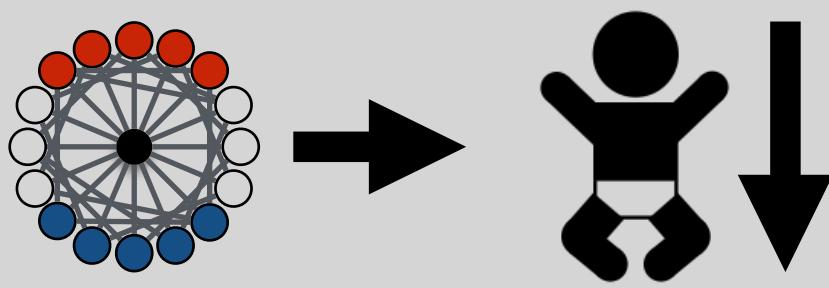


social support



social pressure

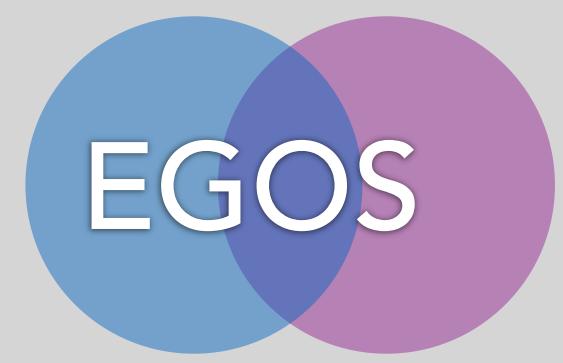
quantifying social influences

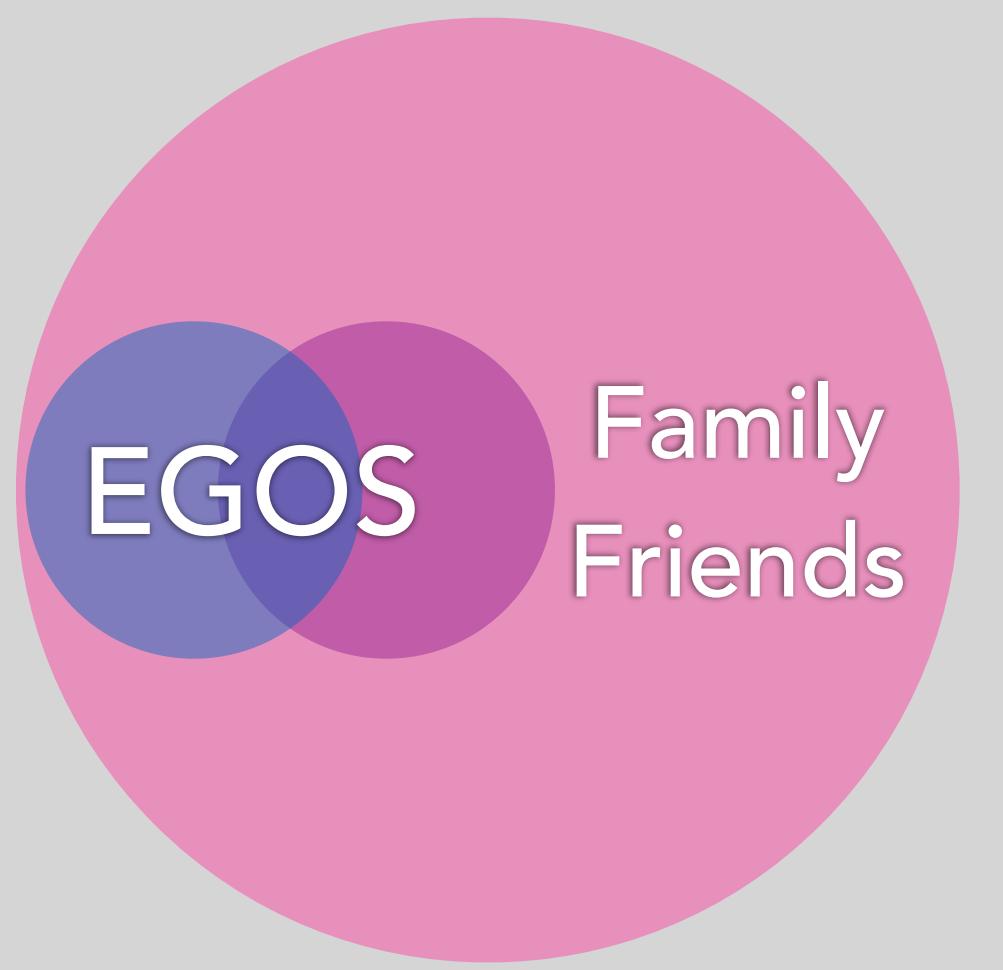


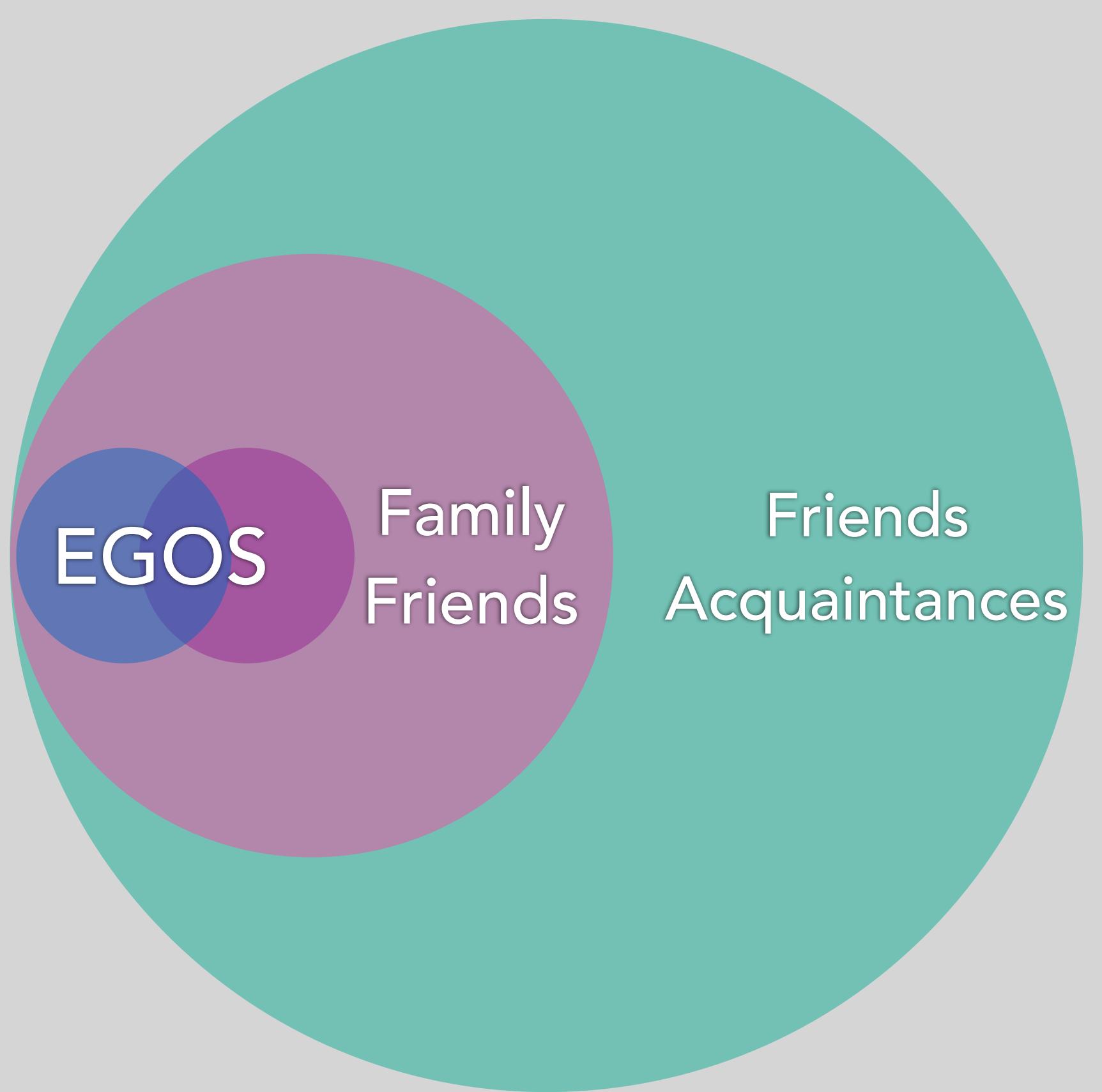
on fertility behaviour

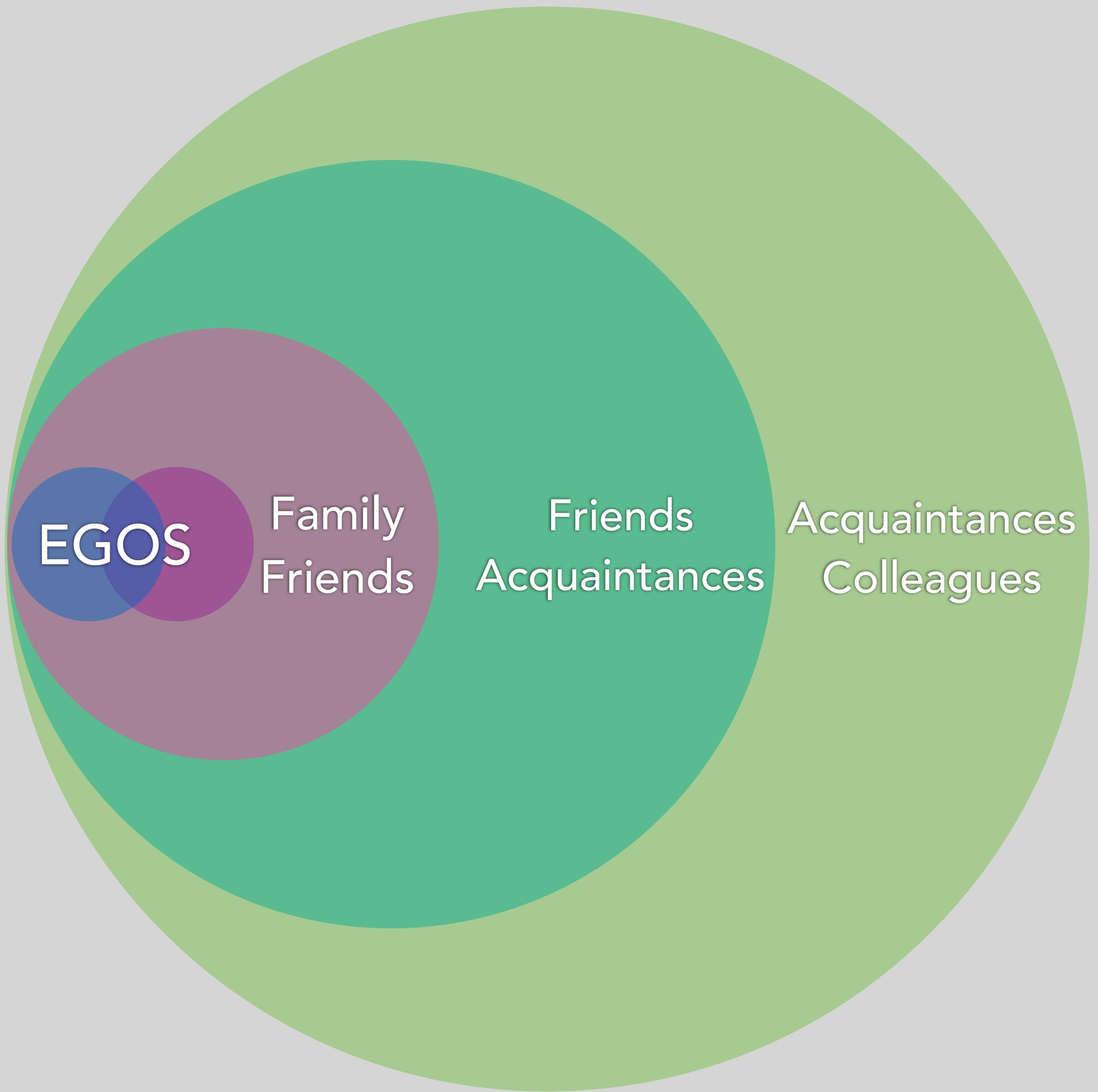
using personal network data

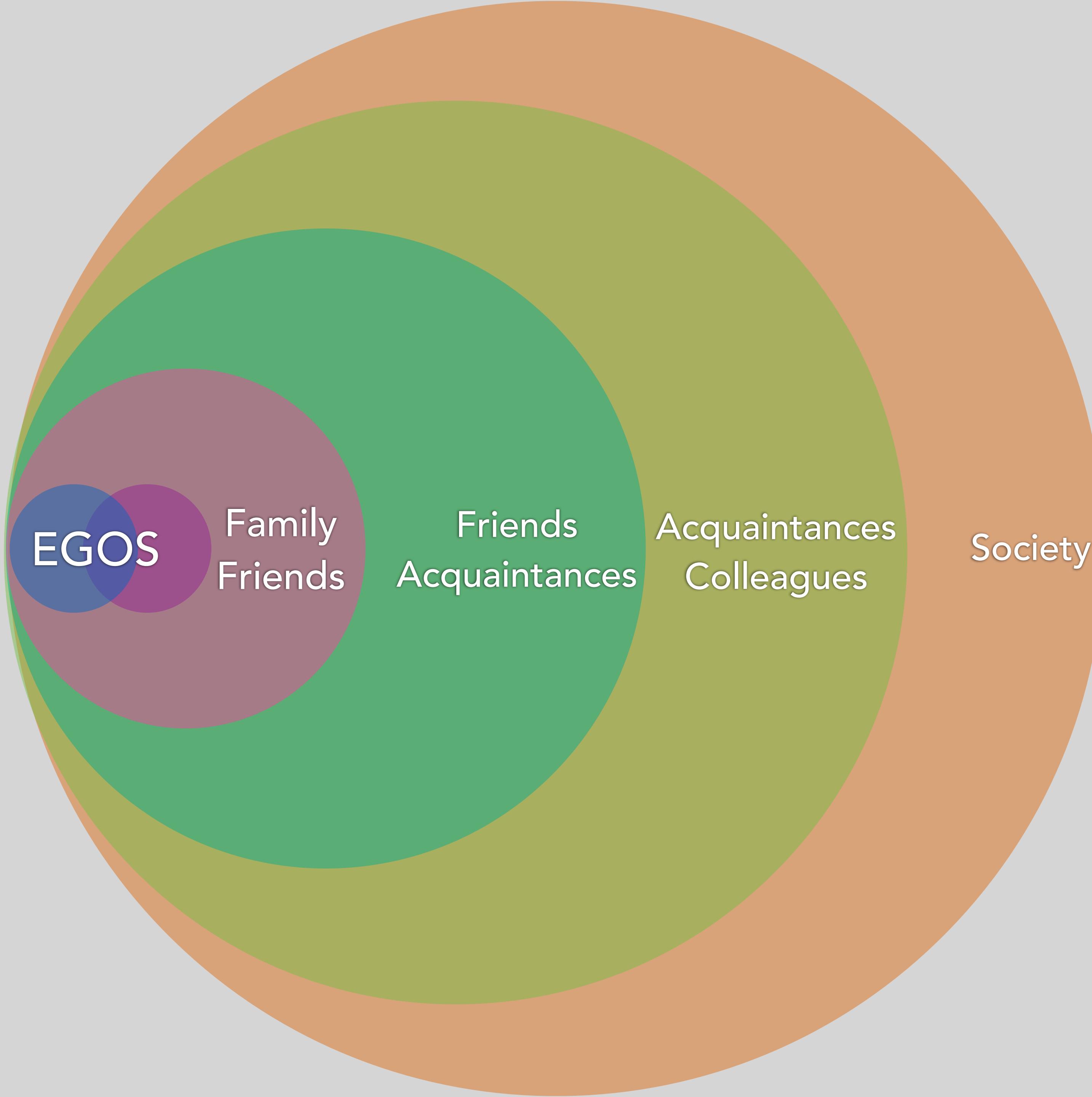
EGO

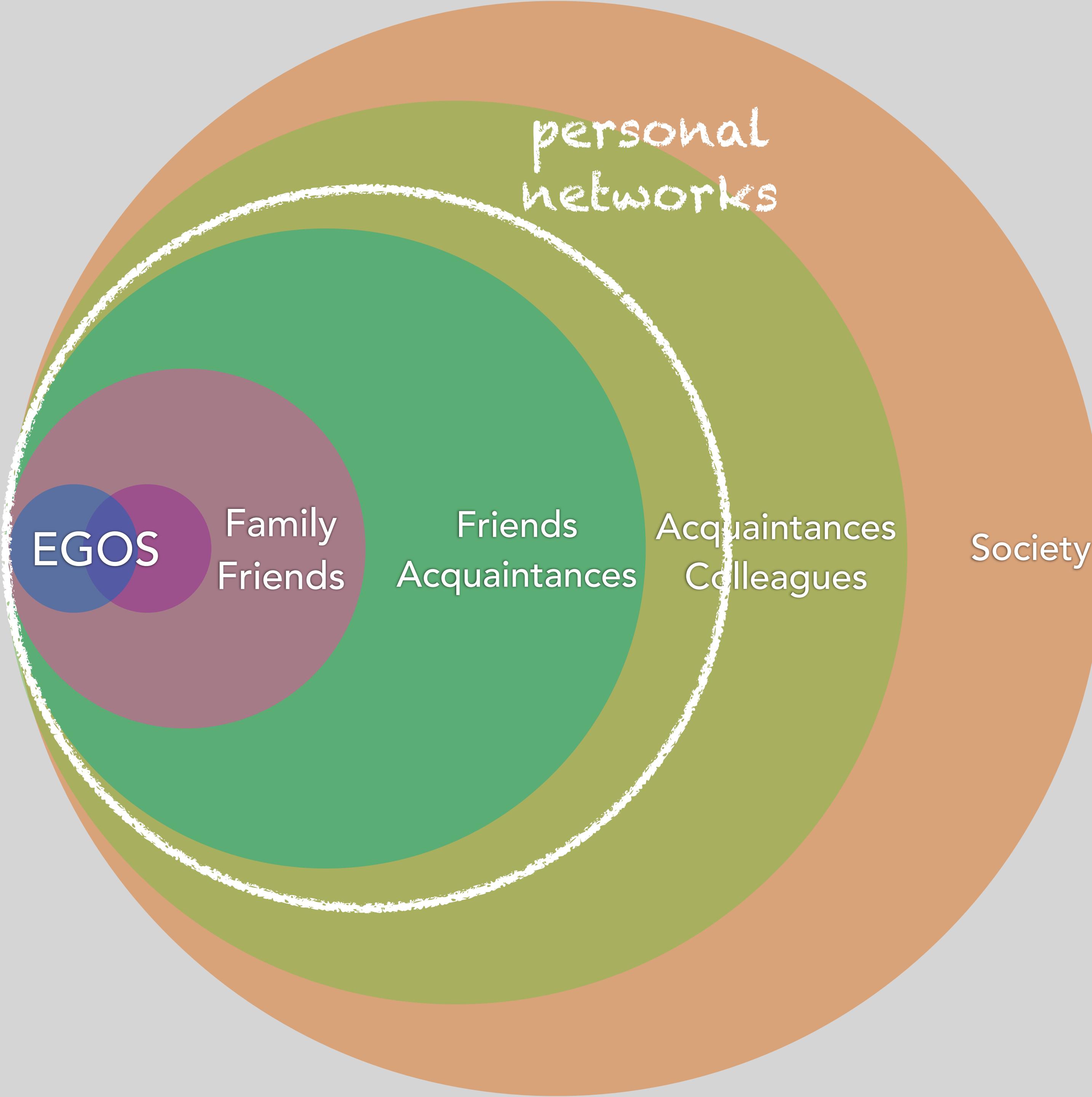






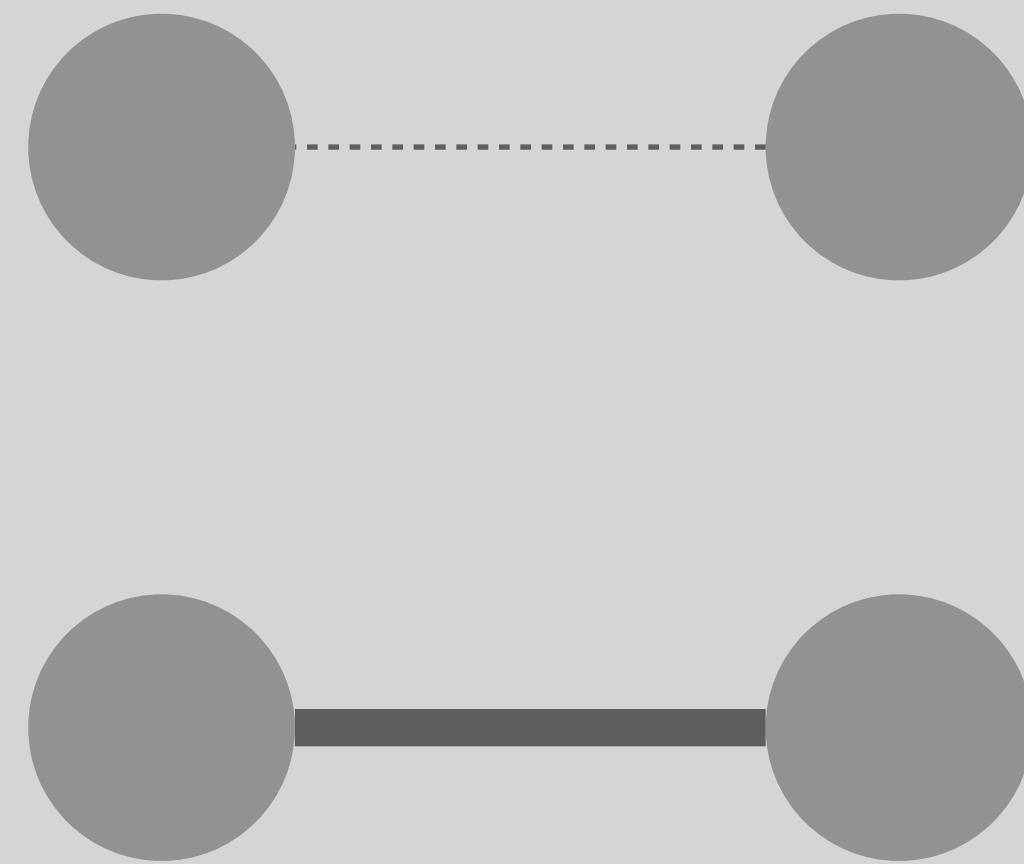






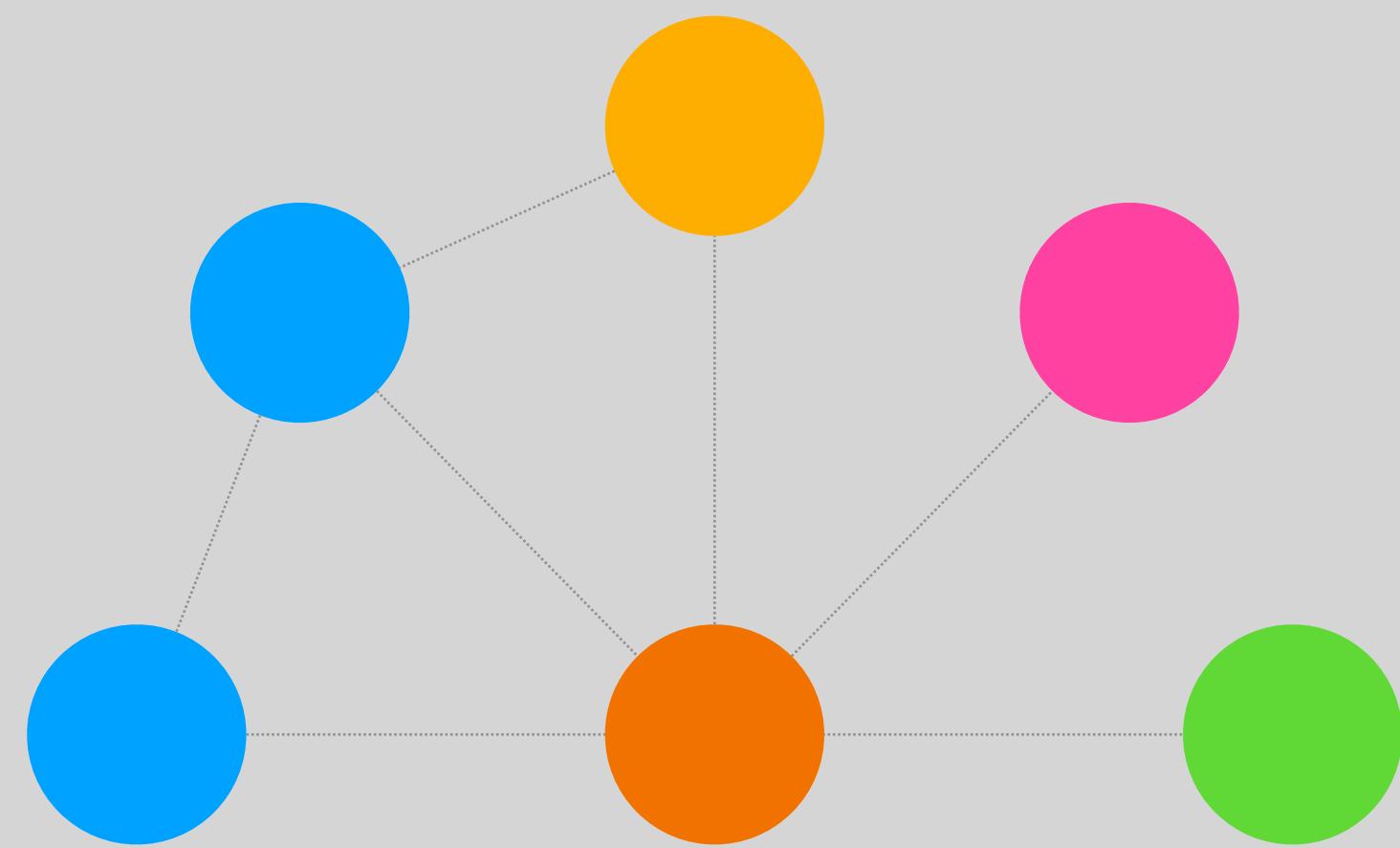
# Personal Networks

tie (strength)



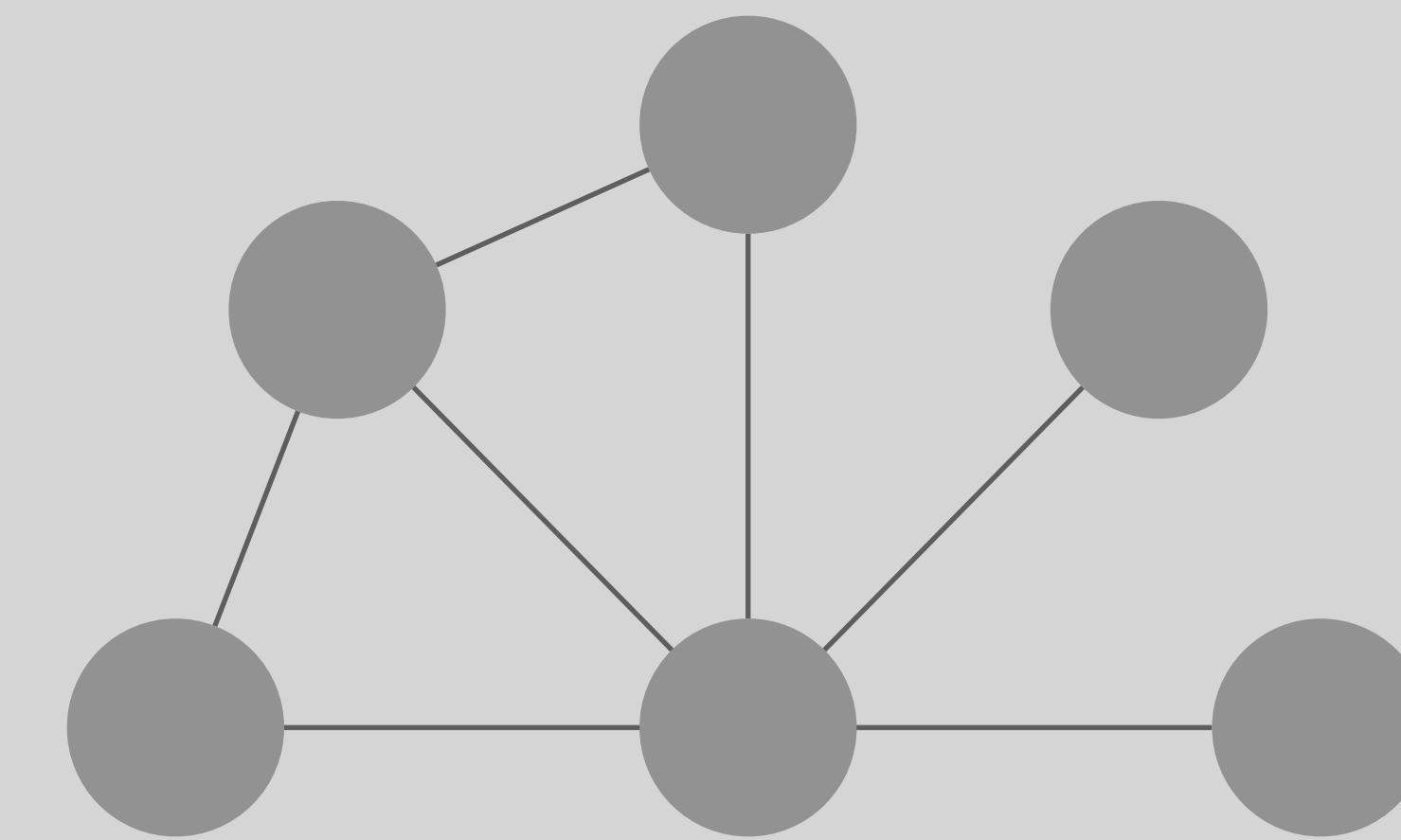
strong tie, more support/pressure  
e.g., quality of relation with parent

composition



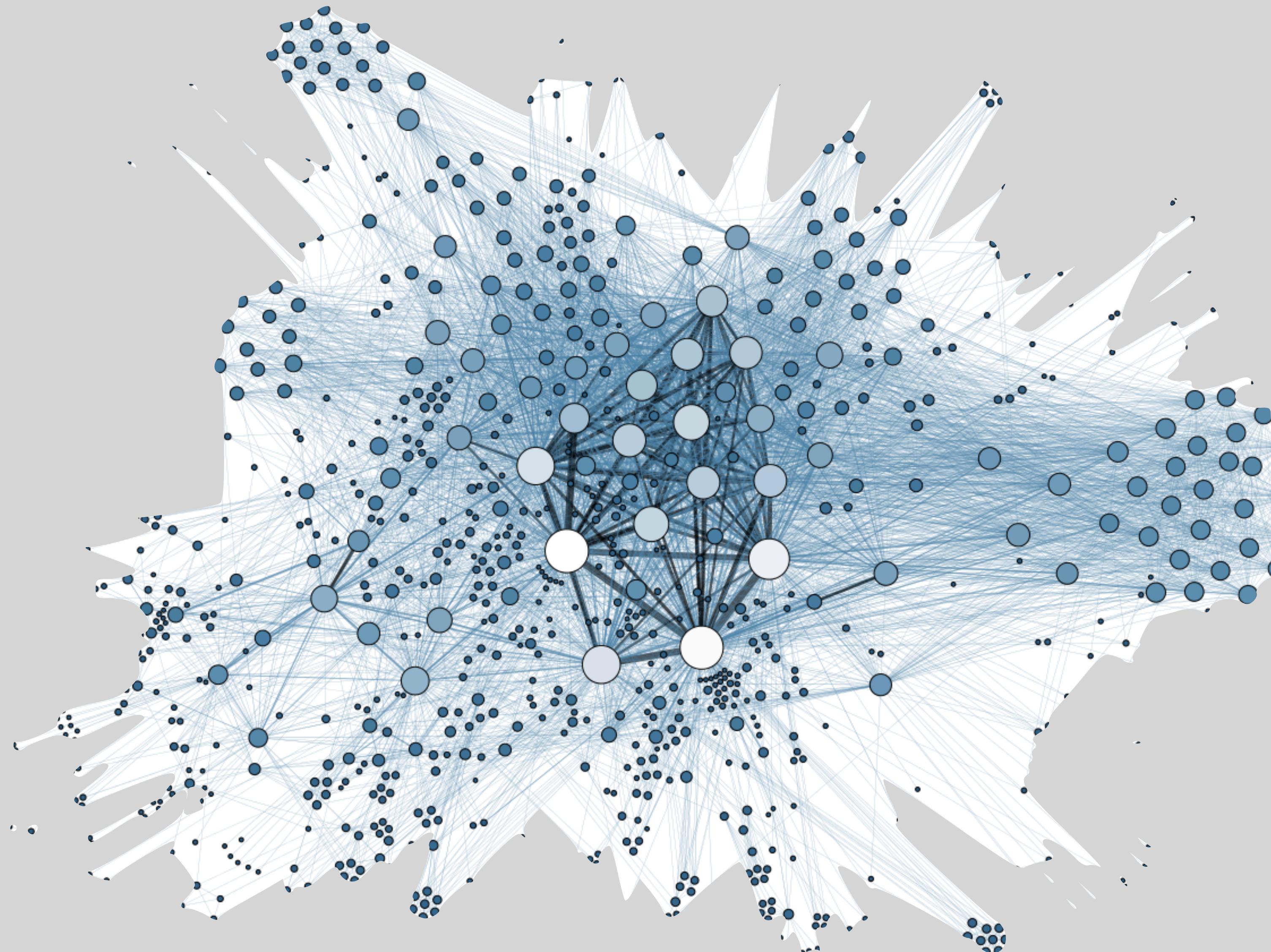
support network, diversity in ideas  
e.g., # kin, # friends, # can help

structure



reinforcing norms, flow information  
e.g., density, # cliques

# Network size



weak ties  
structure

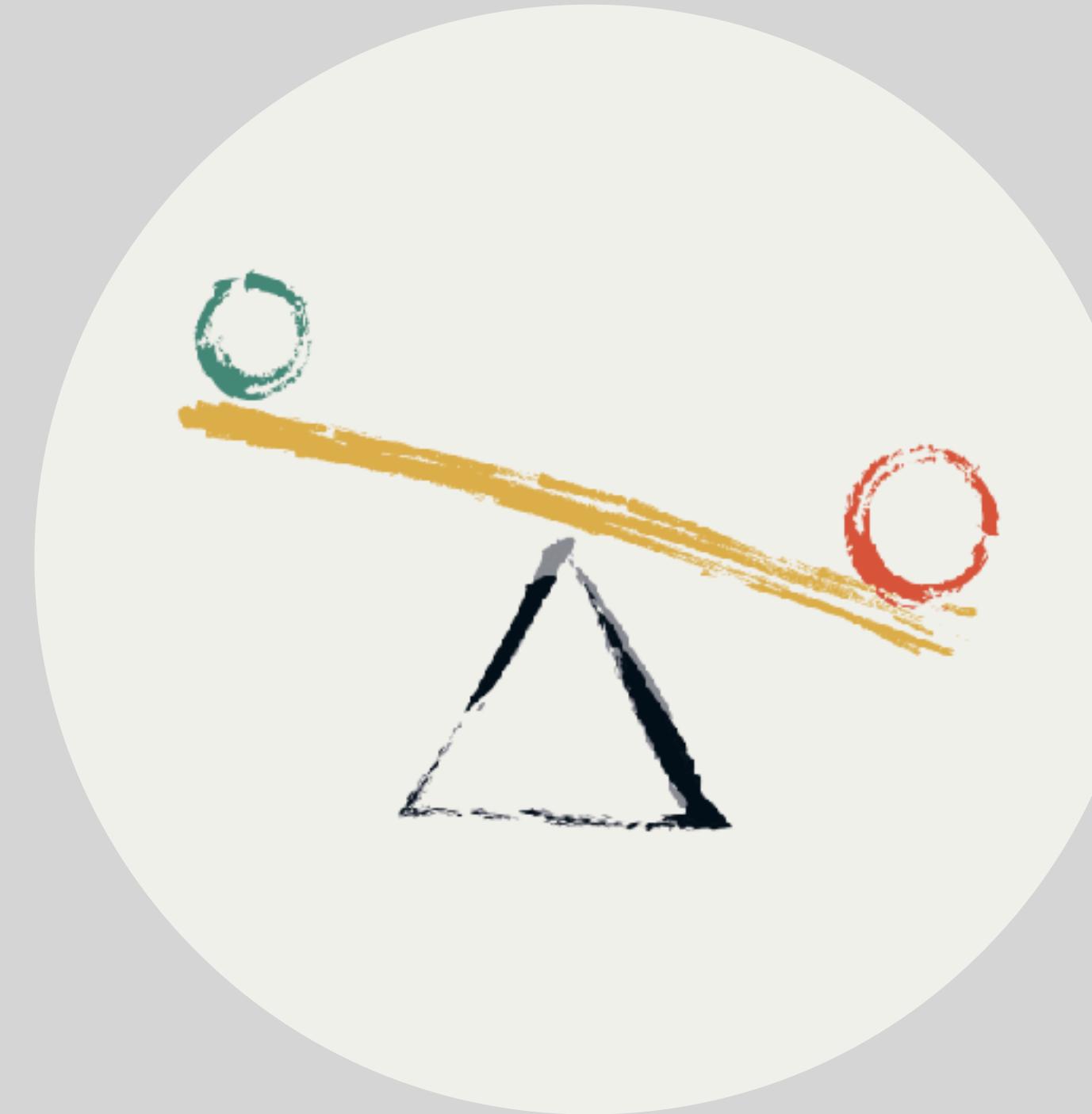
# The Right Answer is 25

scientific interest

weak ties

network structure

network composition



respondent burden

time

boredom

poor(er) response

# Methodology

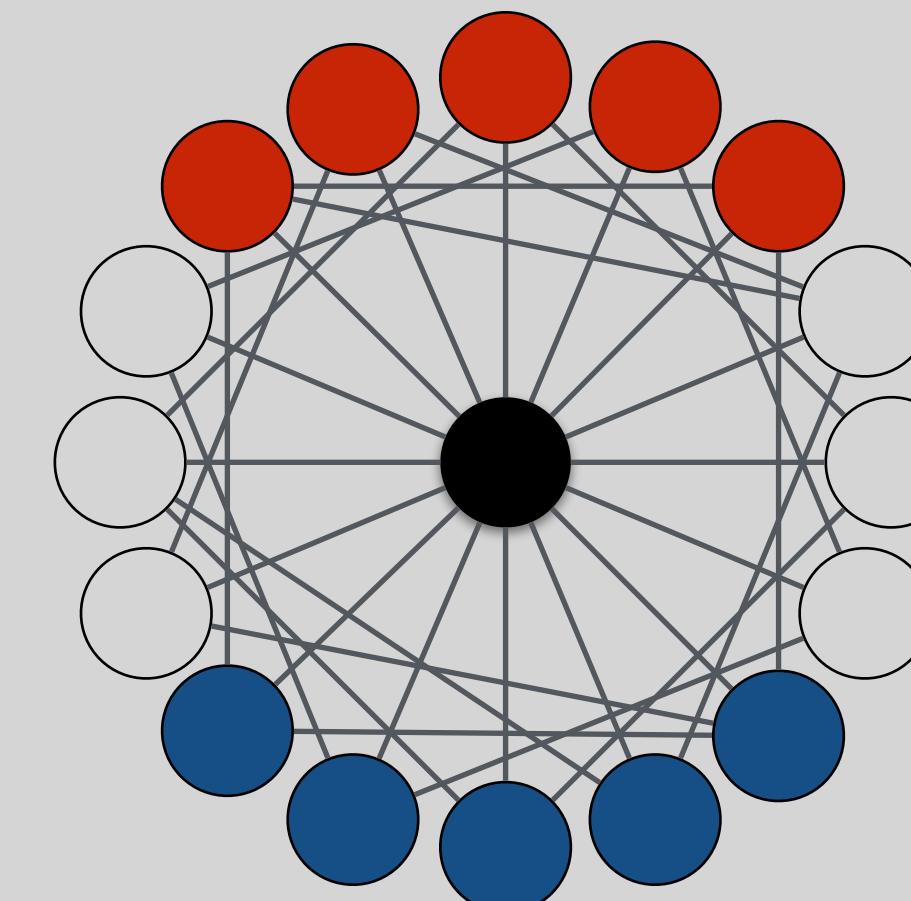


Longitudinal Internet  
Studies for the Social  
sciences

~750 women  
age: 18 - 40

Ego  
●

Alters (25)



Age	
Education	
Income	
Partnership status	
# Children	
Detailed fertility preferences	
Sex	Number and age of children
Age	Friend
Education	Wants children
Relationship type	Does not want children
Closeness	Help with children
Frequency of contact F2F	Talk about children
Frequency of other contact	Relationship with other alters

# Methodology

Please list 25 names of individuals 18 years or older with whom you have had contact in the last year. This can be face-to-face contact, but also contact via phone, internet, or email. You know these people and these people also know you from your name or face (think of friends, family, acquaintances, et cetera). You could reach out to these people if you would have to. Please name your partner in case you have one.

The image shows a digital form for listing contacts. At the top, there are two buttons: 'Naam' (Name) and 'Voeg toe' (Add). Below these are 25 light blue circular input fields, each containing a number from 1 to 25, arranged in a roughly triangular pattern. The numbers are: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25. At the bottom right of the form is a button labeled 'Ga door' (Continue).

# Methodology

Which of these 25 individuals could you ask for help

How close are you to these people?

Als het gaat om ANNE

Met wie heeft ANNE contact? Met contact bedoelen we alle vormen van contact, zoals face-to-face contact, contact via (mobiele) telefoon, post, email, sms, en andere manieren van online en offline communicatie.

Selecteer de personen die contact met elkaar hebben door met de muis op het bolletje te klikken. Er zal een lijn ontstaan die aangeeft dat de personen contact met elkaar hebben. Druk nogmaals op het bolletje om de lijn weer te laten verdwijnen, als de personen geen contact met elkaar hebben.

Anne

# Methodology

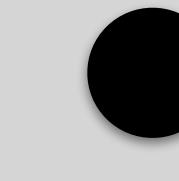


Longitudinal Internet  
Studies for the Social  
sciences

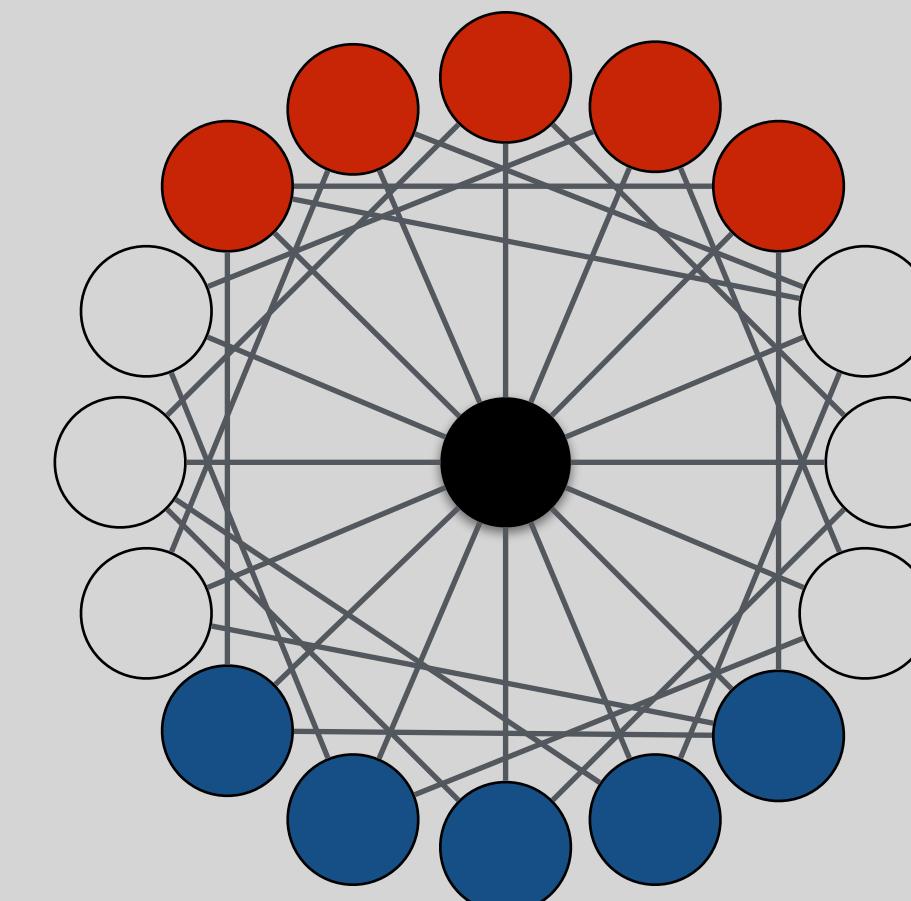


~750 women  
age: 18 - 40

Ego



Alters (25)



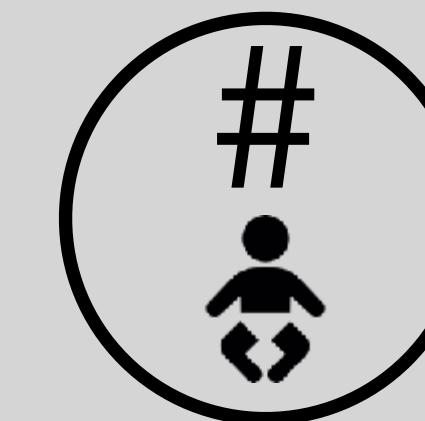
Age  
Education  
Income  
Partnership status  
# Children  
Detailed fertility preferences

## OUTCOMES

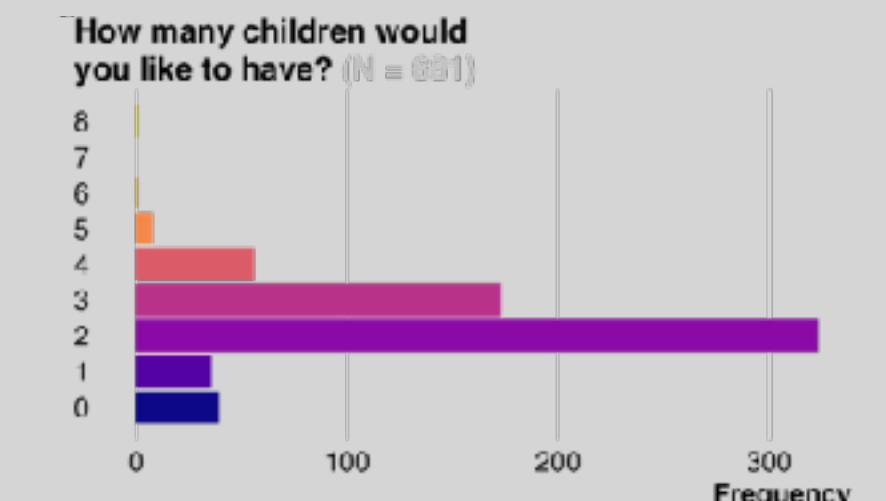
Sex  
Age  
Education  
Relationship type  
Closeness  
Frequency of contact F2F  
Frequency of other contact

Number and age of children  
Friend  
Wants children  
Does not want children  
Help with children  
Talk about children  
Relationship with other alters

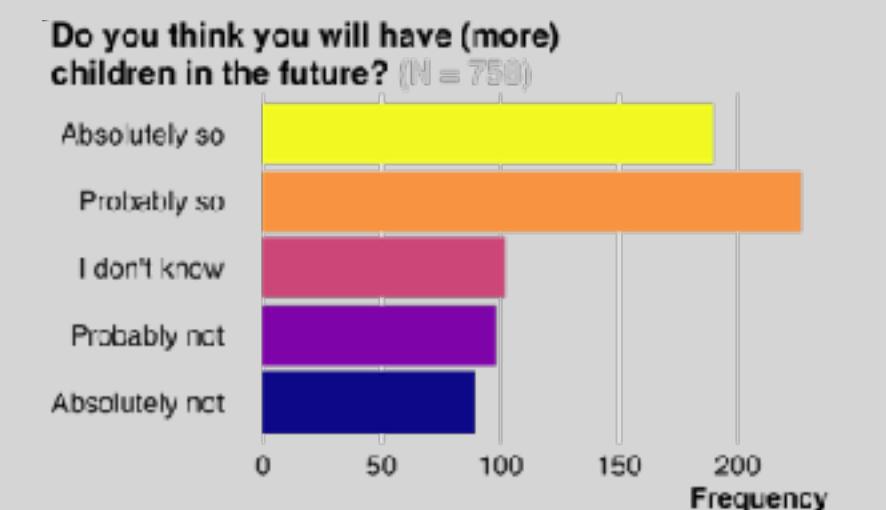
# Outcomes



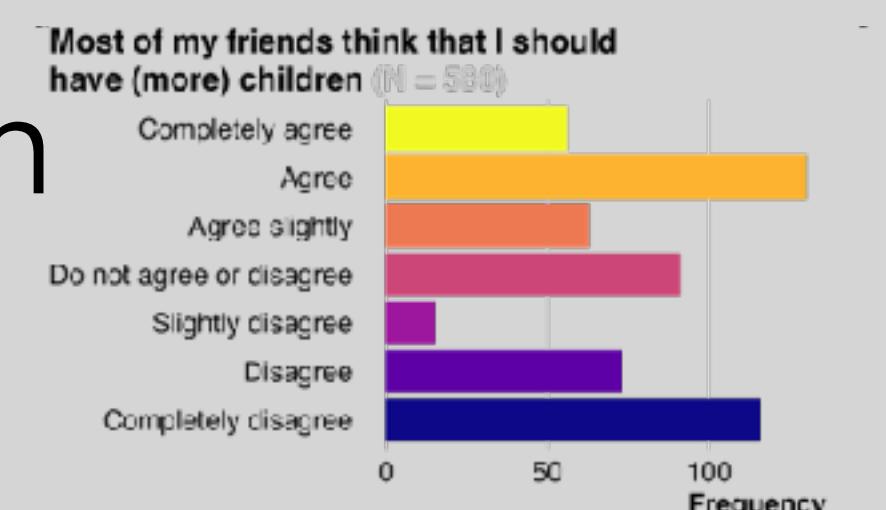
How many children would you like to have?



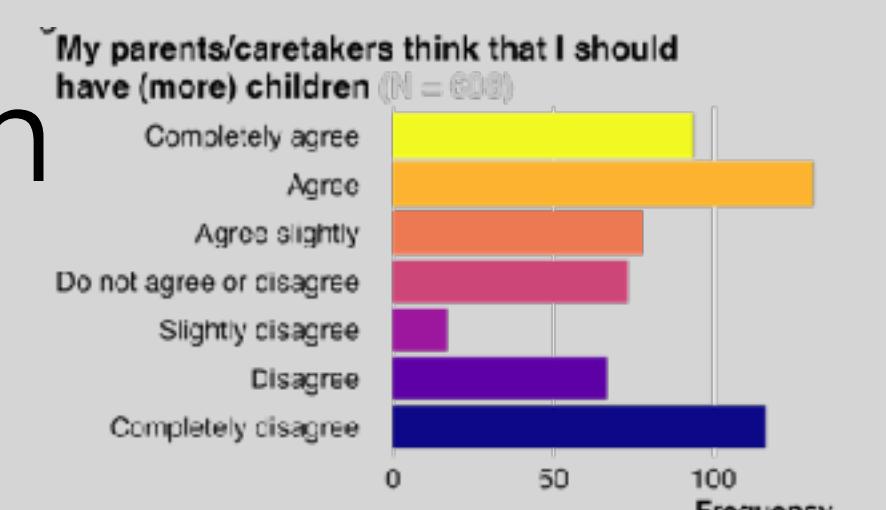
Do you think you will have (more) children in the future?



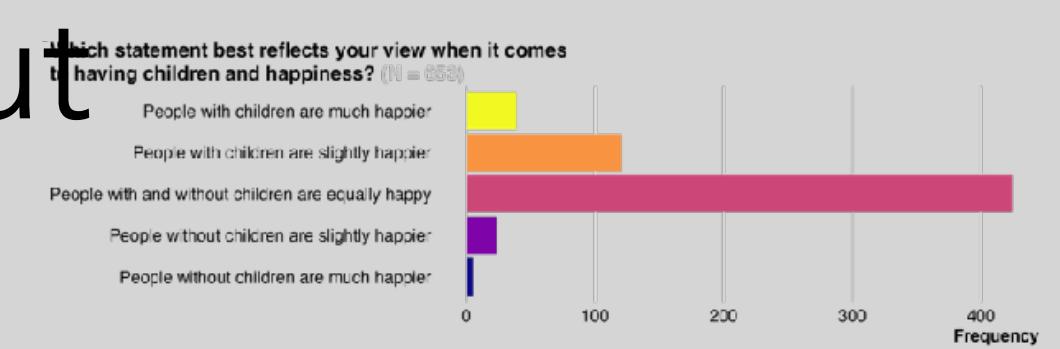
Perceived pressure to have children from friends



Perceived pressure to have children from parents/caretakers



Do you think people with or without children are happier?



# Methodology



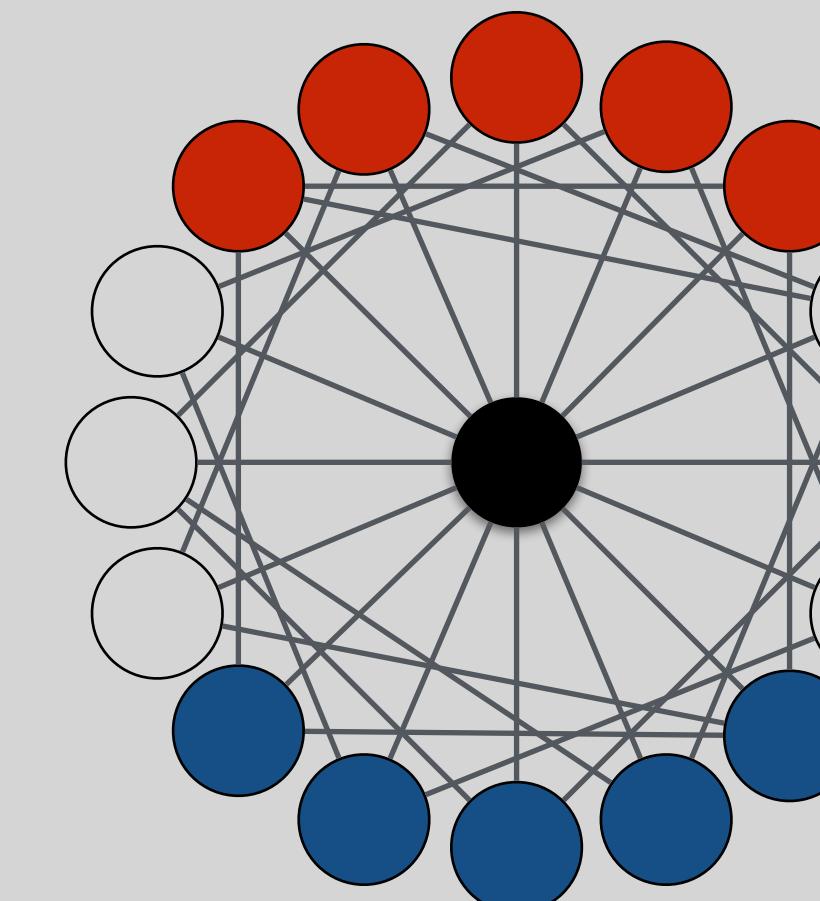
Longitudinal Internet  
Studies for the Social  
sciences



~750 women  
age: 18 - 40

Ego  
●

Alters (25)



## EGO VARIABLES

Age  
Education  
Income  
Partnership status  
# Children  
Detailed fertility preferences

## NETWORK VARIABLES

Sex	Number and age of children
Age	Friend
Education	Wants children
Relationship type	Does not want children
Closeness	Help with children
Frequency of contact F2F	Talk about children
Frequency of other contact	Relationship with other alters

# Personal Networks



tie (strength)

average closeness  
average f2f contact  
average other contact

average closeness **family**  
average closeness **friends**  
average closeness **childfree**

...

composition

% **family**  
% **friends**  
% **childfree**  
% with children  
% who want children  
% childfree  
% highly educated  
% women  
% can provide childcare  
% can talk to about children

...

structure

density  
# cliques  
# isolates and duos  
# communities  
modularity  
degree centralisation  
betweenness centralisation

...

density among **family**  
density among **friends**  
density among **childfree**

...

24 variables

13 variables

20 variables

# Personal Networks



tie (strength)

average closeness  
average f2f contact  
average other contact

average closeness *family*  
average closeness *friends*  
average closeness *childfree*

...

composition

% *family*  
% *friends*

...

HOW TO CHOOSE  
WHICH VARIABLES  
TO FOCUS ON?

structure

density  
# cliques  
and duos  
unities  
y  
centralisation  
less centralisation

density among *family*  
density among *friends*  
density among *childfree*

...

24 variables

13 variables

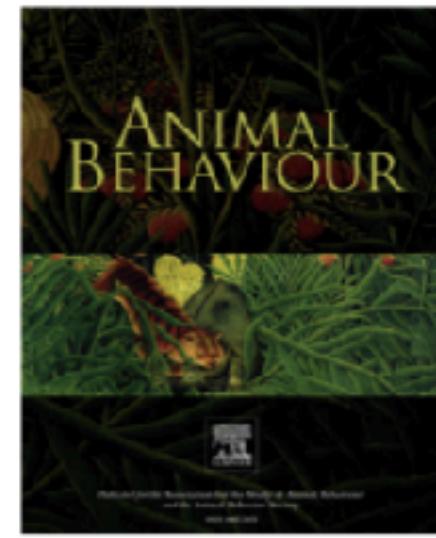
20 variables



Contents lists available at [ScienceDirect](#)

## Animal Behaviour

journal homepage: [www.elsevier.com/locate/anbehav](http://www.elsevier.com/locate/anbehav)



### Commentary

## Is less more? A commentary on the practice of ‘metric hacking’ in animal social network analysis



Quinn M. R. Webber <sup>a,\*</sup>, David C. Schneider <sup>a,b,c</sup>, Eric Vander Wal <sup>a,c</sup>

<sup>a</sup> Cognitive and Behavioural Ecology Interdisciplinary Program, Memorial University of Newfoundland, St John's, NL, Canada

<sup>b</sup> Department of Ocean Sciences, Ocean Sciences Centre, Memorial University of Newfoundland, St John's, NL, Canada

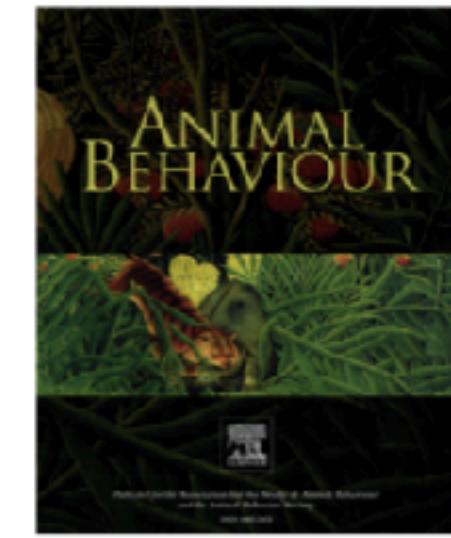
<sup>c</sup> Department of Biology, Memorial University of Newfoundland, St John's, NL, Canada



Contents lists available at [ScienceDirect](#)

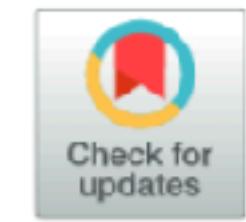
## Animal Behaviour

journal homepage: [www.elsevier.com/locate/anbehav](http://www.elsevier.com/locate/anbehav)



### Commentary

## Is less more? A commentary on the practice of ‘metric hacking’ in animal social network analysis



Quinn M. R. Webber <sup>a,\*</sup>, David C.

<sup>a</sup> Cognitive and Behavioural Ecology Interdisciplinary Program

<sup>b</sup> Department of Ocean Sciences, Ocean Sciences Centre, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

<sup>c</sup> Department of Biology, Memorial University of Newfoundland, St. John's, NF A1C 5S7, Canada



General Article

## False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Joseph P. Simmons<sup>1</sup>, Leif D. Nelson<sup>2</sup>, and Uri Simonsohn<sup>1</sup>

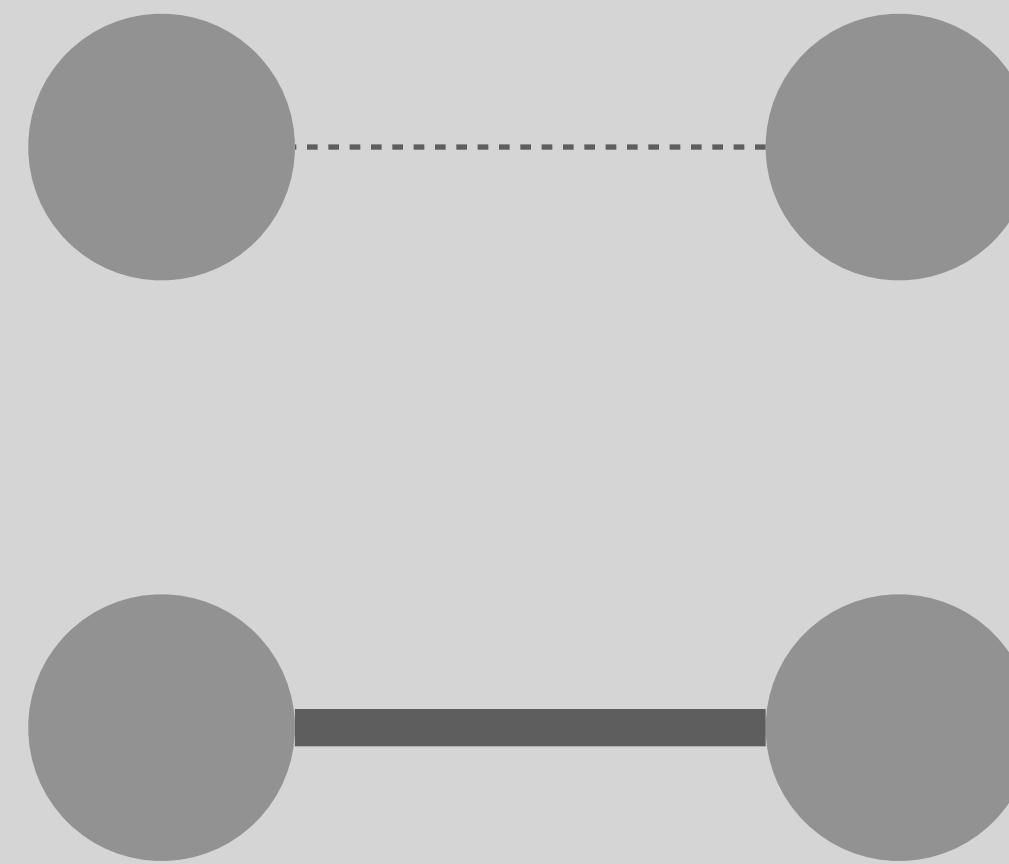
<sup>1</sup>The Wharton School, University of Pennsylvania, and <sup>2</sup>Haas School of Business, University of California, Berkeley



Psychological Science  
22(11) 1359–1366  
© The Author(s) 2011  
Reprints and permission:  
[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)  
DOI: 10.1177/0956797611417632  
<http://pss.sagepub.com>  
\$SAGE

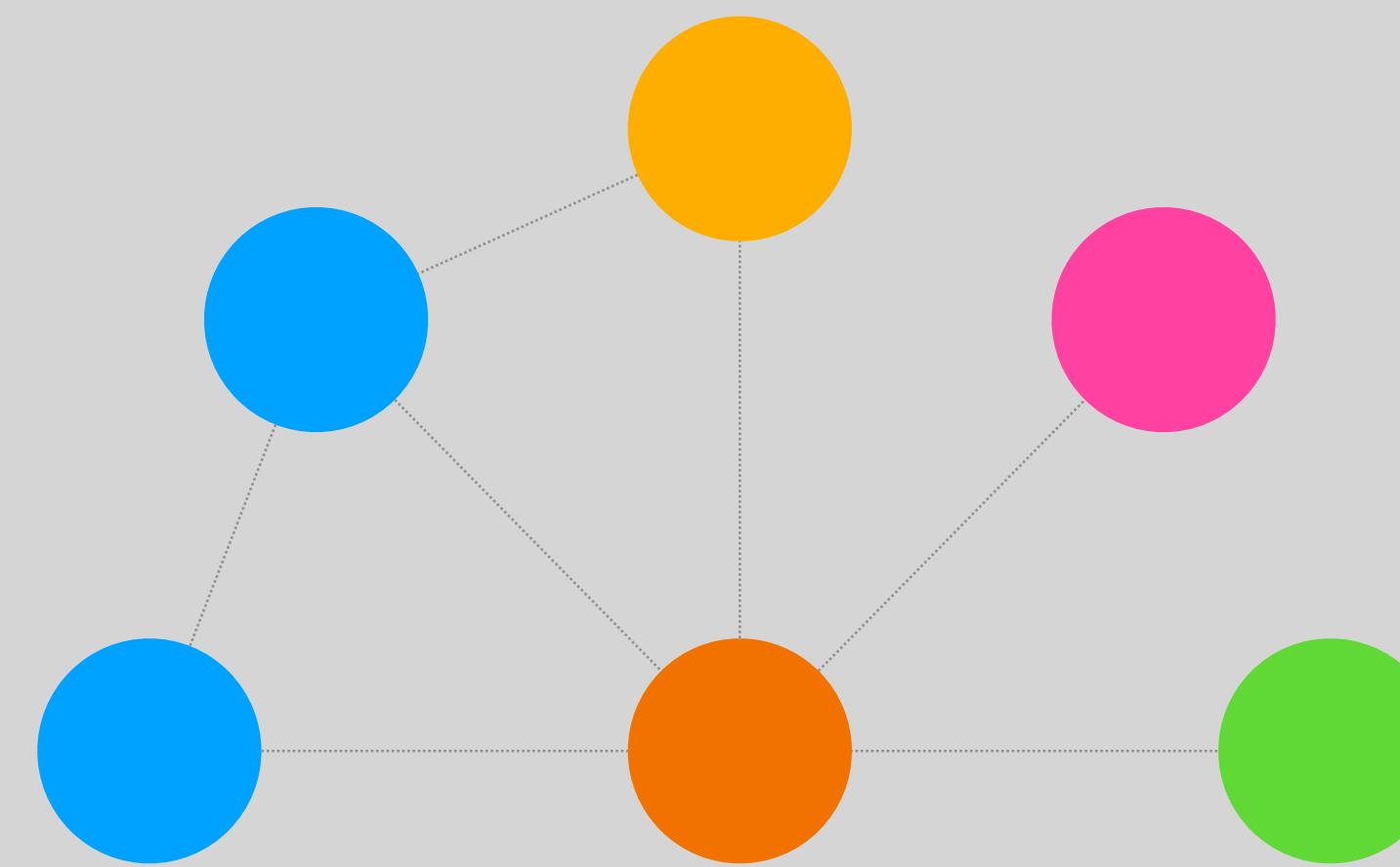
# Personal Networks

tie (strength)



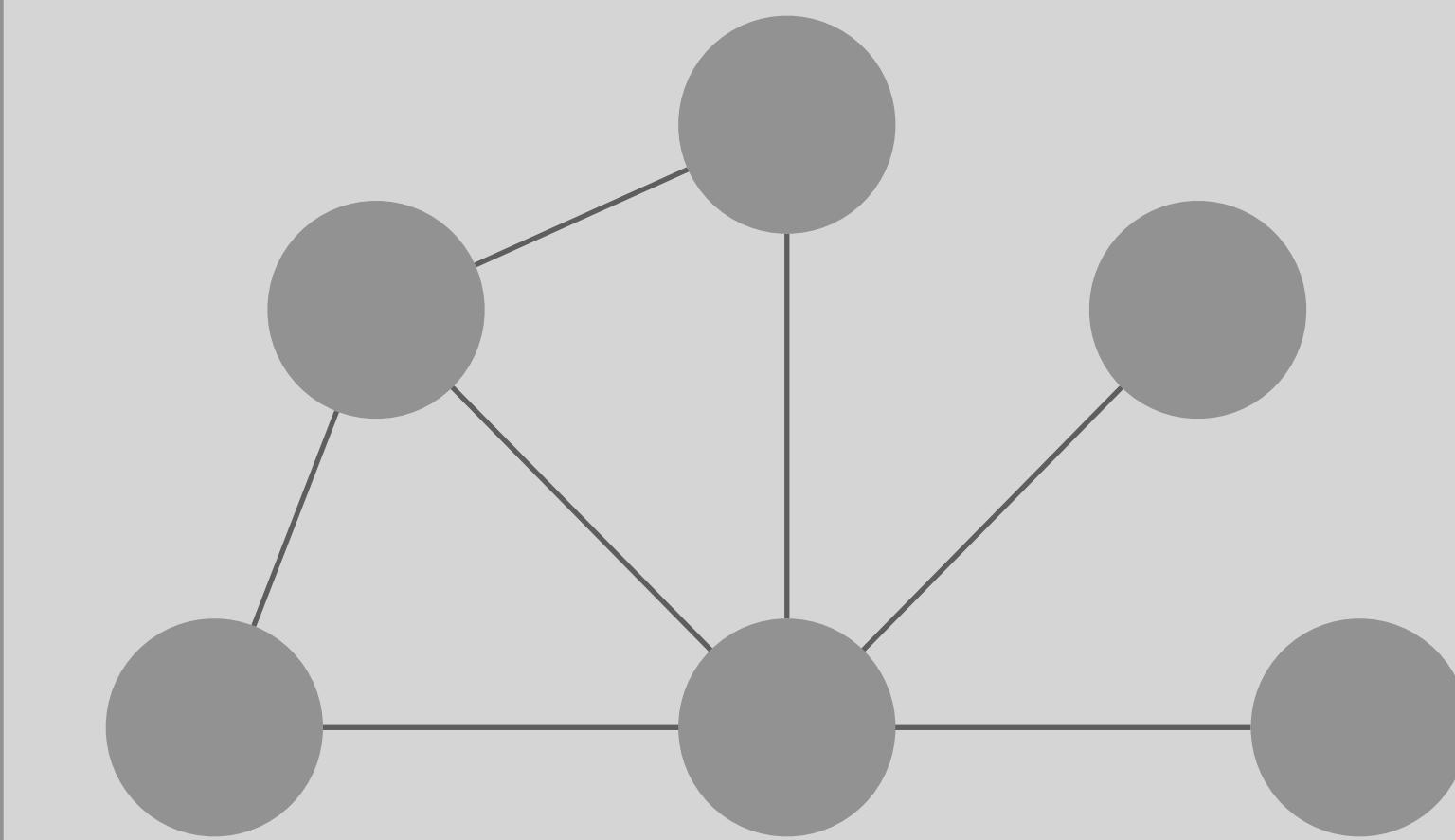
strong tie, more support/pressure  
e.g., quality of relation with parent

composition

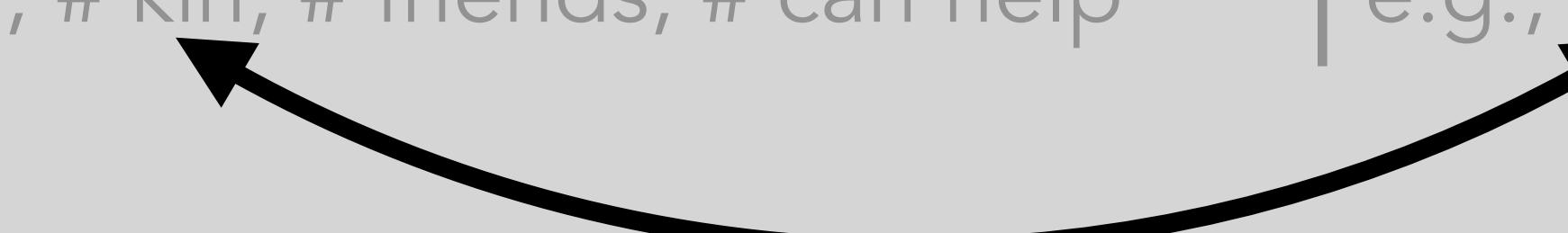


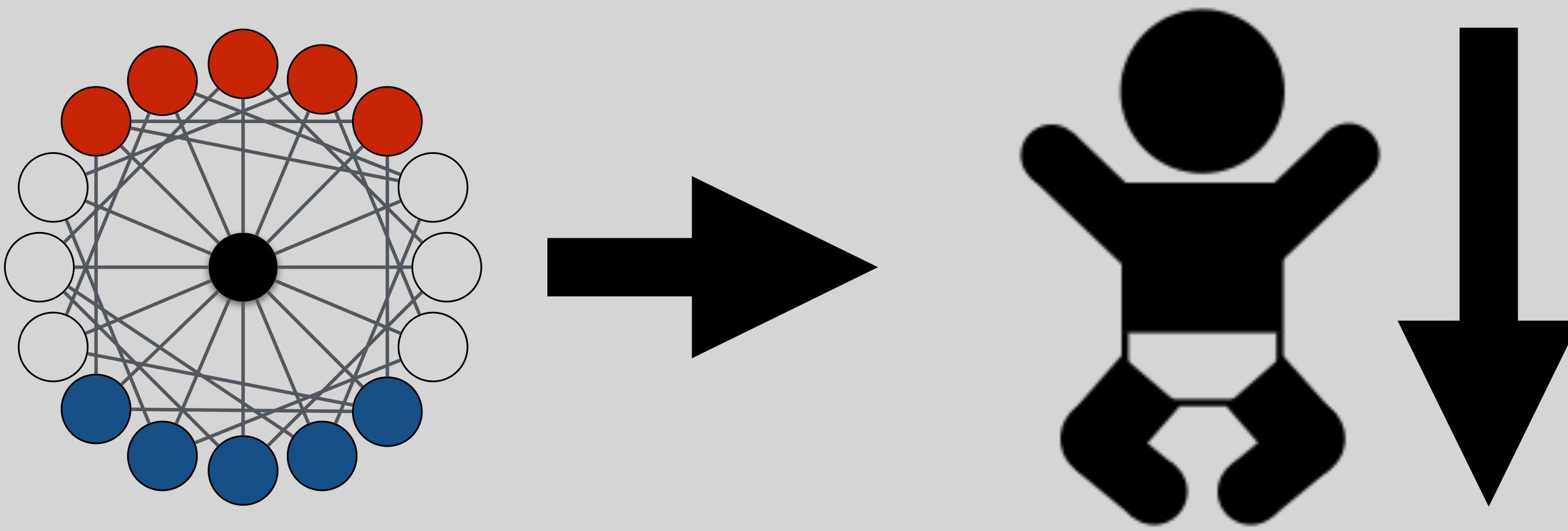
support network, diversity in ideas  
e.g., # kin, # friends, # can help

structure



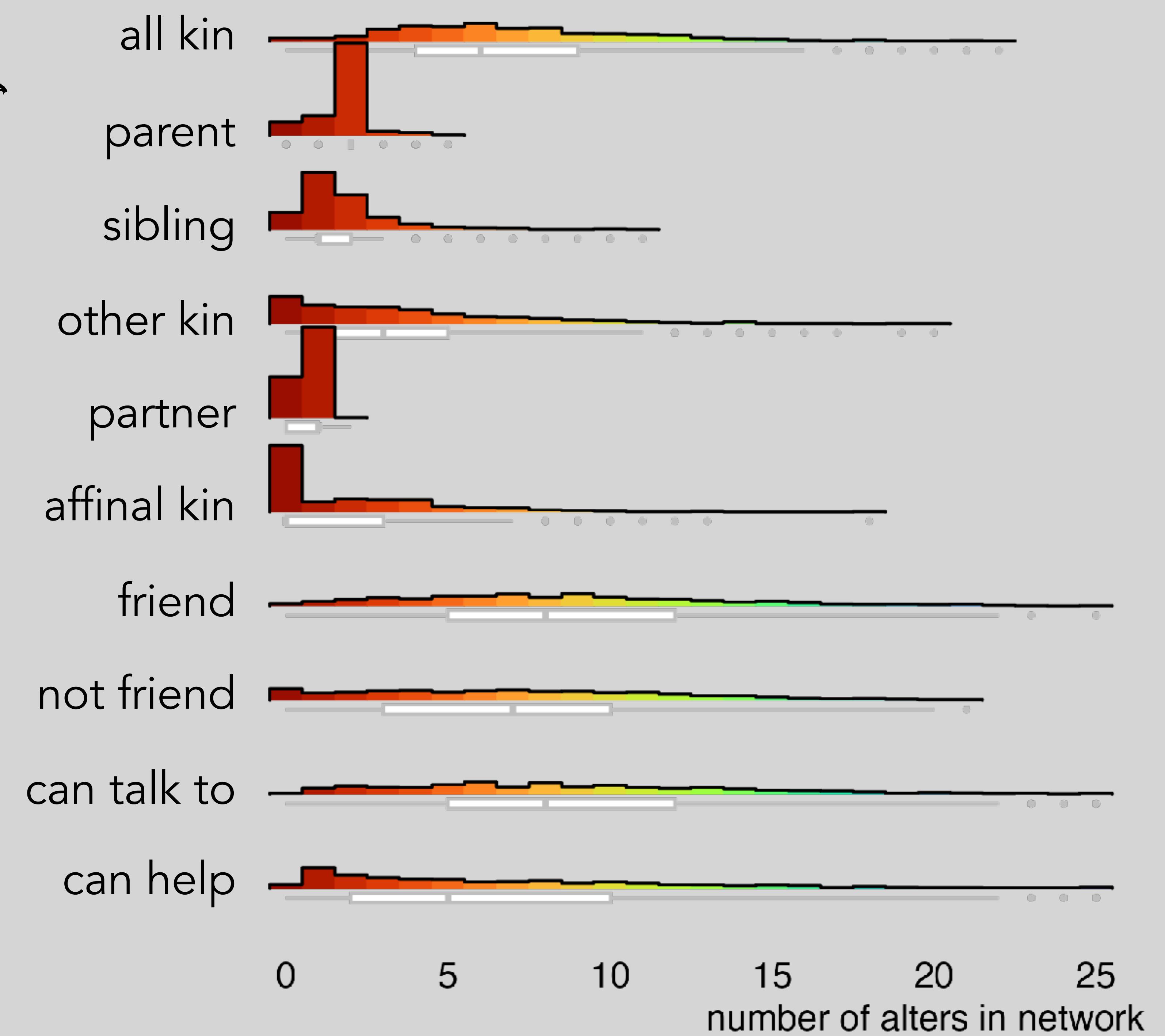
reinforcing norms, flow information  
e.g., density, # cliques



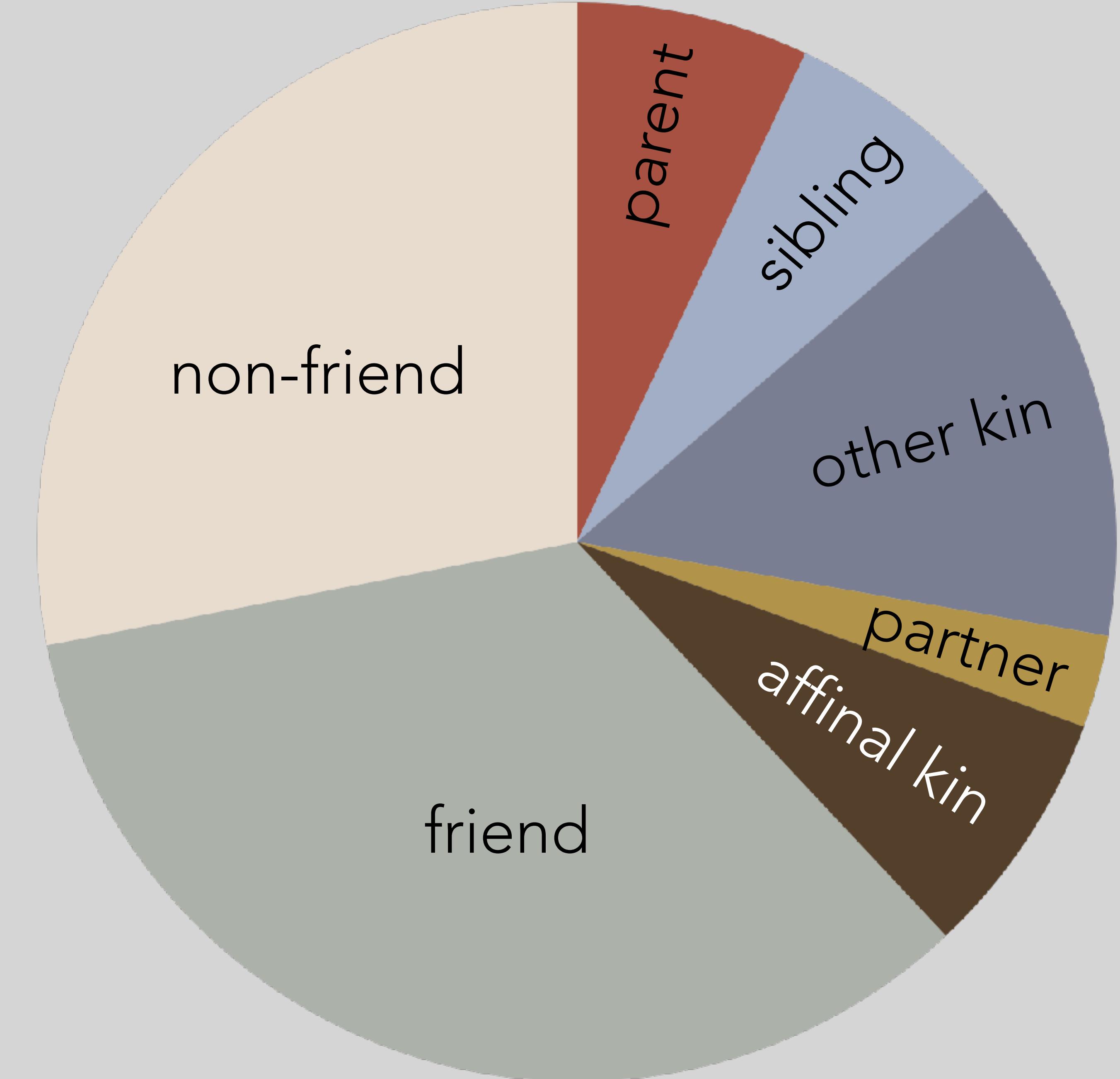


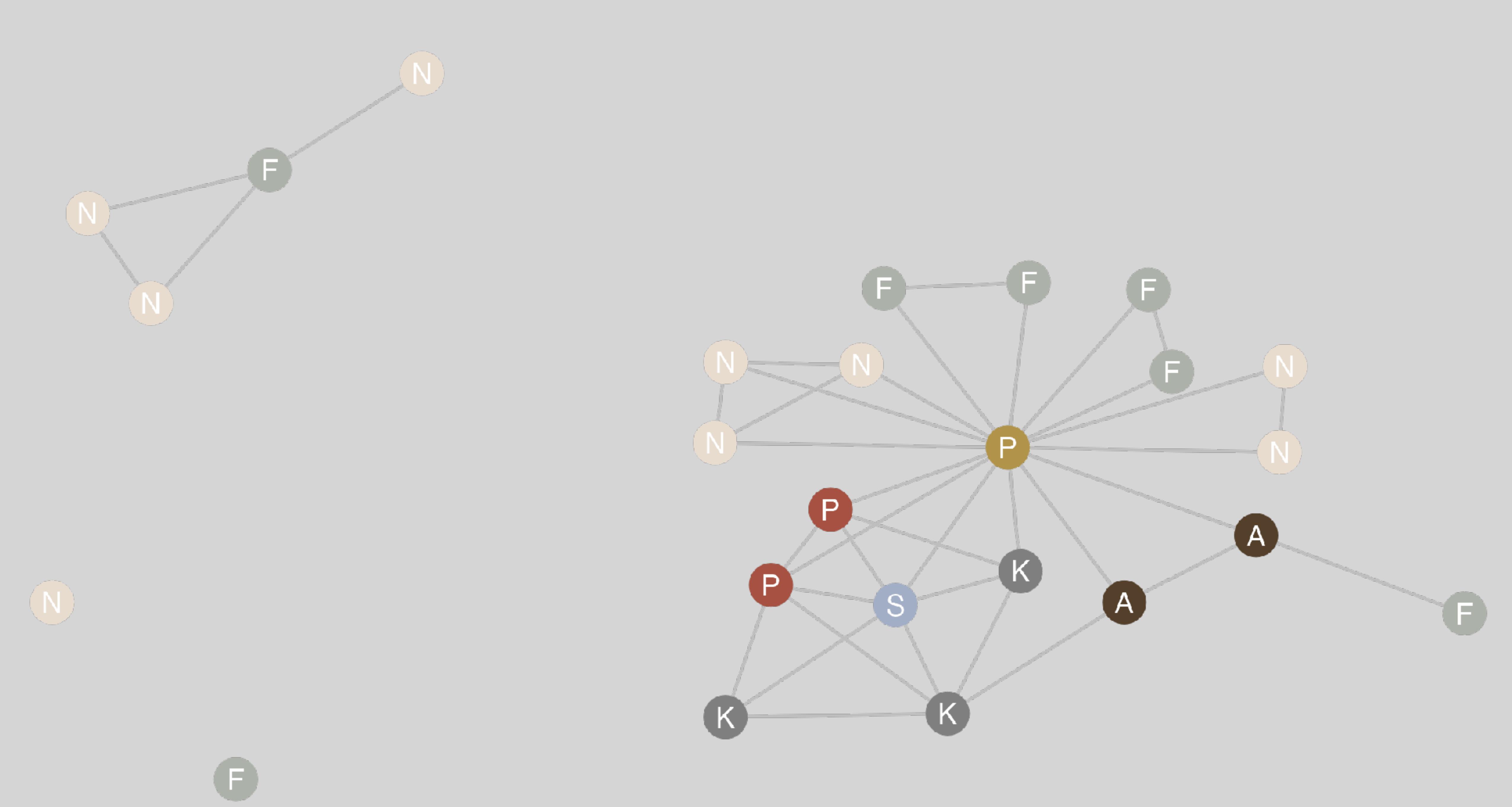
predicting fertility outcomes  
using personal network data

# Composition



kin make up a  
substantial fraction  
of the network

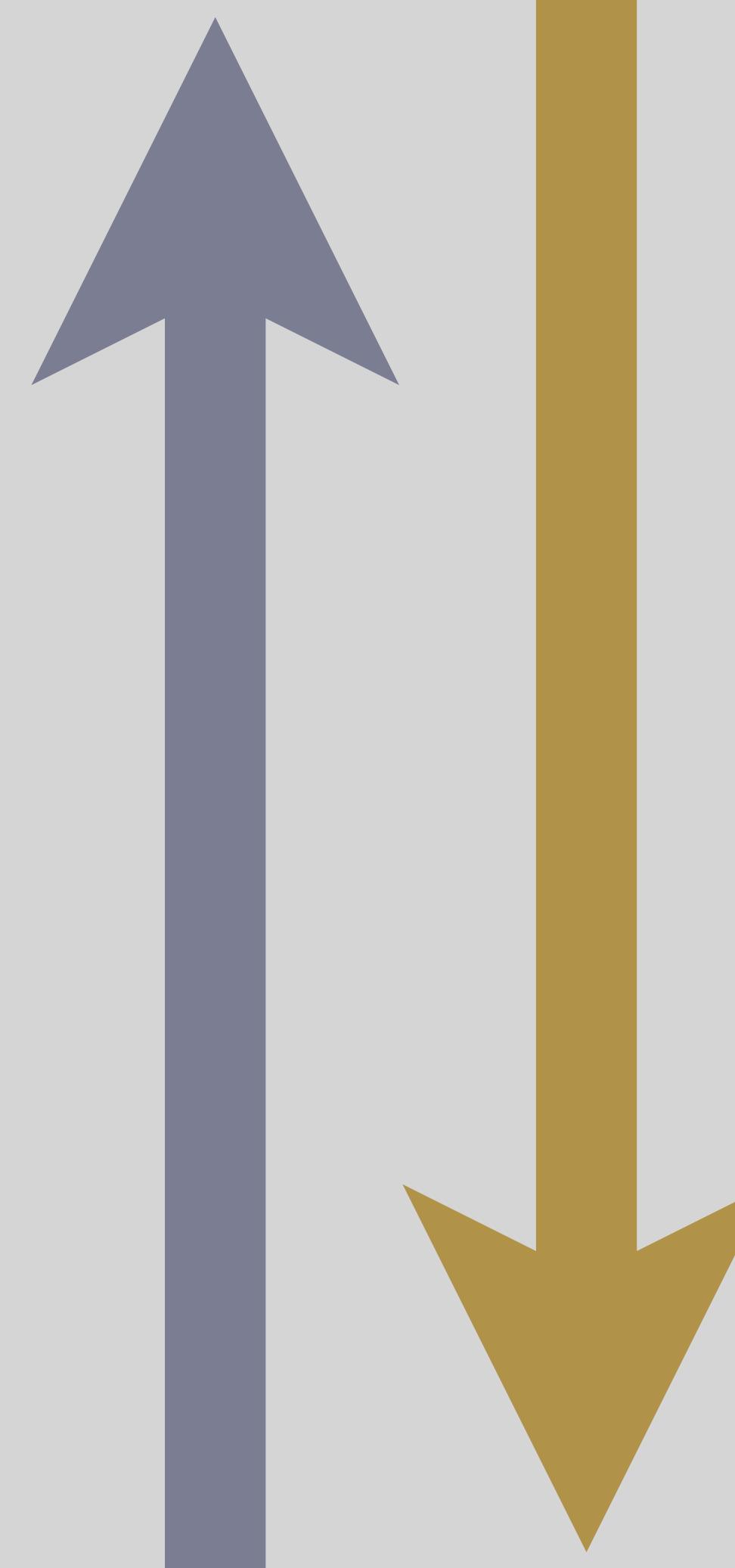






# INTERPRETABILITY

LASSO regression



XGBoost  
Support Vector Machines

Graph Neural Networks

# COMPLEXITY

# Lasso Regression

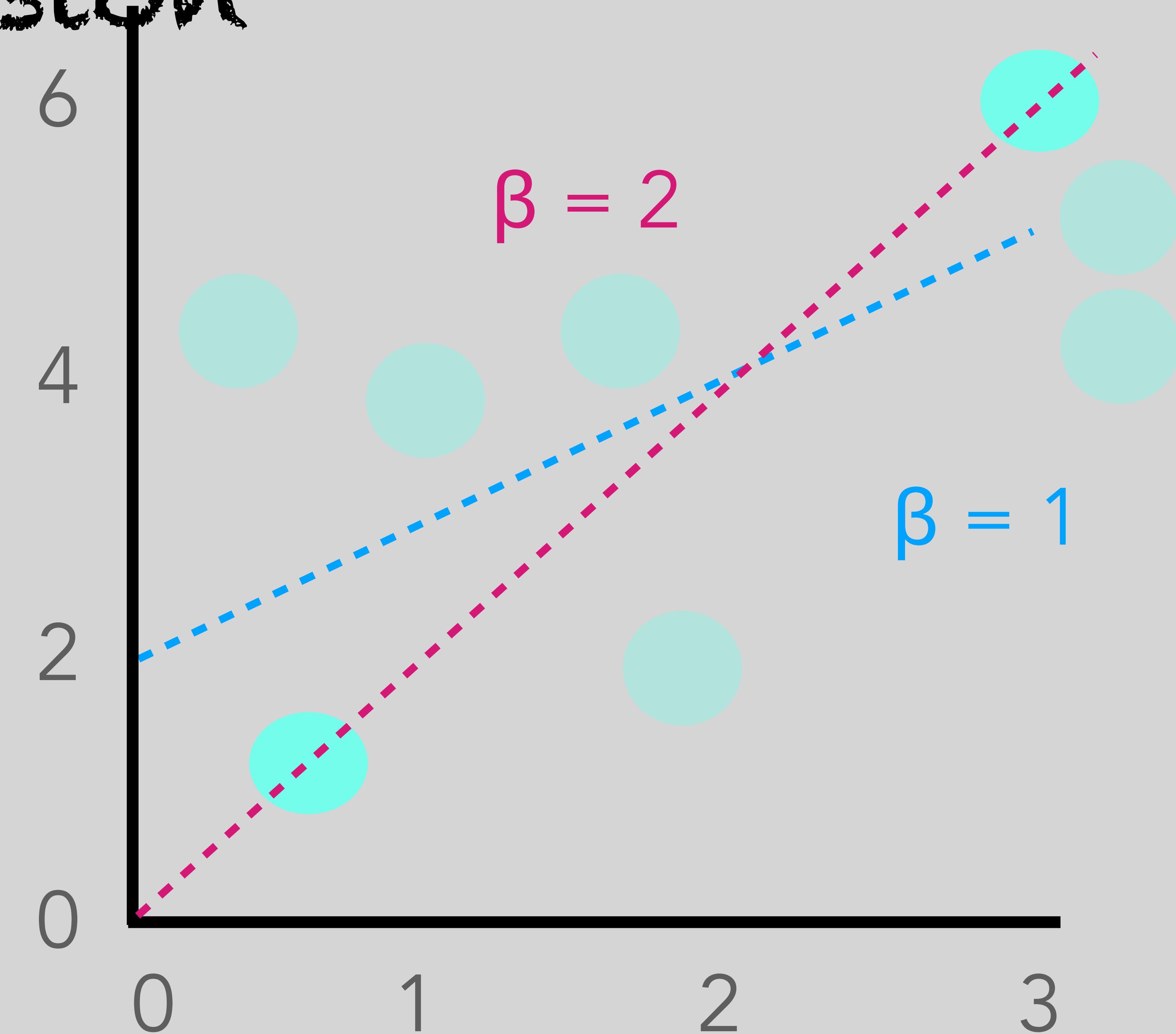
$$\sum_{i=0}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^p |\beta_j|$$


linear regression      penalty term

- ✓ can handle many, correlated variables
- ✓ leads to sparse, predictive, interpretable models
- (✗) reduced variance through increased bias

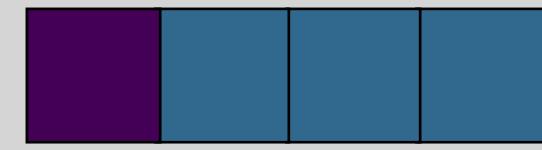
# LASSO Regression

$$\sum_{i=0}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^p |\beta_j|$$

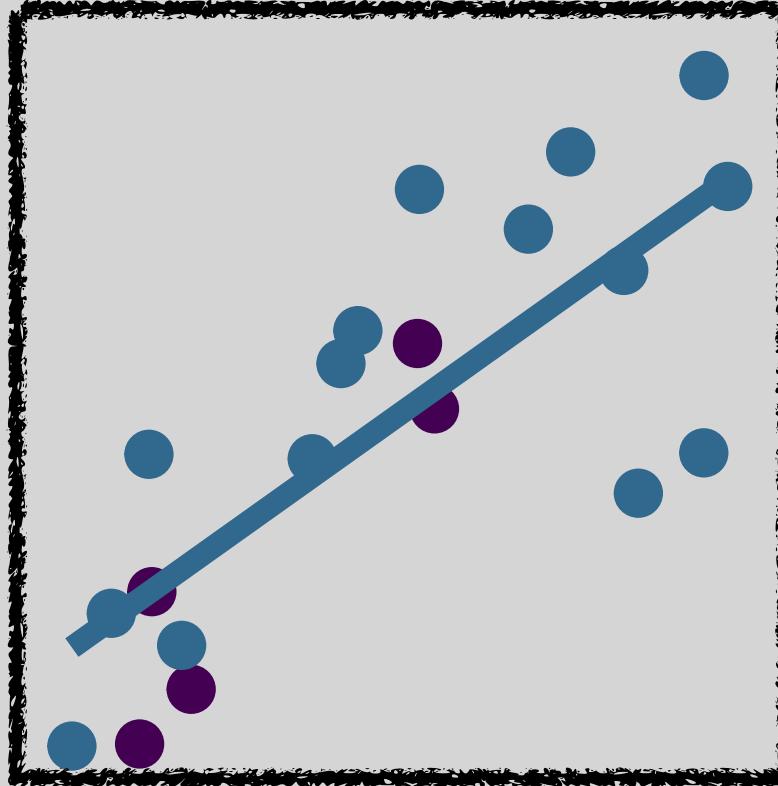


# Cross-Validation

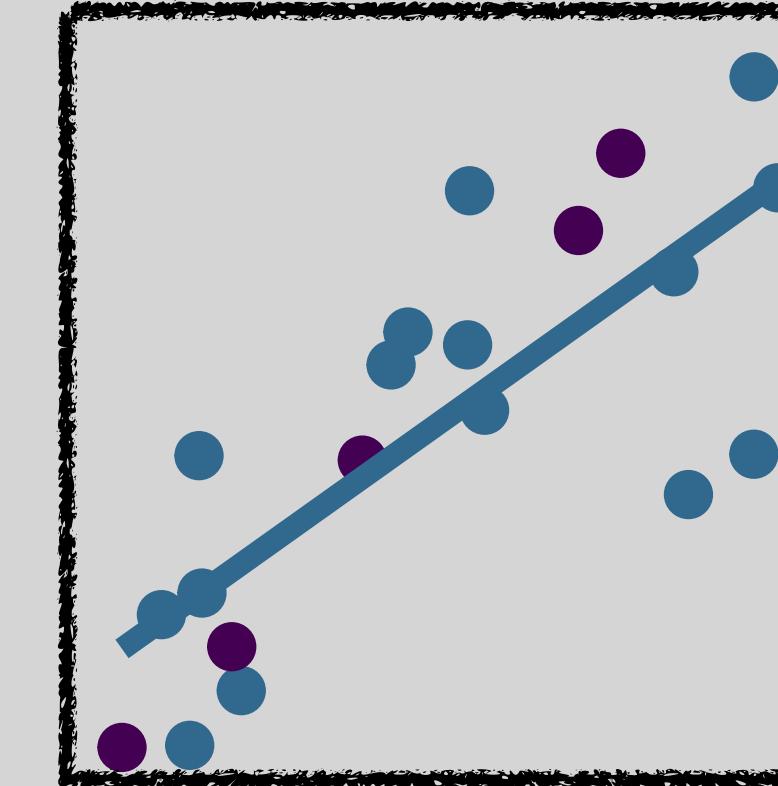
$\lambda$  is determined through cross-validation and **out-of-sample predictive ability**



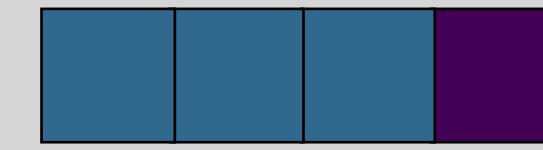
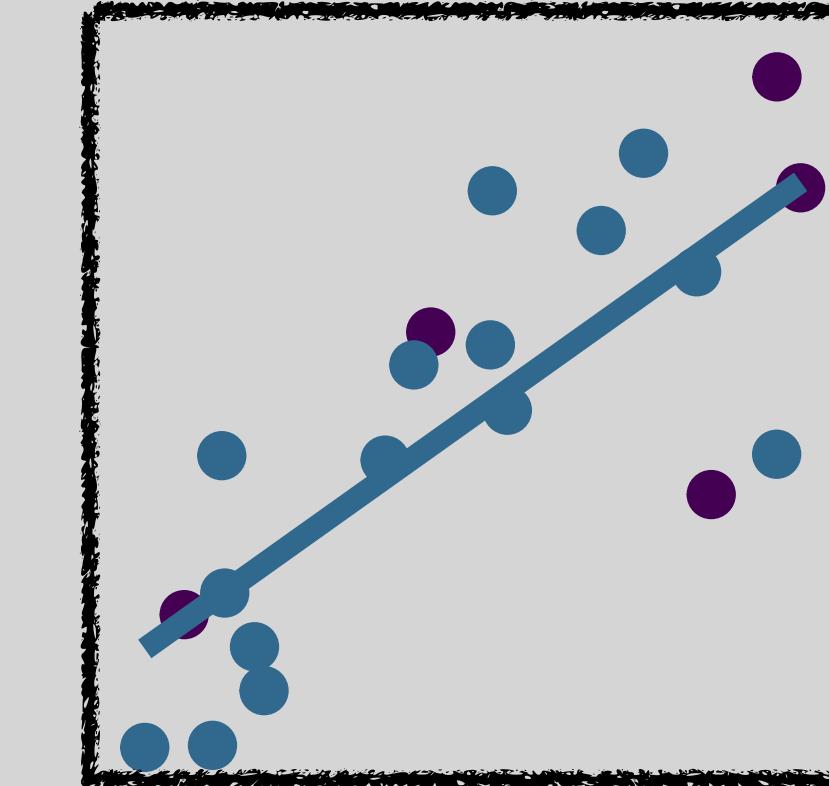
fold 1



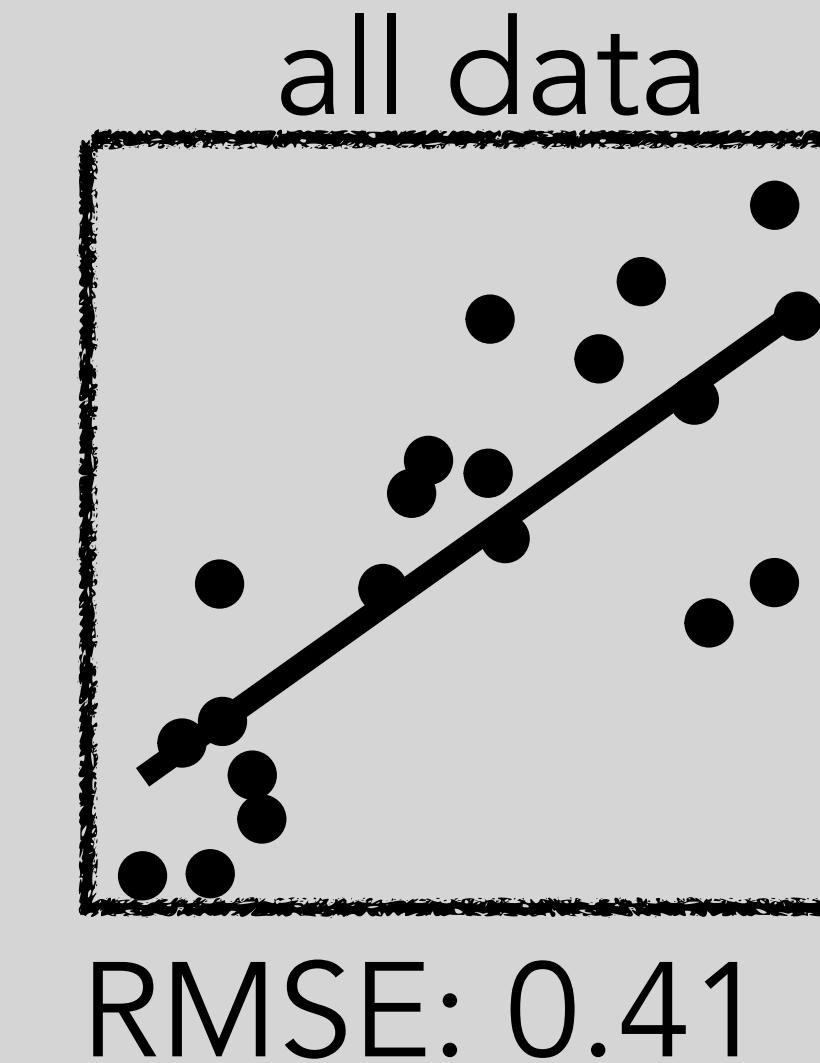
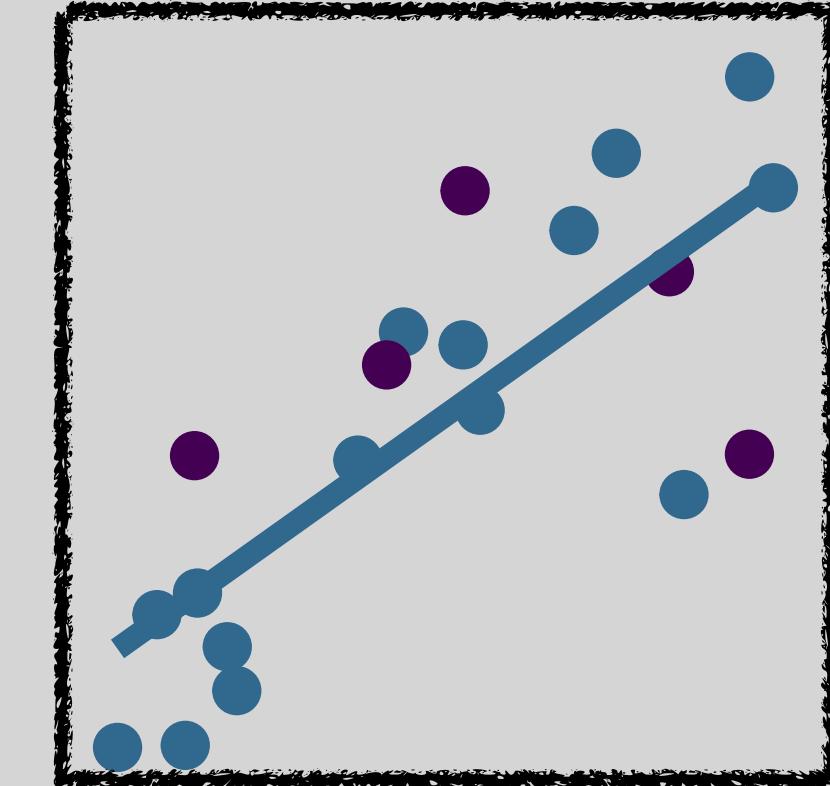
fold 2



fold 3



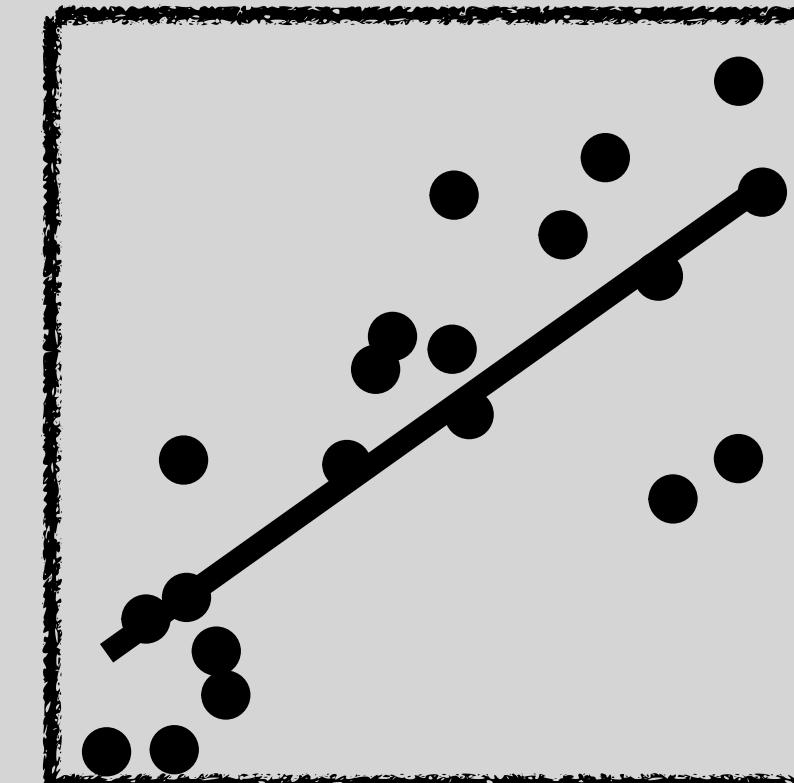
fold 4



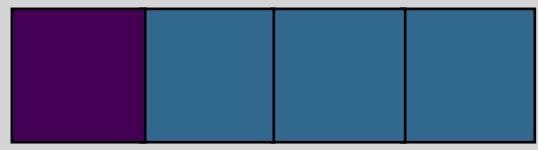
# Cross-Validation

strength of model  
**quantified by out-of-sample predictive ability**

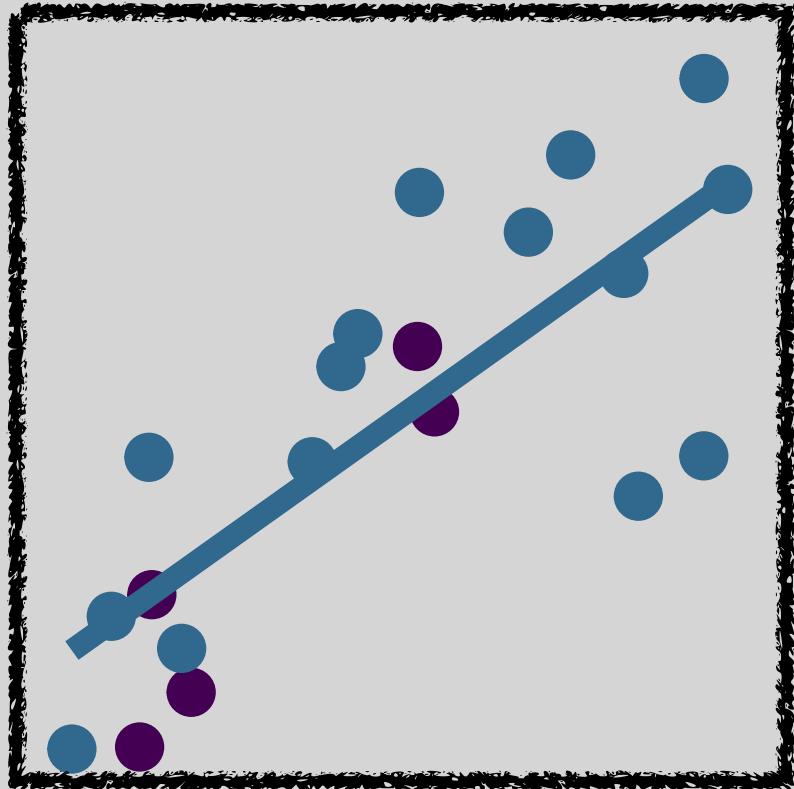
all data



RMSE: 0.41



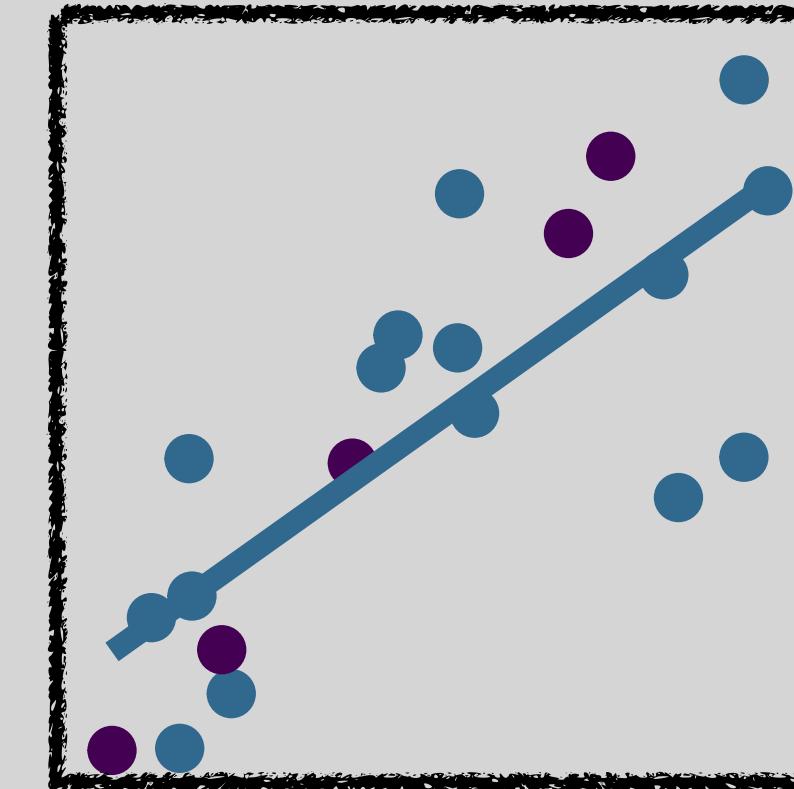
fold 1



RMSE: 0.38



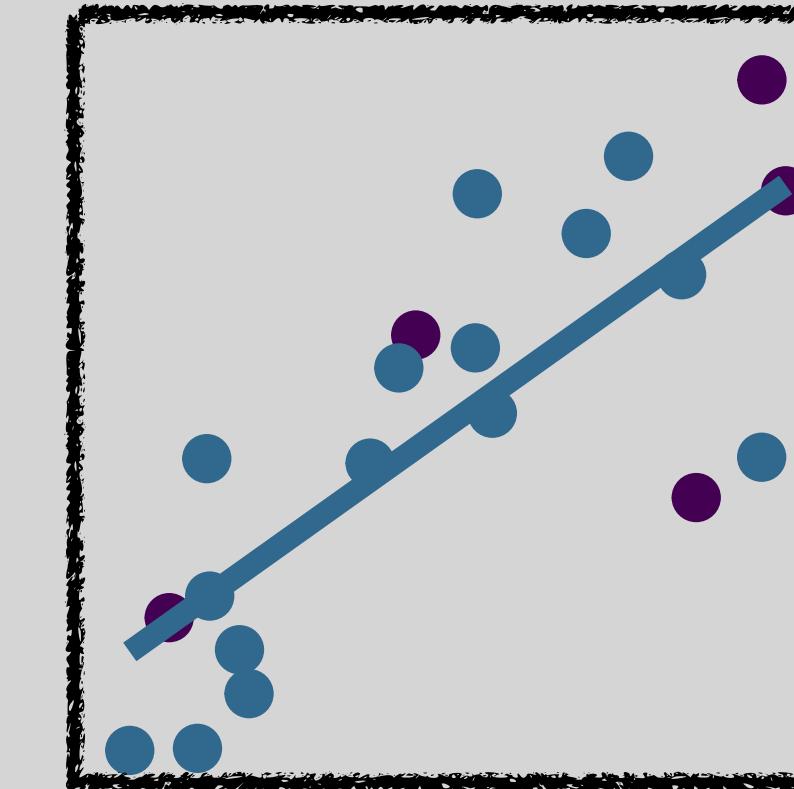
fold 2



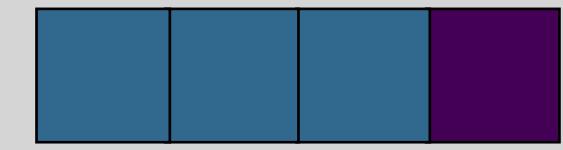
RMSE: 0.38



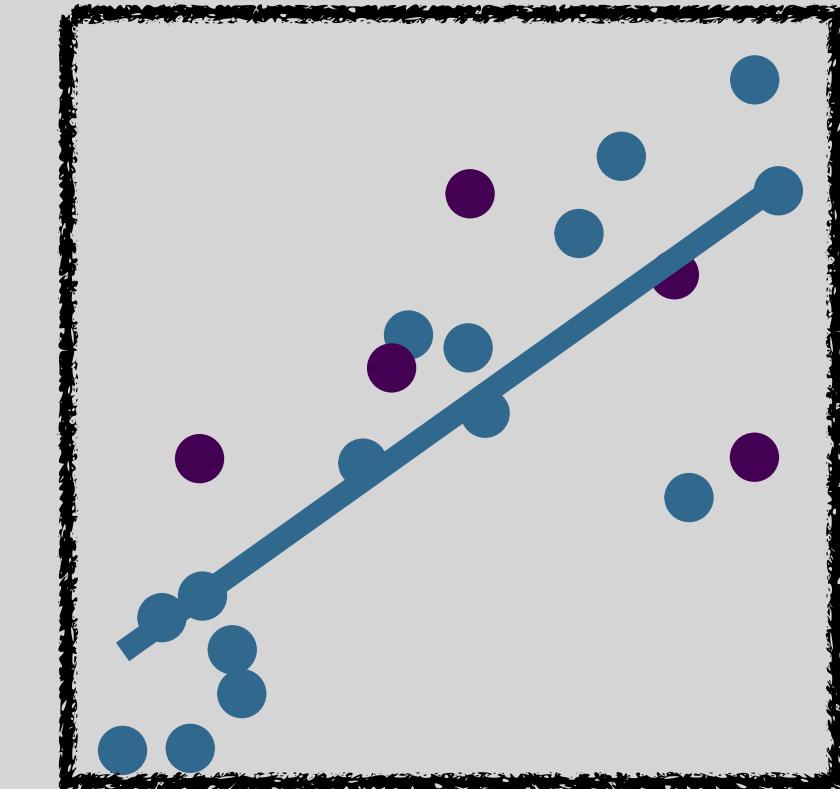
fold 3



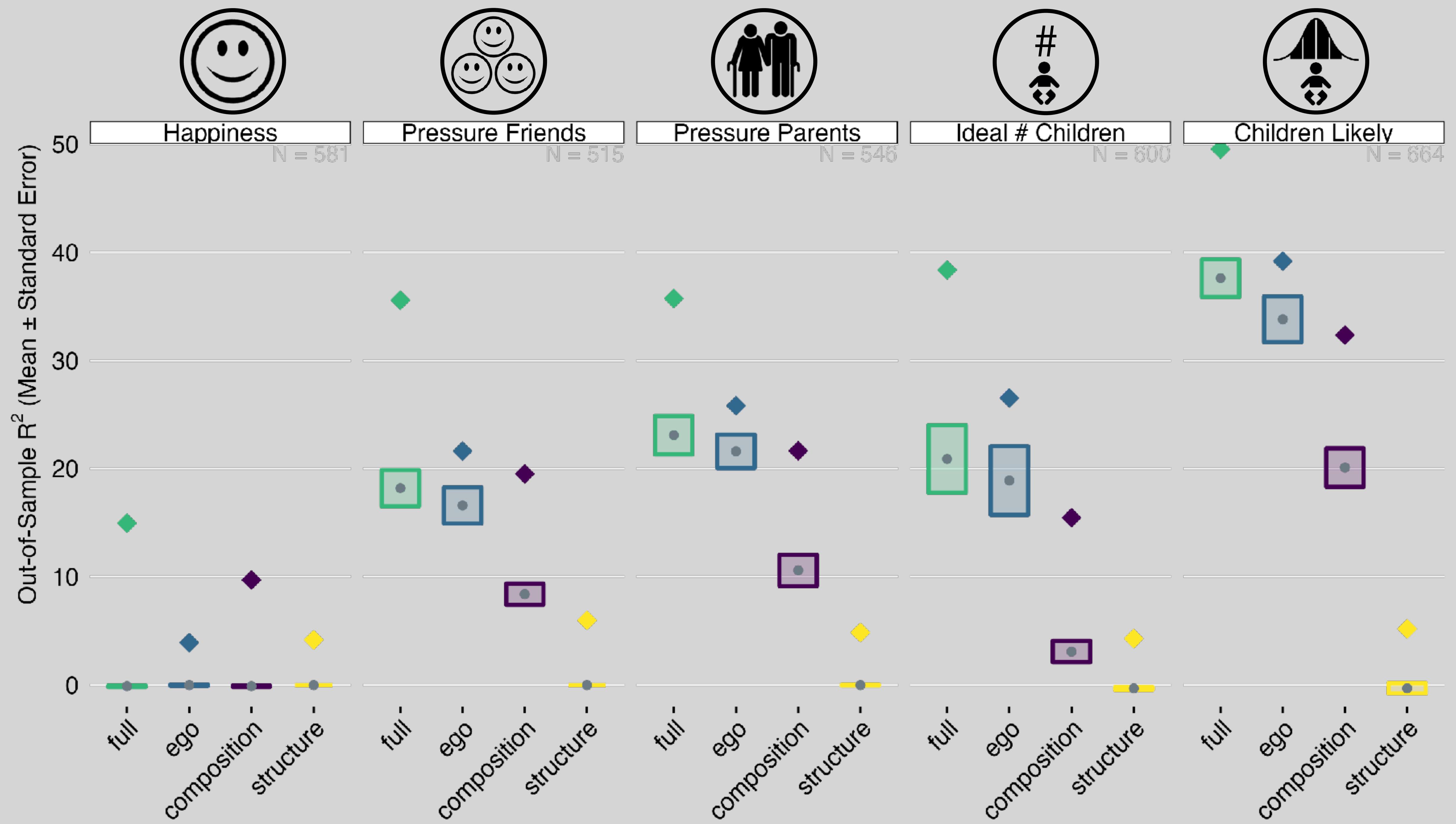
RMSE: 0.45



fold 4



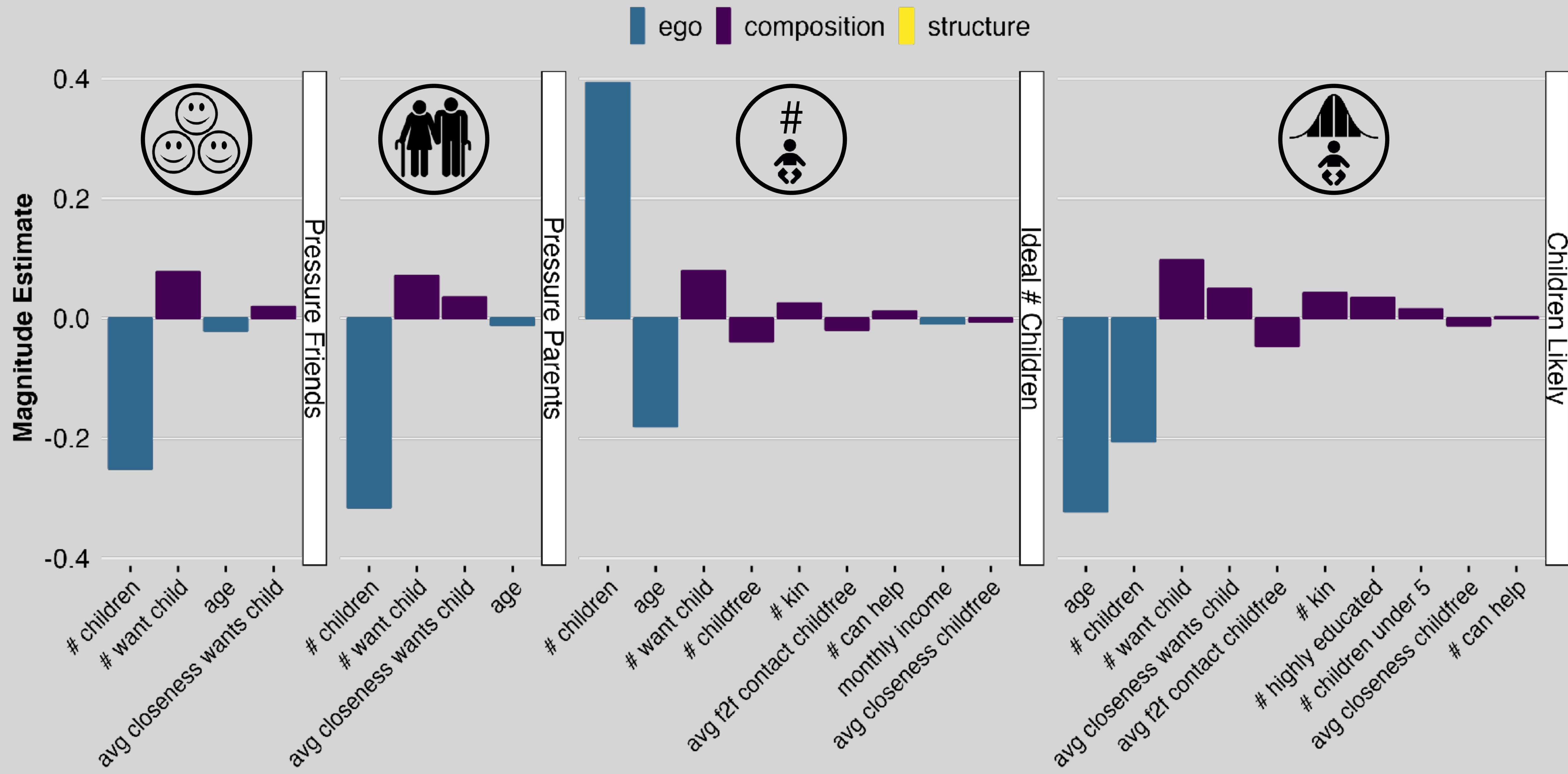
RMSE: 0.62



# Take-home messages

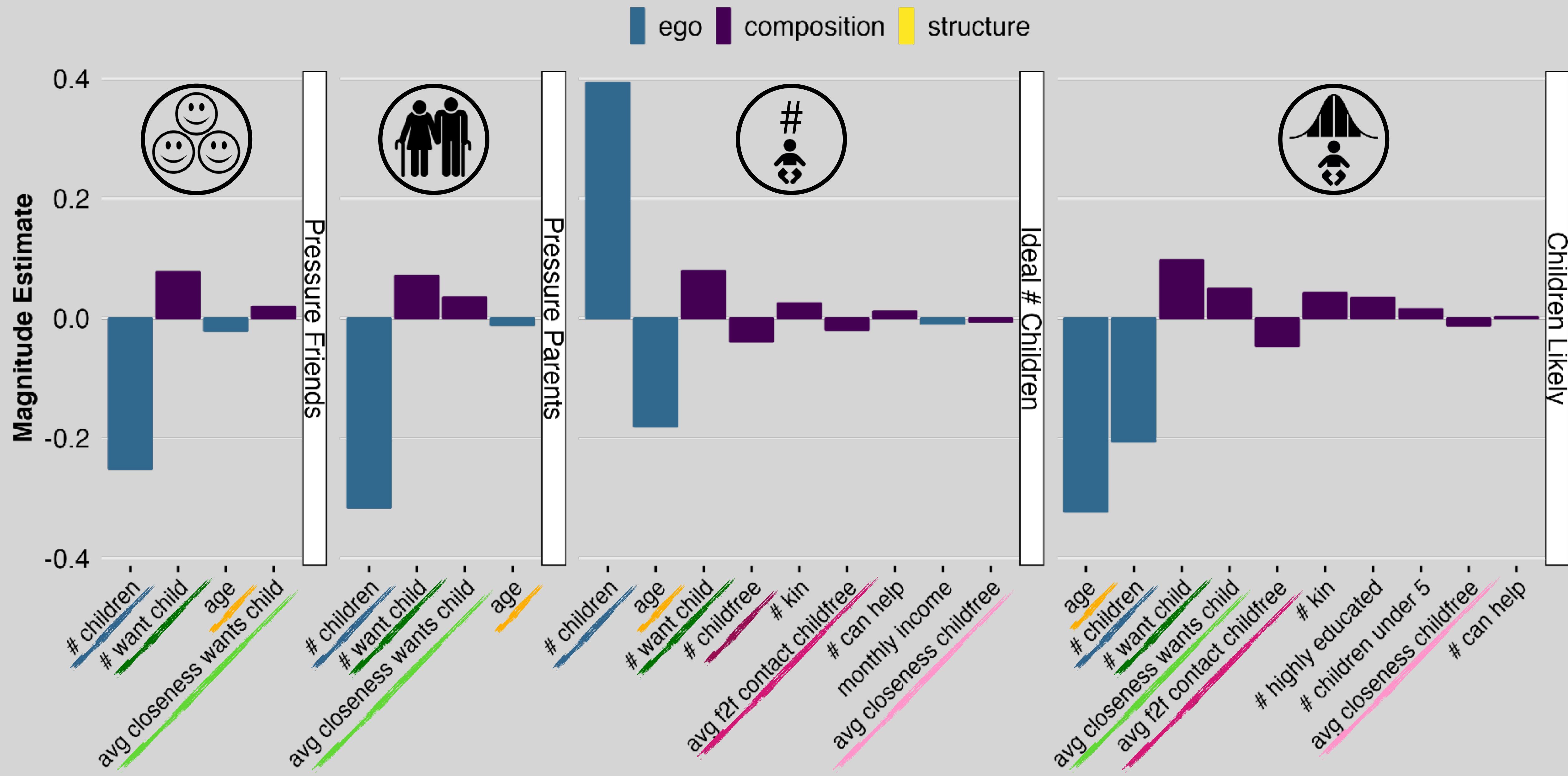
✓ predicting pretty well!

(✗) massive overfitting (~15 %-points)



# Take-home messages

- ✓ predicting pretty well!
- (✗) massive overfitting (~15 %-points)
- ✓ personal variables important, composition so-so, structure not



# Important Variables



- age
- # children
- # alters who **do** want children
- # alters who **do not** want children
- strength of relationship to these people

# Take-home messages

- ✓ predicting pretty well!
- (✗) massive overfitting (~15 %-points)
- ✓ personal variables important, composition so-so, structure not
- ✓ people who want children and who do not important

# Take-home messages

 predicting pretty well!

difficult to assess how well

 massive overfitting (~15 %-points)

potentially misleading conclusions

 personal variables important, composition so-so, structure not

networks may not be unimportant, few ego variables

 people who want children and who do not important

understudied

## *social learning*

- ✓ # people with childwish,  
ties to them
- ✓ # childfree people,  
ties to them

## *social contagion*

- ✓ # children under 5

- ✓ # kin
- ✓ # people that can help

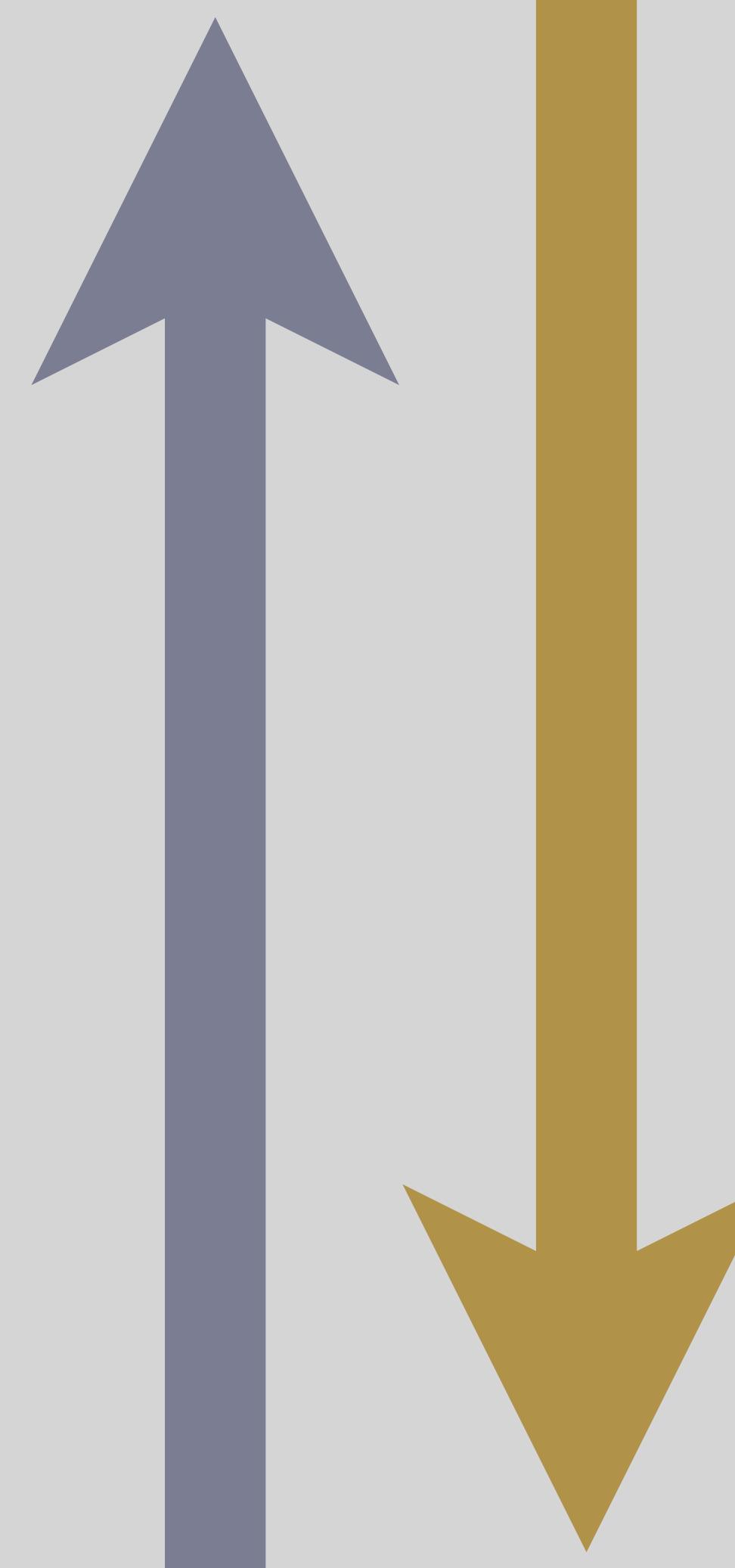
## *social support*

- ✓ people felt pressure
- ✓ # people with childwish

## *social pressure*

# INTERPRETABILITY

LASSO regression

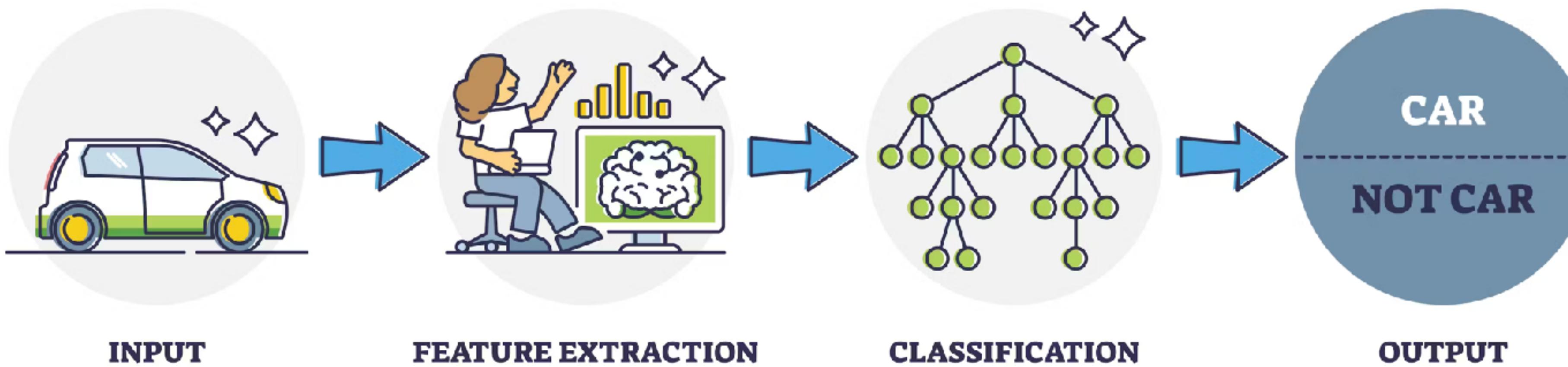


XGBoost  
Support Vector Machines

Graph Neural Networks

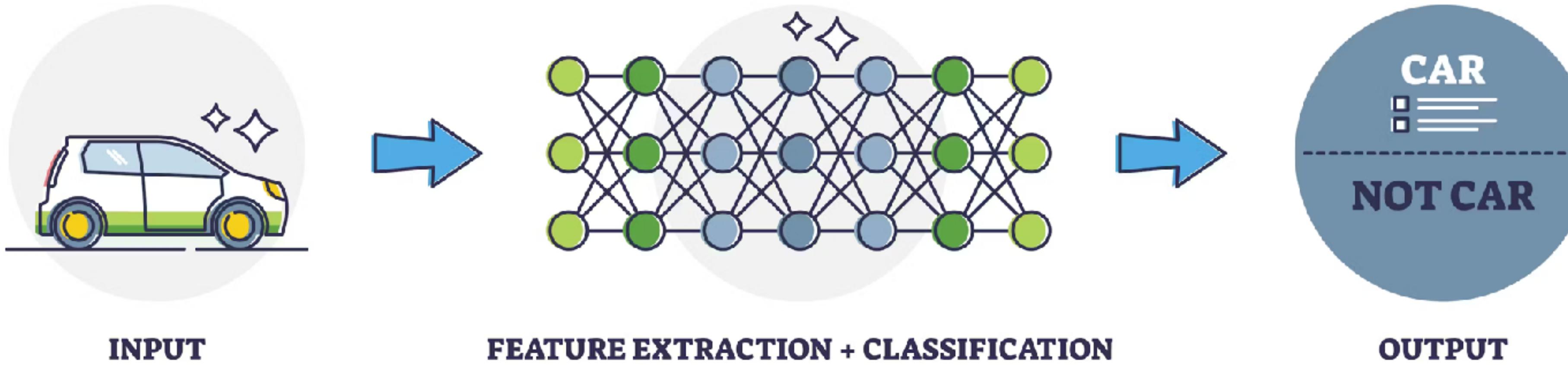
# COMPLEXITY

## MACHINE LEARNING



LASSO  
XGBoost  
SVM

## DEEP LEARNING



GNN

# Graph Neural Networks

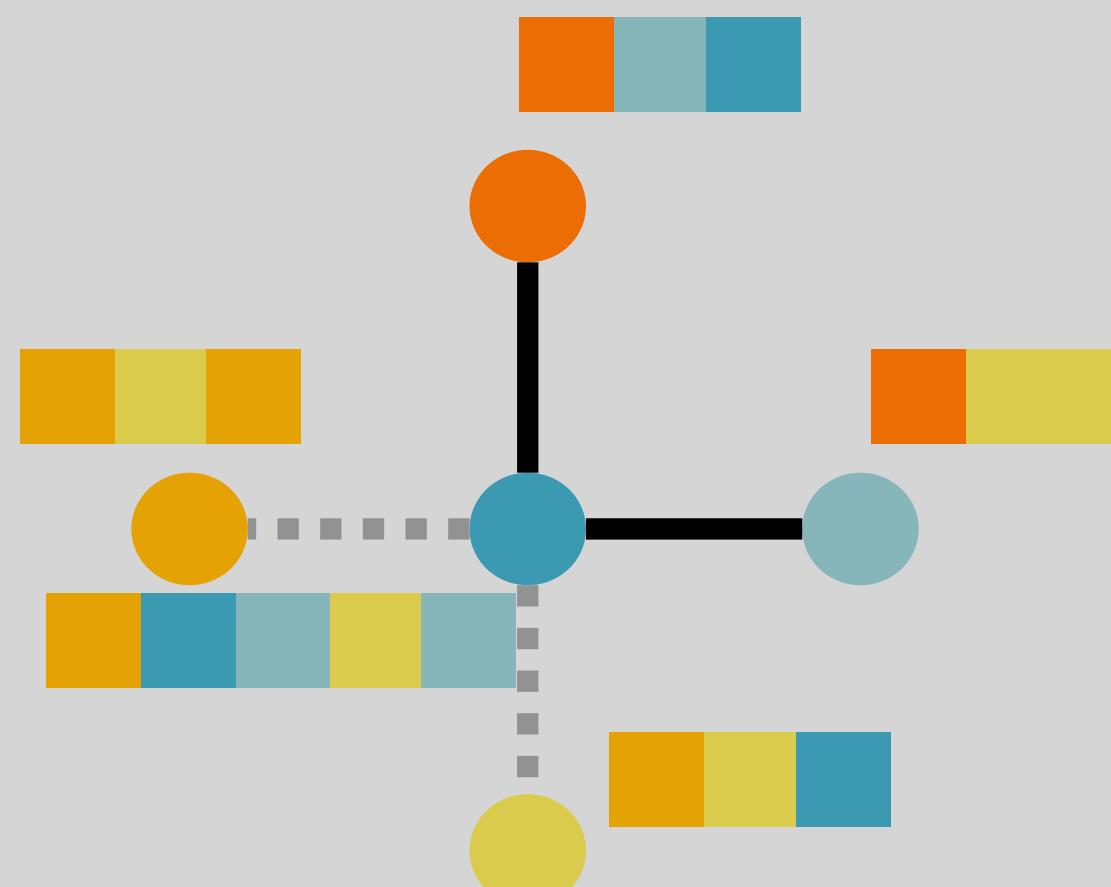


Pau  
Vila Soler

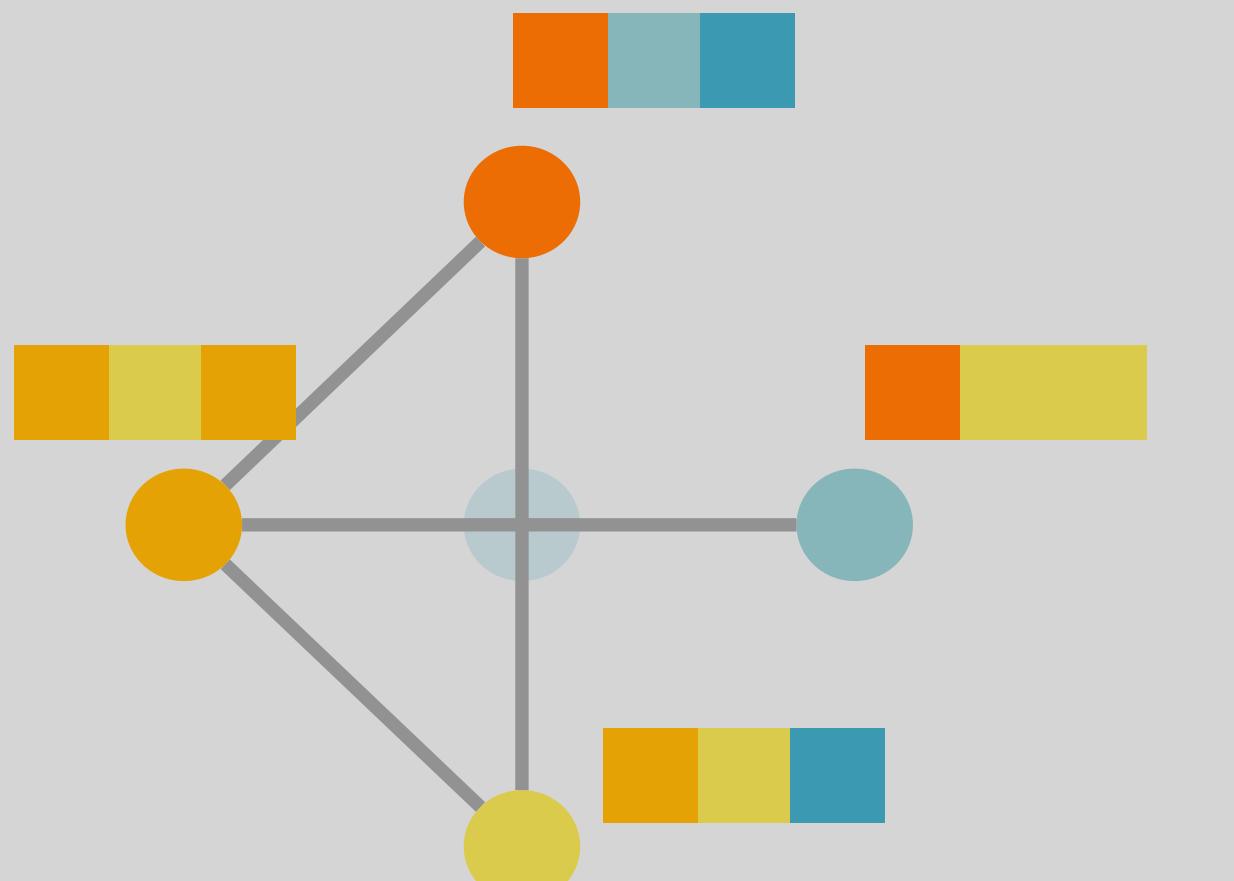


Javier  
Garcia-Bernardo

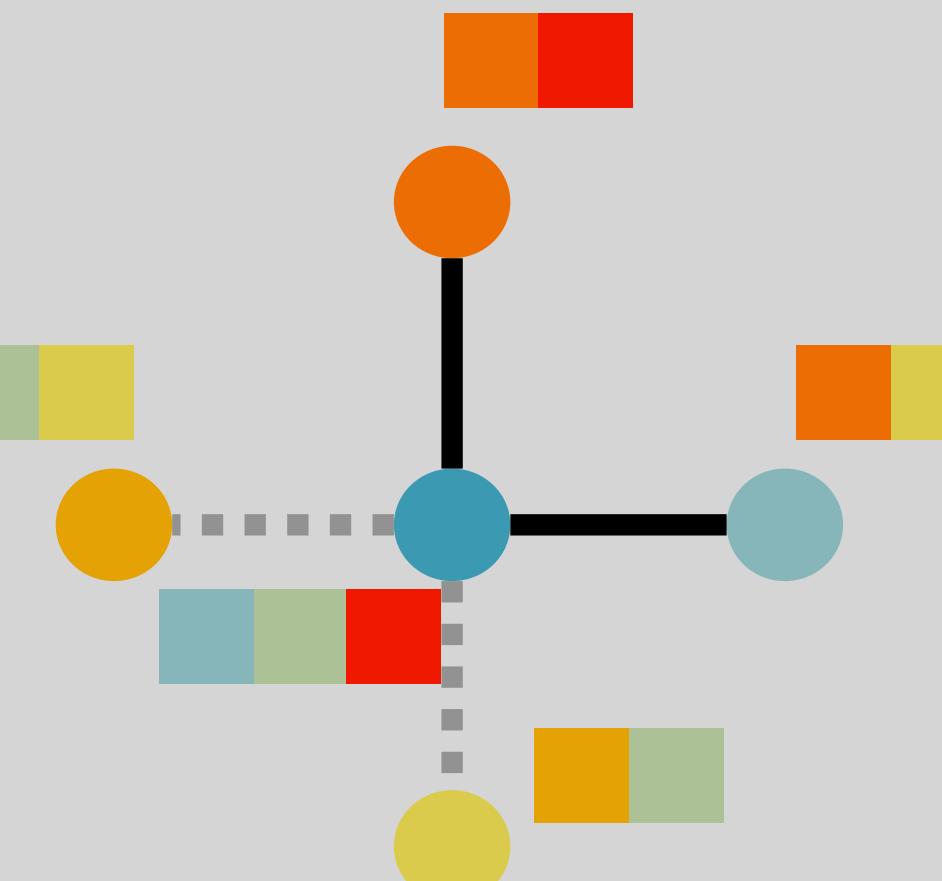
Combining ego-alter  
information



Combining alter-alter  
information



Combining alter-ego  
information

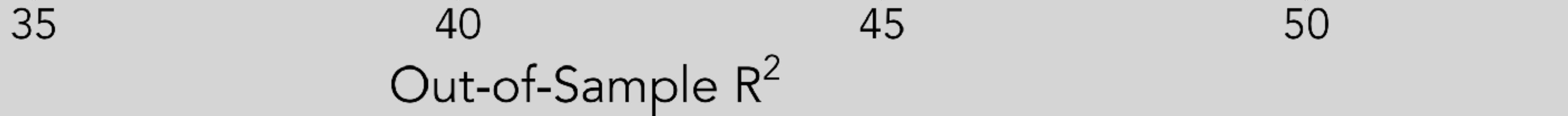


prediction on  
graph level

children likely?



SVM  
RIDGE  
ELASTIC  
LASSO  
GNN  
CUBIST  
XGB



children likely?



SVM  
RIDGE  
ELASTIC  
LASSO  
GNN  
CUBIST  
XGB

30 35 40 45 50

Out-of-Sample  $R^2$



LASSO already  
pretty good

children likely?



SVM  
RIDGE  
ELASTIC  
LASSO  
GNN  
CUBIST  
XGB

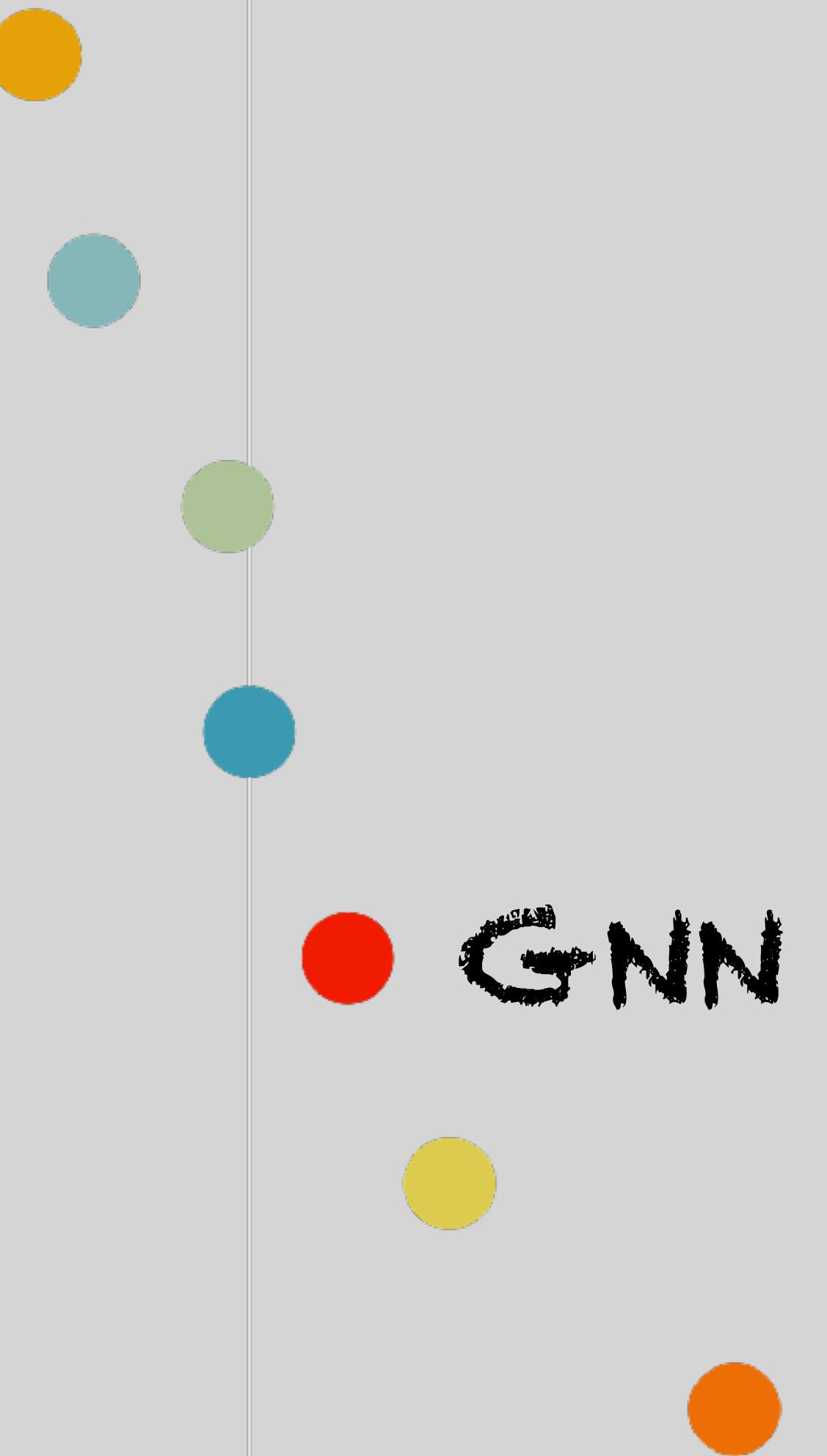
35

40

45

50

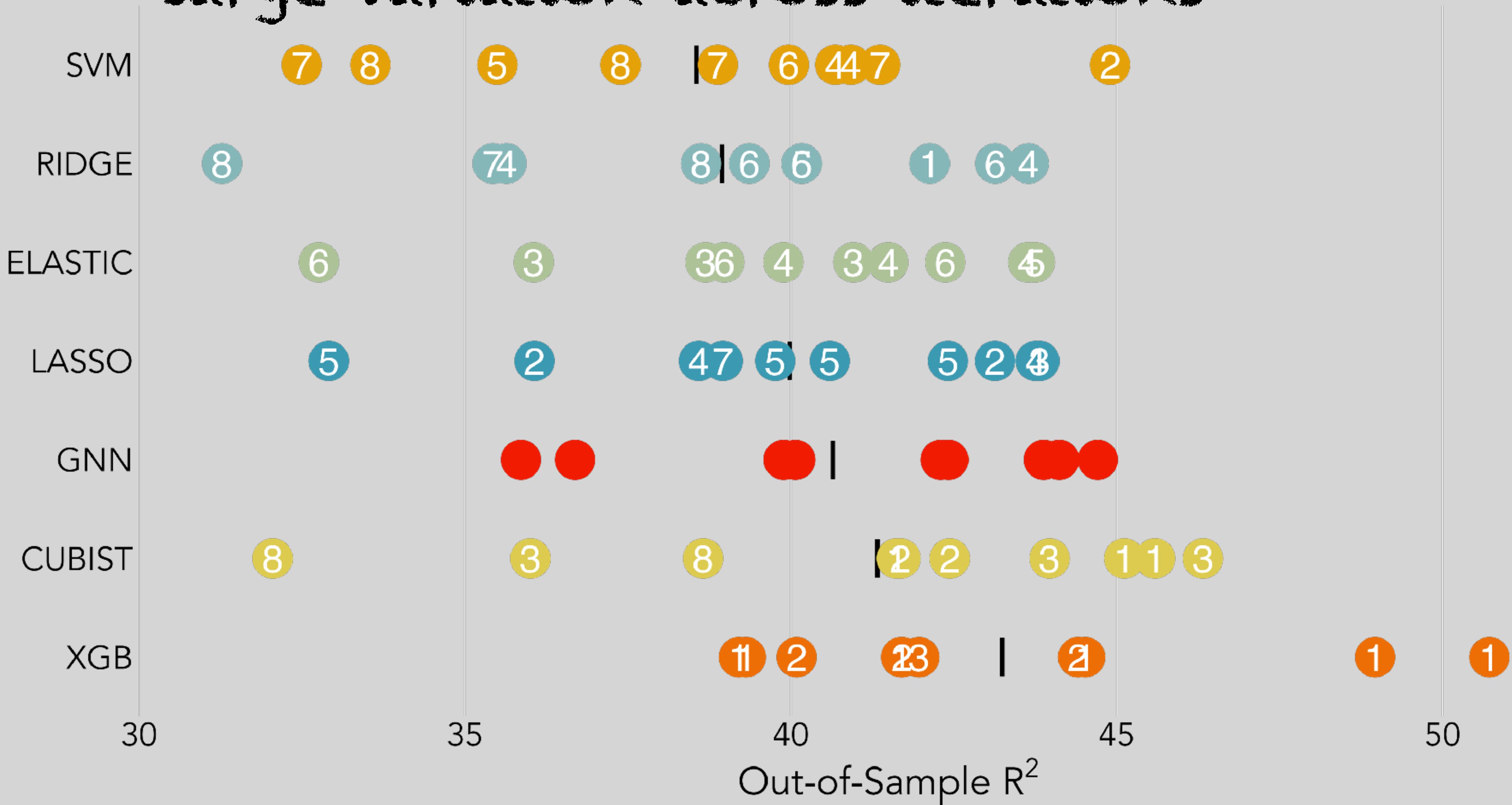
Out-of-Sample  $R^2$



● GNN pretty good

Large variation across iterations

# children likely? (



# Take-home messages

-  some improvement over LASSO regression
  - some evidence for non-linearities
-  improvement over LASSO regression not impressive
  - is lack of interpretability and dozens of hours compute worth it?
-  large variation across iterations
  - sample size clearly a constraint
-  GNN performed well!
  - requires fewer decisions, capitalises on network structure

# R package FertNet

## FertNet: Process Data from the Social Networks and Fertility Survey

Processes data from The Social Networks and Fertility Survey, downloaded from <<https://dataarchive.lissdata.nl>>, including correcting respondent errors and transforming network data into network objects to facilitate analyses and visualisation.

Version: 0.1.1  
Imports: [haven](#) (≥ 2.5.1)  
Suggests: [testthat](#) (≥ 3.0.0), [tidygraph](#) (≥ 1.2.2)  
Published: 2023-03-16  
Author: Stulp Gert  [aut, cre]  
Maintainer: Stulp Gert <g.stulp at rug.nl>  
License: [CC BY 4.0](#)  
NeedsCompilation: no  
Materials: [README](#) [NEWS](#)  
CRAN checks: [FertNet results](#)

### Documentation:

Reference manual: [FertNet.pdf](#)

### Downloads:

Package source: [FertNet 0.1.1.tar.gz](#)

Windows binaries: r-devel: [FertNet 0.1.1.zip](#), r-release: [FertNet 0.1.1.zip](#), r-oldrel: [FertNet 0.1.1.zip](#)

macOS binaries: r-release (arm64): [FertNet 0.1.1.tgz](#), r-oldrel (arm64): [FertNet 0.1.1.tgz](#), r-release (x86\_64): [FertNet 0.1.1.tgz](#), r-oldrel (x86\_64): [FertNet 0.1.1.tgz](#)

### Linking:

Please use the canonical form <https://CRAN.R-project.org/package=FertNet> to link to this page.



DEMOGRAPHIC RESEARCH  
*A peer-reviewed, open-access journal of population sciences*

## DEMOGRAPHIC RESEARCH

VOLUME 49, ARTICLE 19, PAGES 493–512  
PUBLISHED 8 SEPTEMBER 2023

<https://www.demographic-research.org/Volumes/Vol49/19/>  
DOI: 10.4054/DemRes.2023.49.19

### Data Description

Describing the Dutch Social Networks and Fertility Study and how to process it

Gert Stulp

© 2023 Gert Stulp.

This open-access work is published under the terms of the Creative Commons Attribution 3.0 Germany (CC BY 3.0 DE), which permits use, reproduction, and distribution in any medium, provided the original author(s) and source are given credit.  
See <https://creativecommons.org/licenses/by/3.0/de/legalcode>.

# Predicting fertility outcomes with networks

