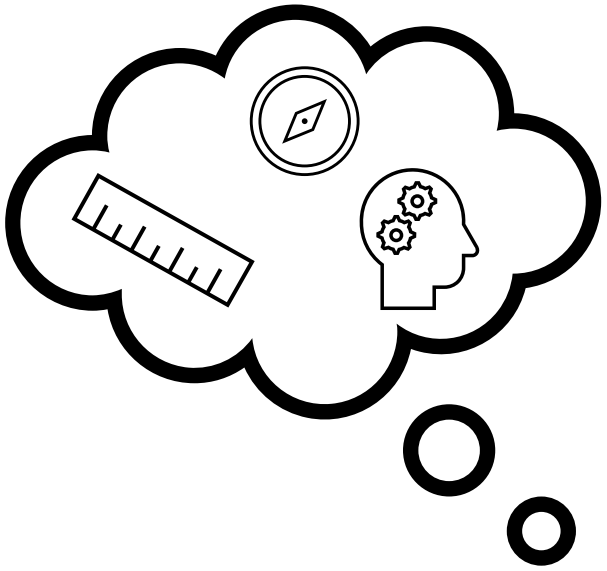
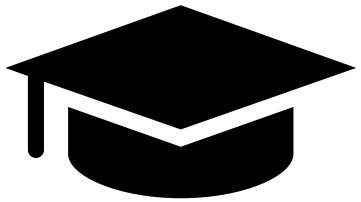
Using Particle Physics to Model Non-linear Expressions of Latent Constructs

- Philip Warncke, UNC Chapel Hill pwarncke@live.unc.edu
- Dino Capentras (ETA Zurich) dino.carpentras@gmail.com
- Adrian Lüders (University of Hohenheim) Adrian.Lueders@uni-hohenheim.de
- Mike Quale (University of Limerick) mike.quayle@ul.ie

Slides: https://philip-warncke.netlify.app/files/PhaFNA_slides.pdf

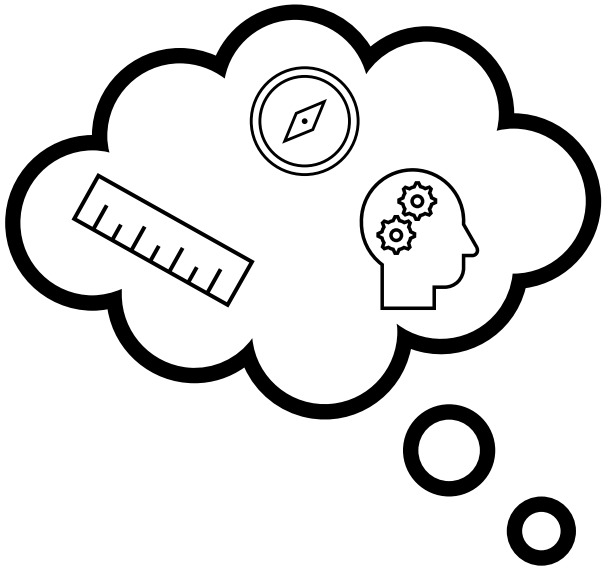
Social attitudes & latent constructs



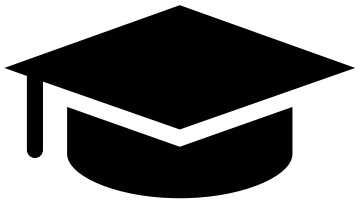


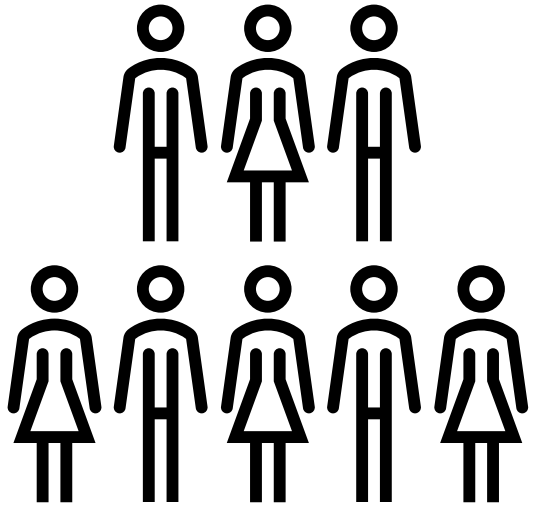
Research on social attitudes often involves **latent variables:**

- Intelligence
- Social trust
- Anxiety, depression, happiness
- Personality
- Political ideology
- …

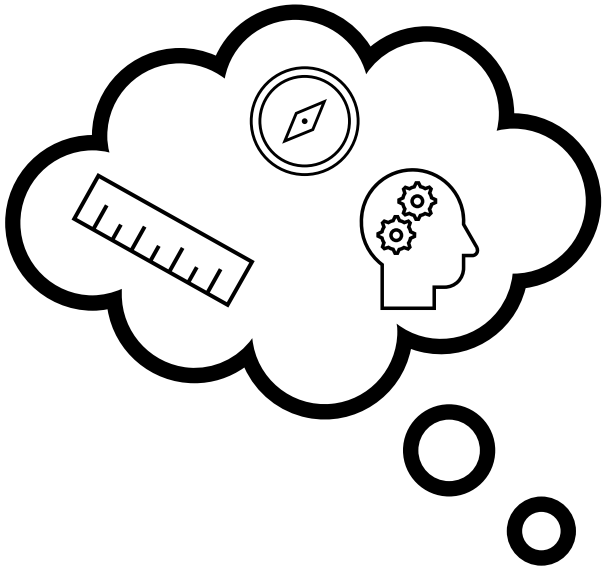


Latent variables

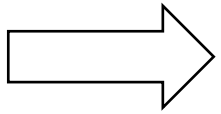


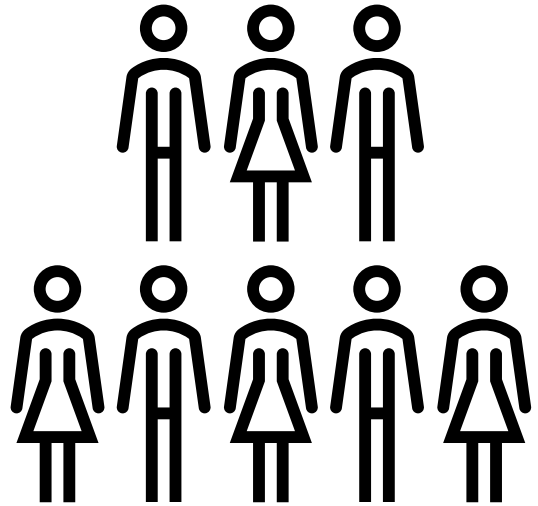


Data people provide

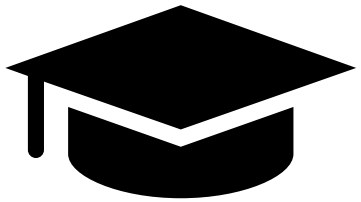


Latent data-generating process

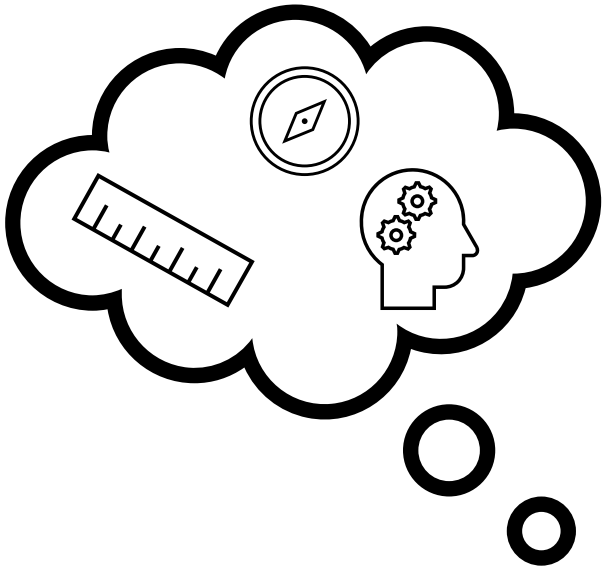




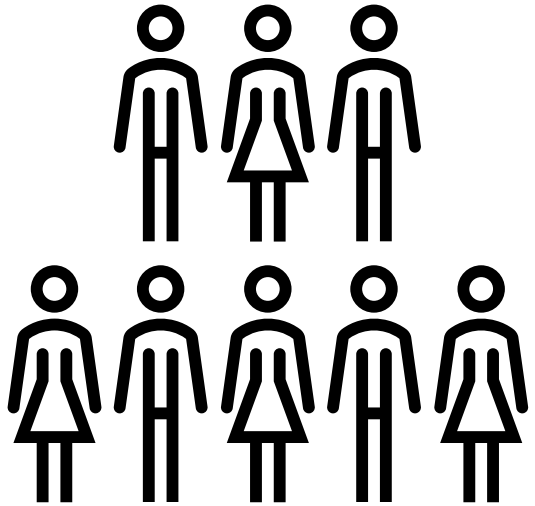
Latent variables



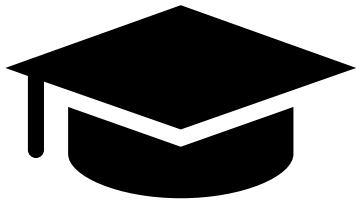
Data people provide



Latent data-generating process

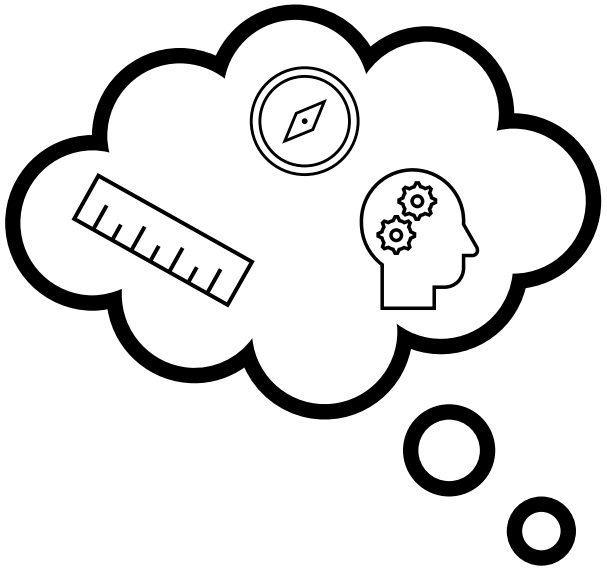


Latent variables

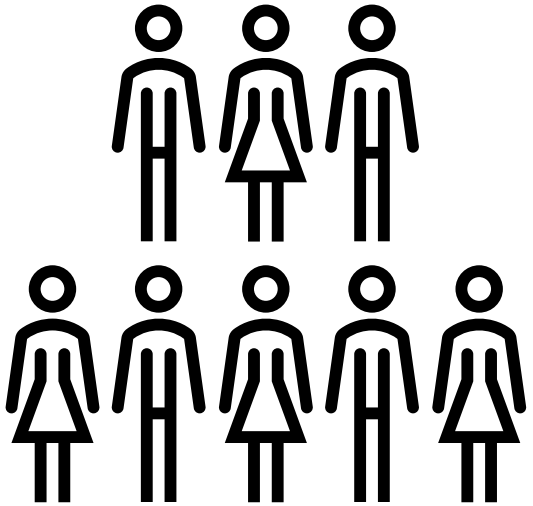


Data people provide

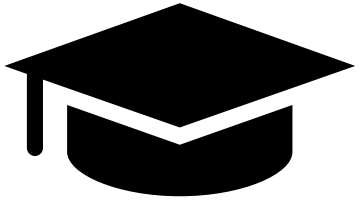
Latent variable models
aim to reconstruct latent data-generating processes from empirical data



Latent data-generating process



Latent variables



Data people provide

Latent variable models
aim to reconstruct latent data-generating processes from empirical data

Assumptions of latent data-generating processes in LVM's

- Linearity
  - Principle-Components analysis
  - Maximum-likelihood factor models
- Monotonicity
  - Item-Response Theory
  - Weighted-least squares factor models
- Linear constructs (latent dimensions form cartesian space)
  - all models

In reality …

- LDGP’s may be
  - Non-linear
  - Non-monotonic

In reality …

- LDGP’s may be
  - Non-linear
  - Non-monotonic
- Latent constructs may be non-linear

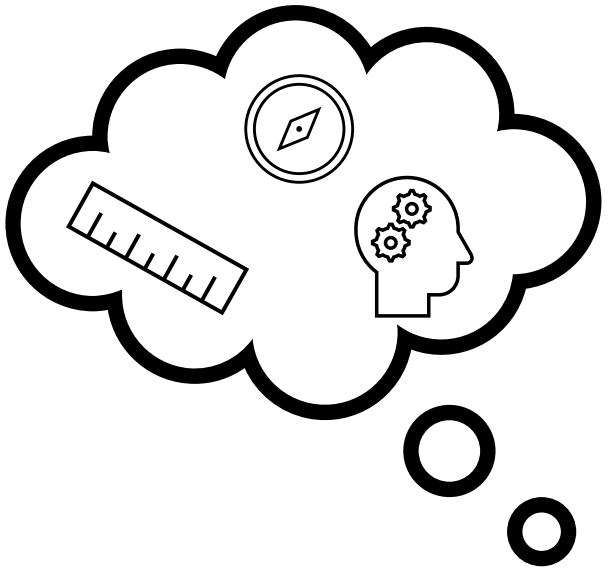
In reality …

- LDGP's may be
  - Non-linear
  - Non-monotonic

- Latent constructs may be non-linear

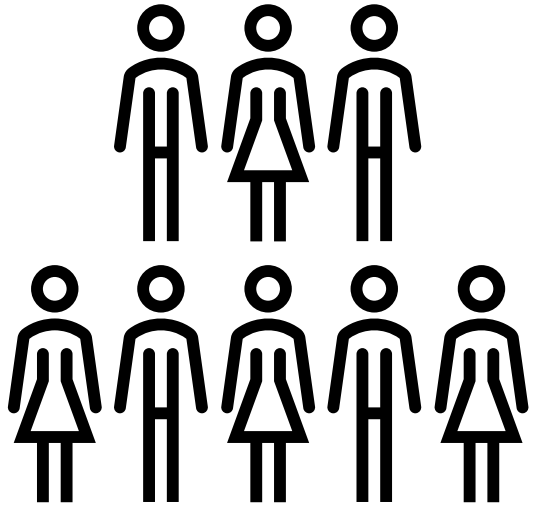
We don't know the properties of the LDGP – need to infer these from manifest data!

Latent variables

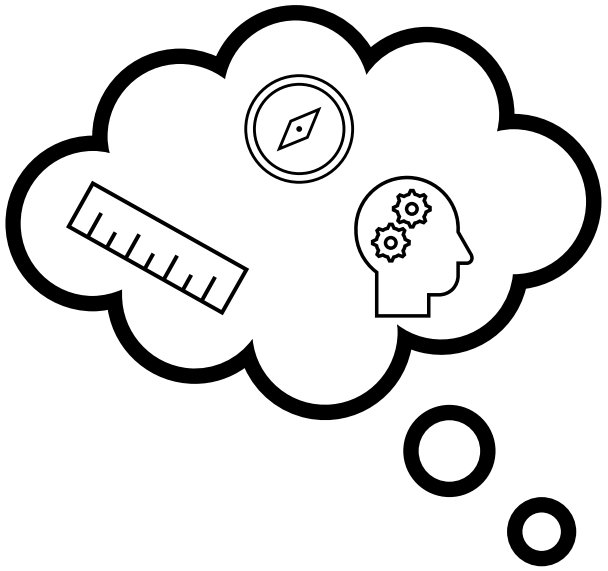


Latent data-generating processes

Manifest data



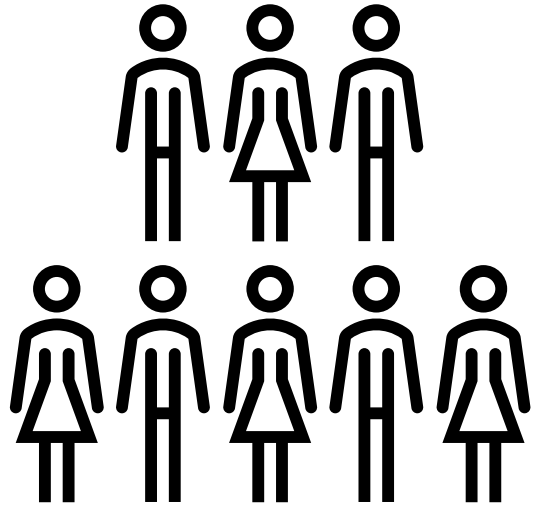
Latent variables



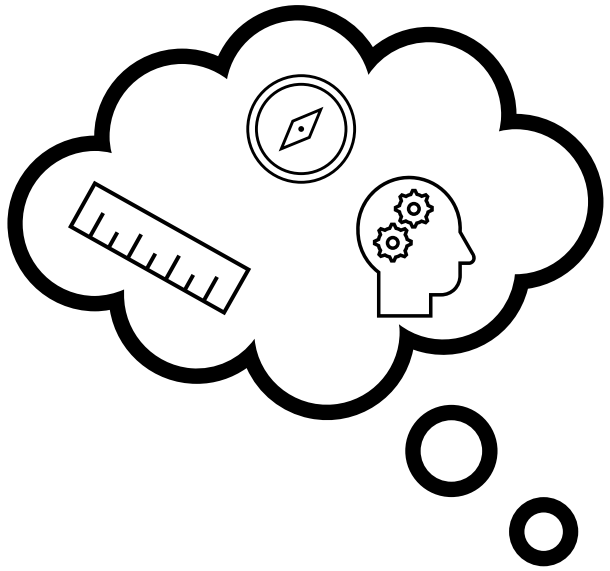
Latent data-generating processes

Assumptions of
- **Linearity**
- **Monotonicity**
- **Dimensionality**

Manifest data



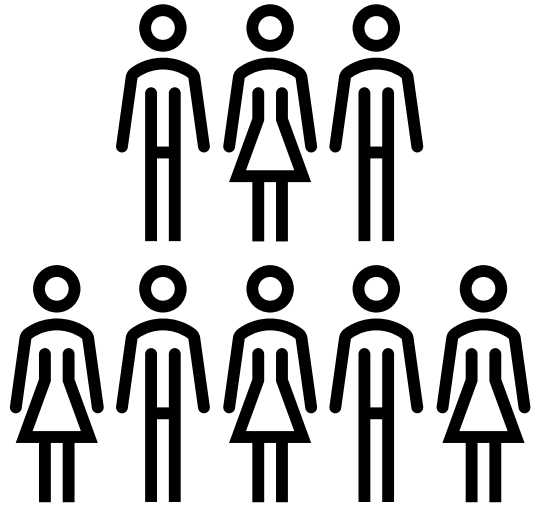
Latent variables



Latent data-generating processes

Assumptions of
- **Linearity**
- **Monotonicity**
- **Dimensionality**

Manifest data

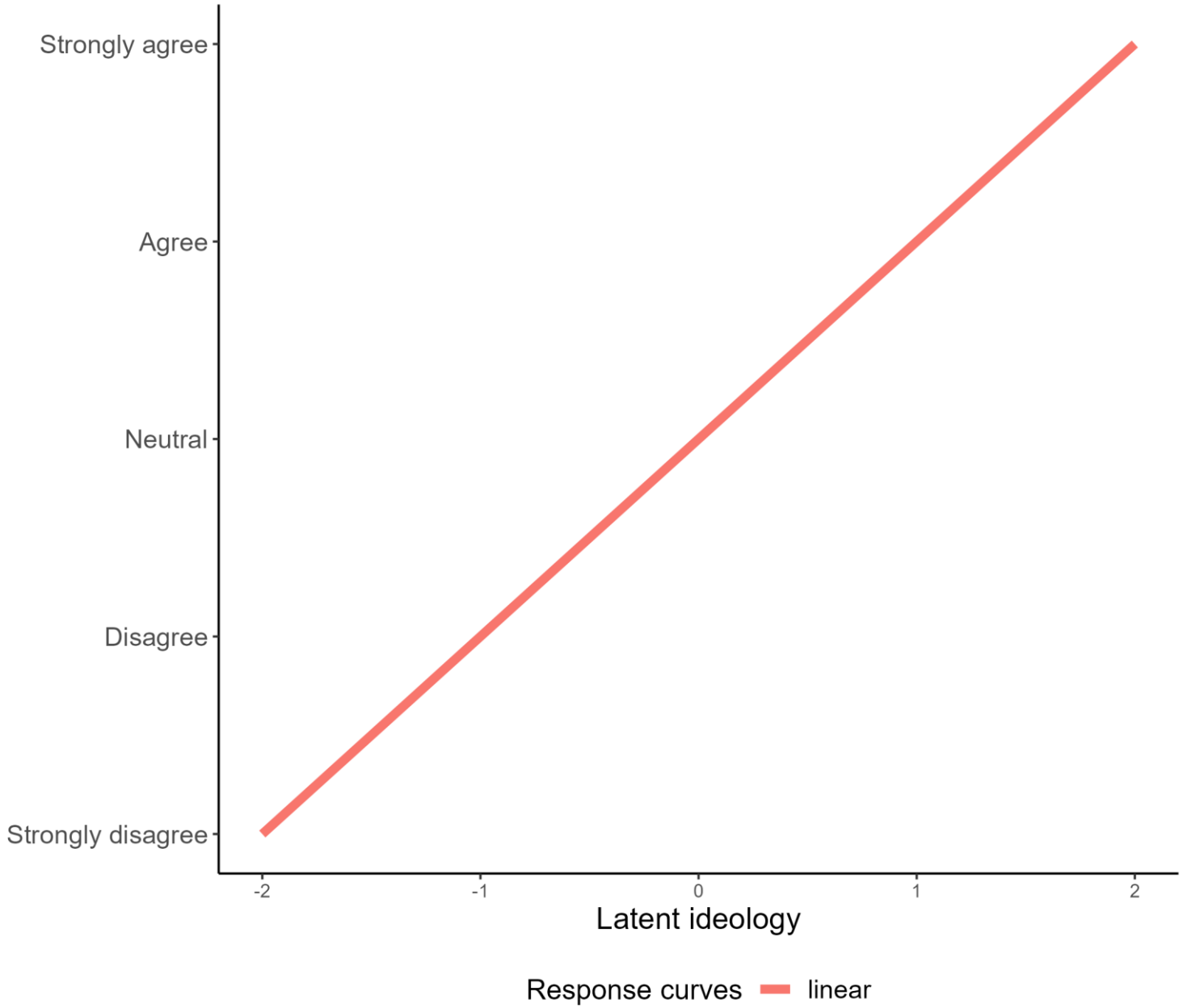


Can latent variable models reproduce latent constructs when assumptions of linearity and/or dimensionality are violated?

Latent variable: Political ideology

Survey items:

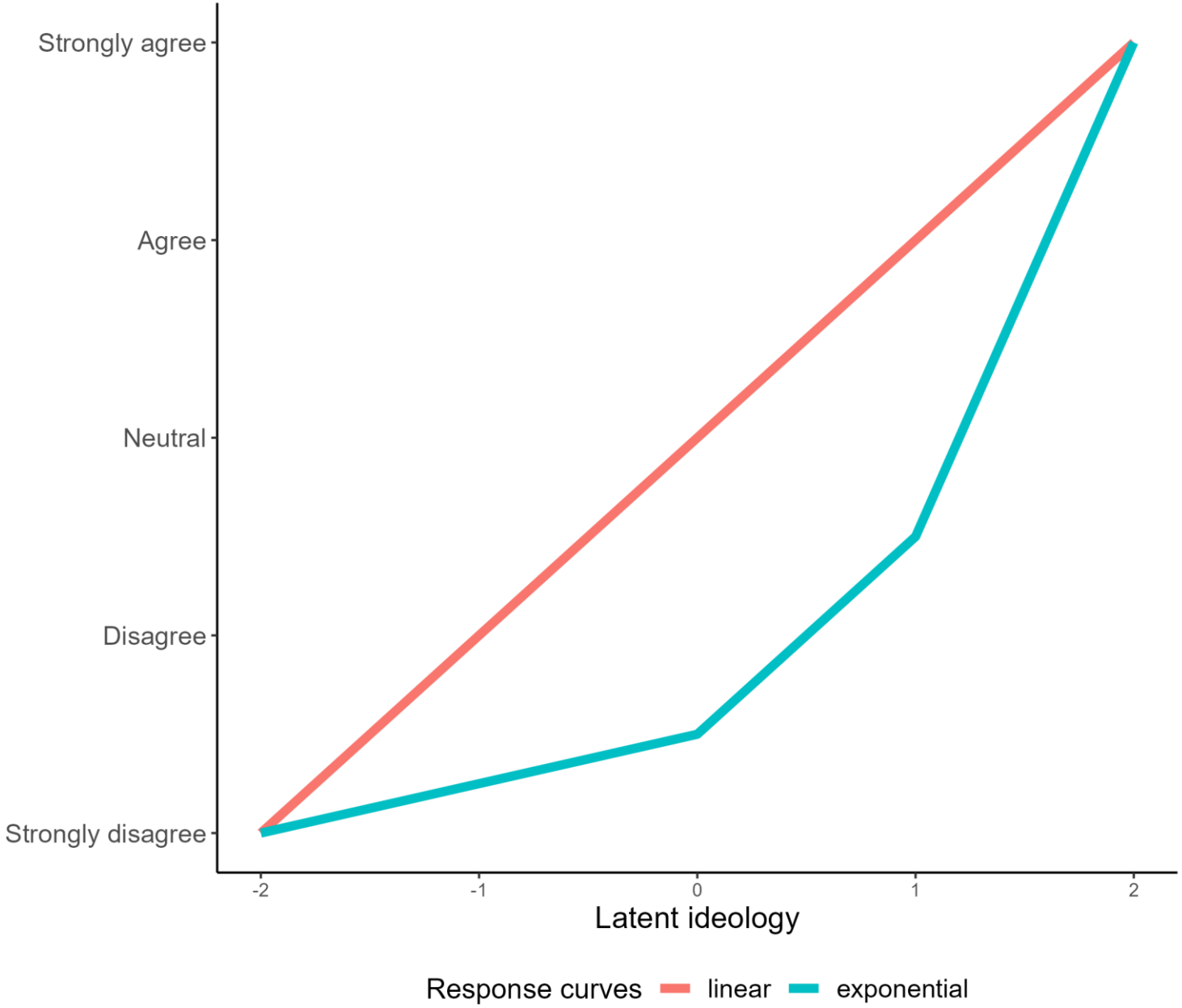
- The United States should limit technology imports from China.
- The United States should increase military support to Ukraine.
- The people I disagree with politically are not evil.



Latent variable: Political ideology

Survey items:

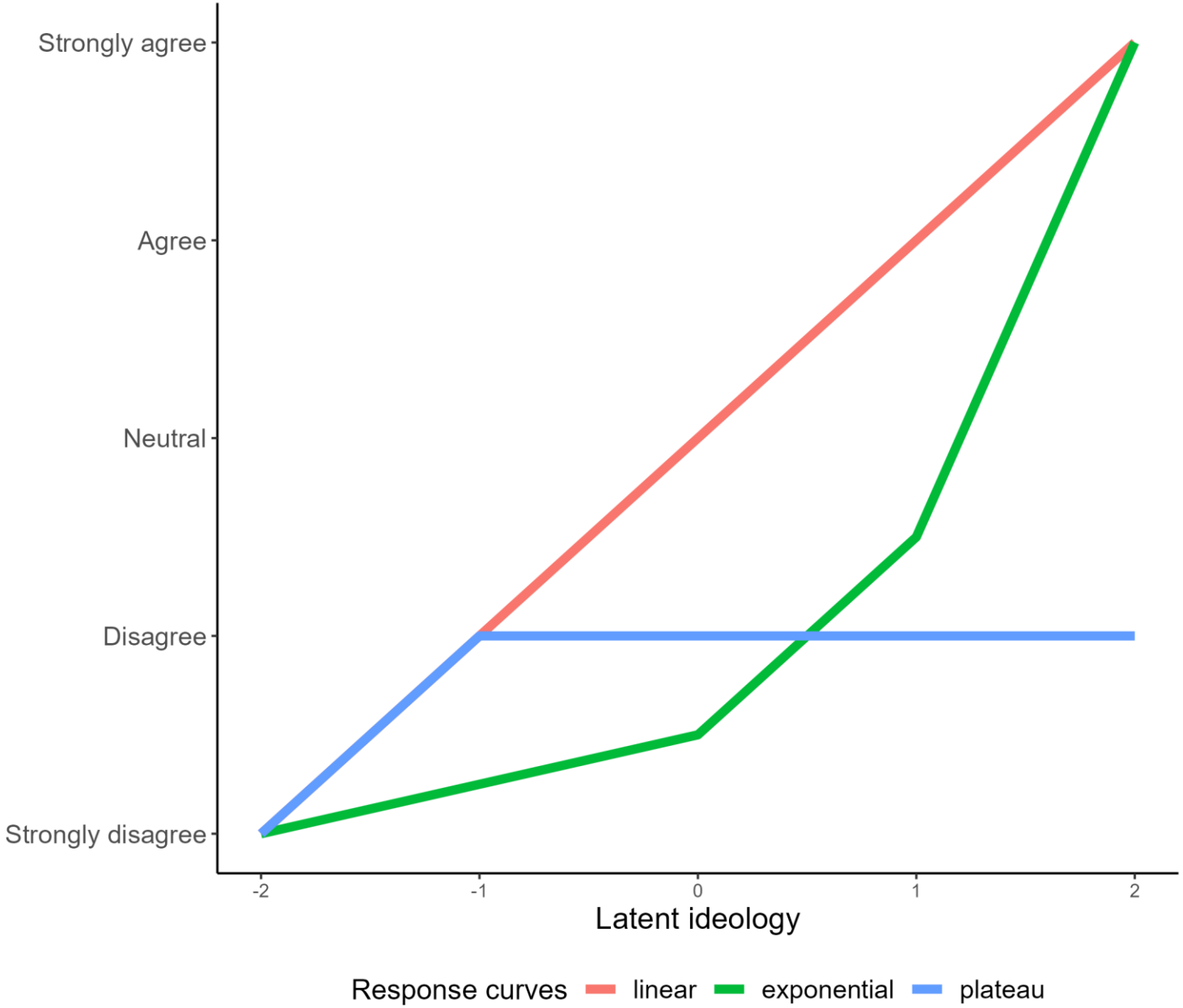
- The United States should limit technology imports from China.
- The United States should increase military support to Ukraine.
- The people I disagree with politically are not evil.



Latent variable: Political ideology

Survey items:

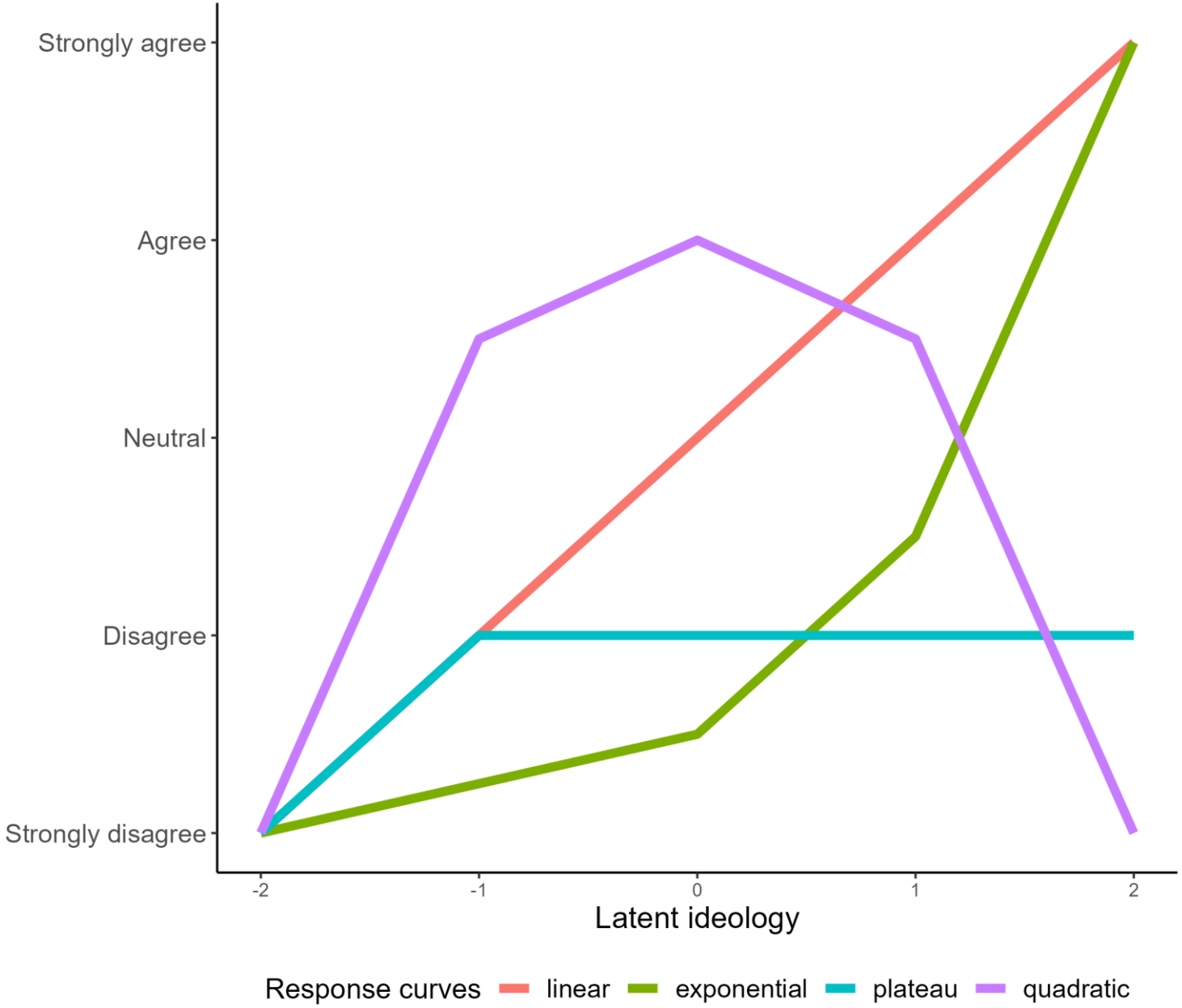
- The United States should limit technology imports from China.
- The United States should increase military support to Ukraine.
- The people I disagree with politically are not evil.



Latent variable: Political ideology

Survey items:

- The United States should limit technology imports from China.
- The United States should increase military support to Ukraine.
- The people I disagree with politically are not evil.



Simulation setup

LDGP

- N = 1,000
- K = 12 items (5-point ordered Likert scales)
- 20 % measurement error

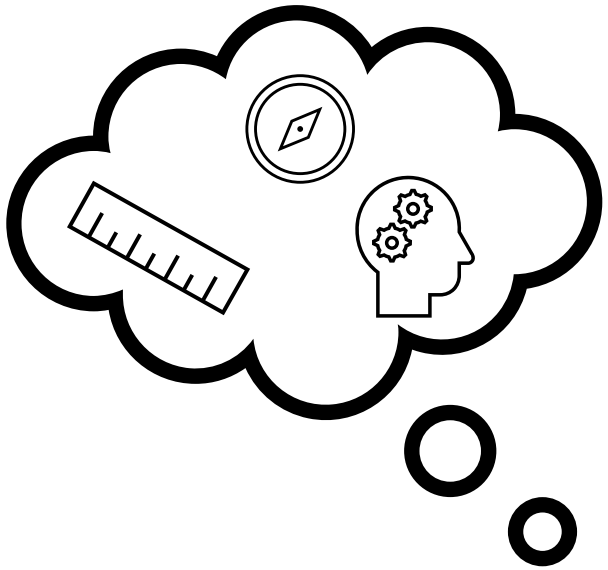
Latent variable models

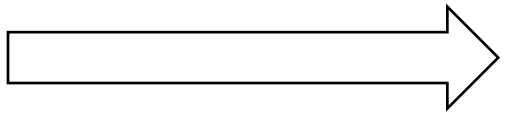
- Principle components (PCA)
- Exploratory factor analysis (EFA)
- Item response theory (IRT)
- Physics-aided factor-network analysis (PhaFNA)

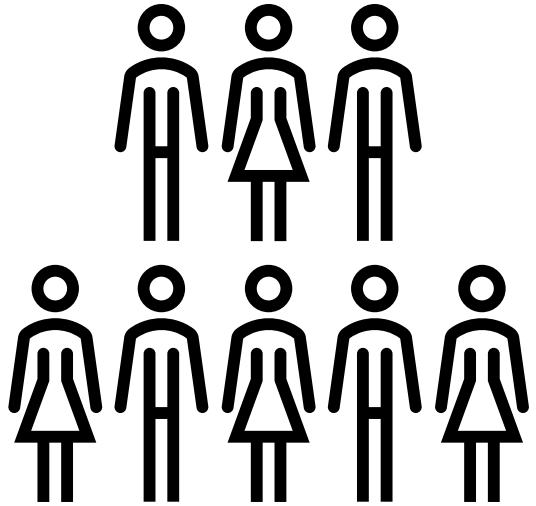
Performance comparison

- Predicative accuracy (regression R-squared of predicted scores on latent construct)
- Computational time

Scenario 1: Ideal case







Latent variable

- Unidimensional
- Continuous
- Linear

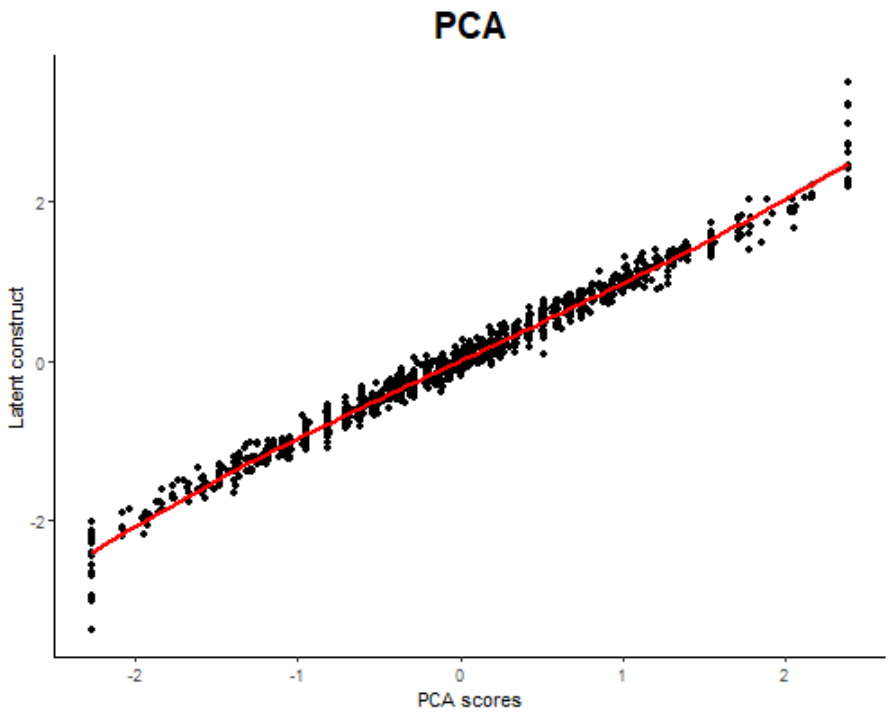
LGDP

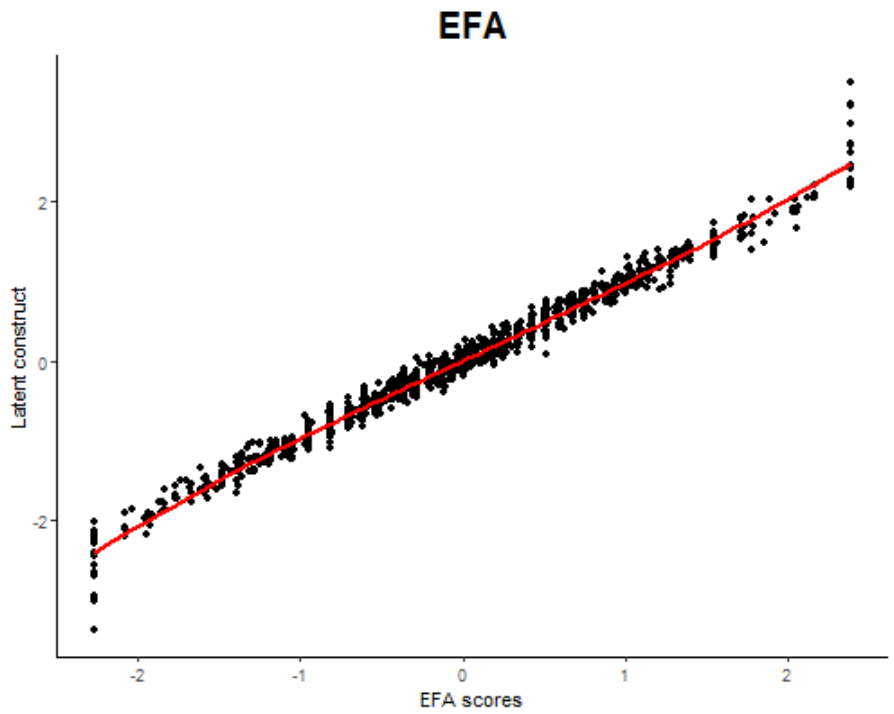
- 20% random measurement error
- Random-uniform transformation to 5-point ordered Likert-scale items

Manifest data

- 12 Likert scale items

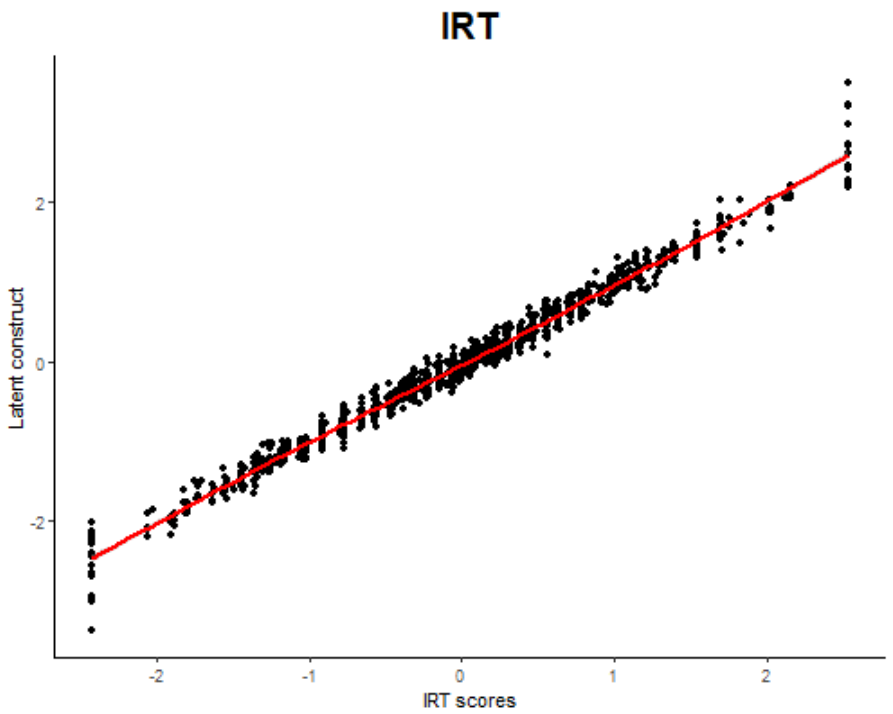
$R^2 = 0.97$
t = 0.3s

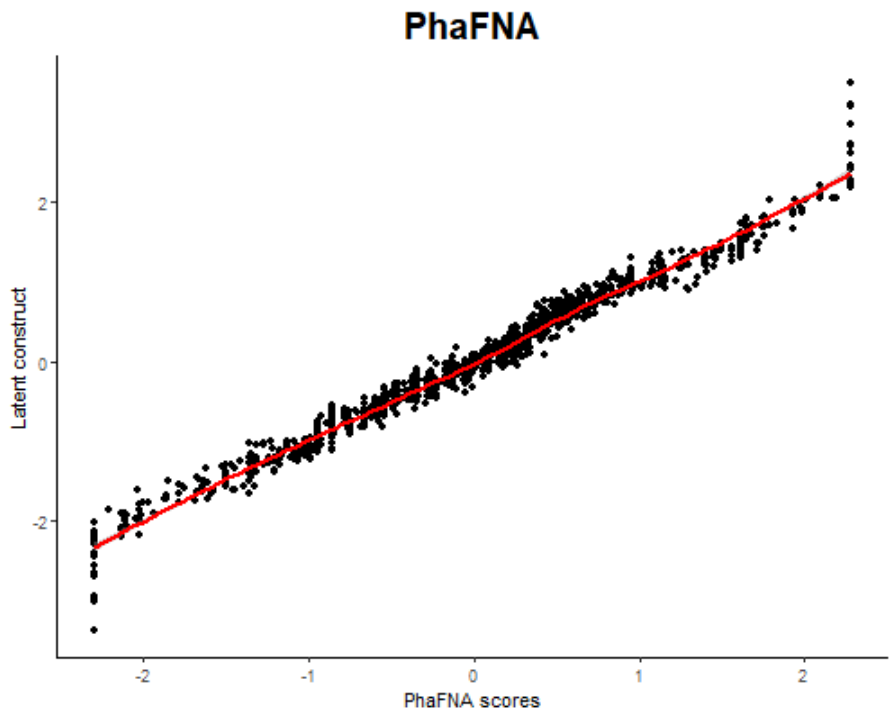




$R^2 = 0.96$
t = 1.5s

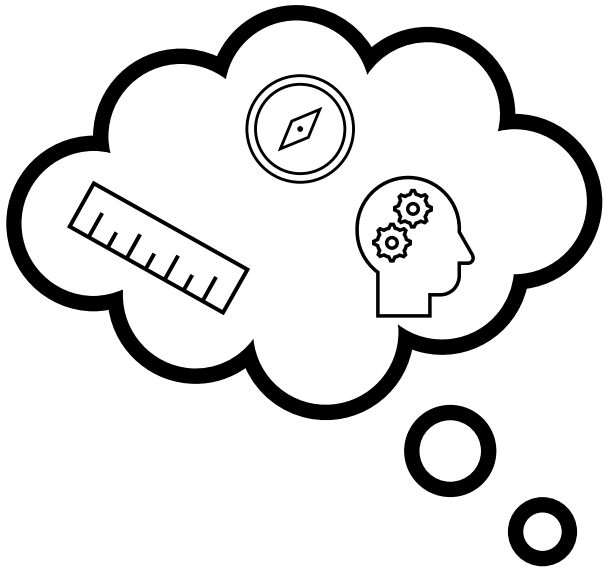
$R^2 = 0.98$
t = 2.1s

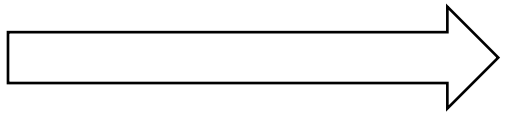


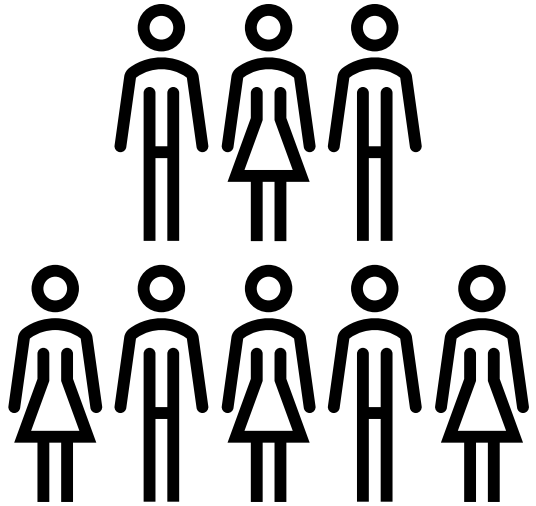


$R^2 = 0.95$
t = 1.6s

Scenario 2: LGDP violates linearity







Latent variable

- Unidimensional
- Continuous
- Linear

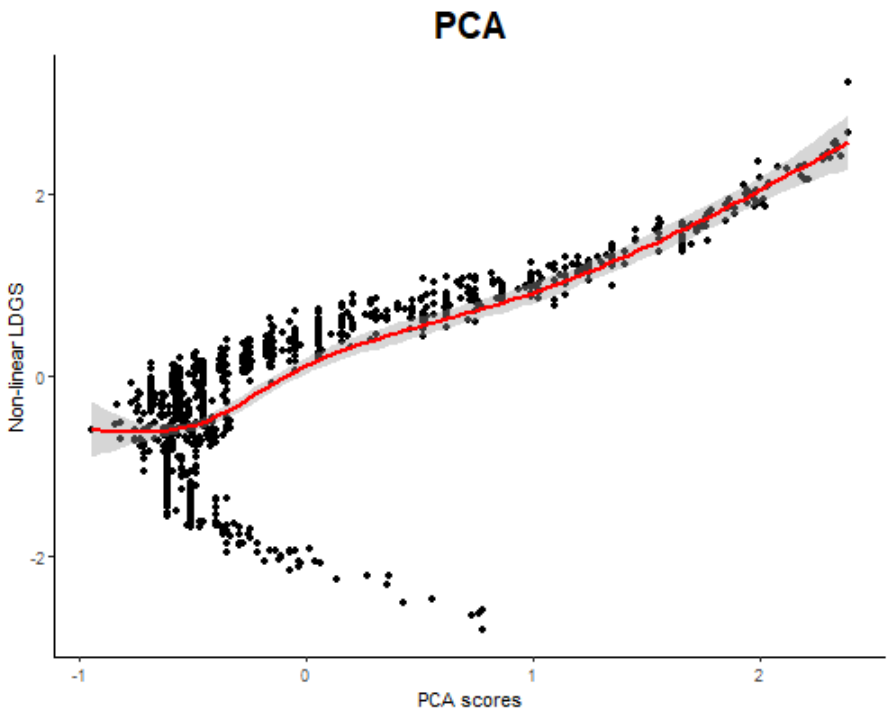
LGDP

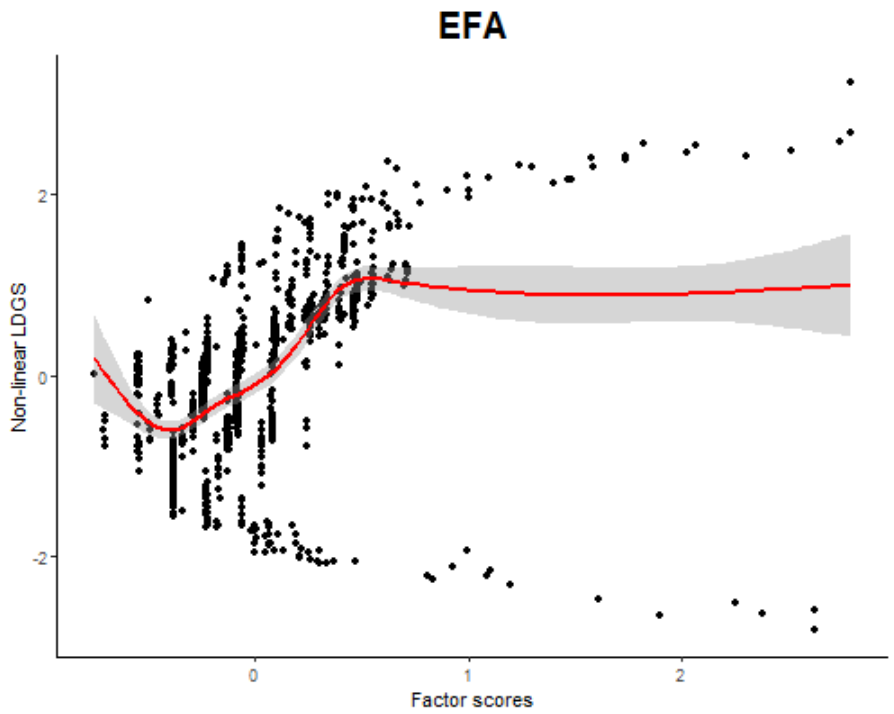
- 20% random measurement error
- **Non-linear Likert transformations**

Manifest data

- 12 Likert scale items
  - 2/12: Linear
  - 4/12: Quadratic
  - 4/12: 50% linear, 50% quadratic
  - 2/12: Exponential

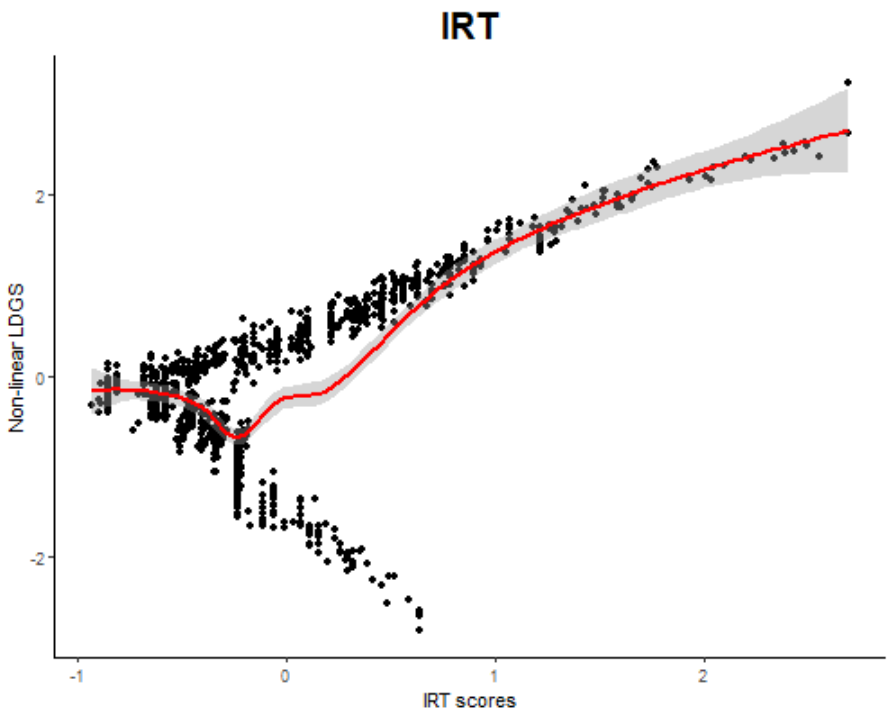
$R^2 = 0.48$
t = 0.3s

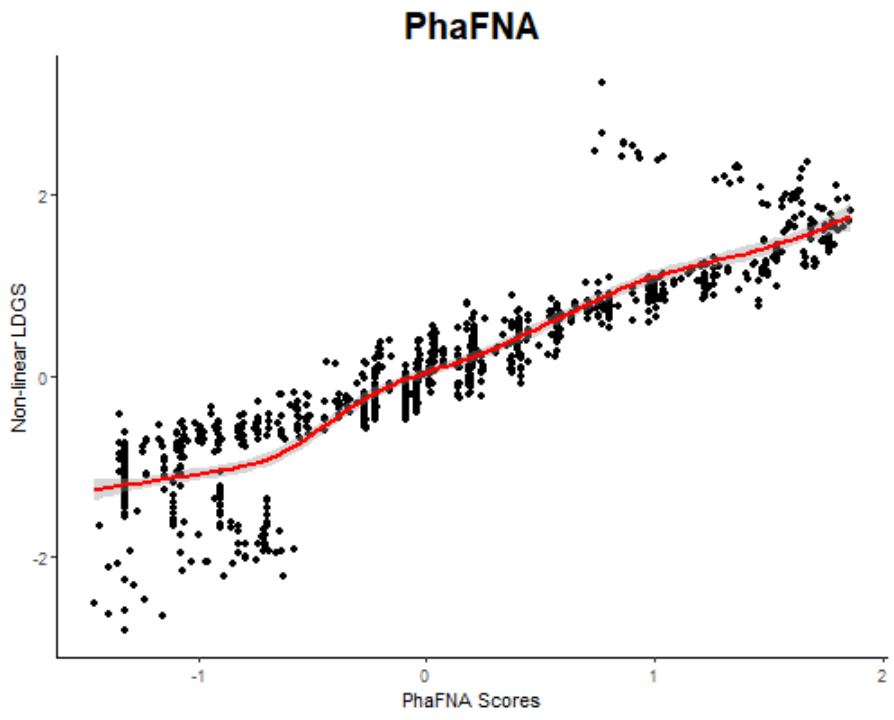




$R^2 = 0.32$
t = 3.4s

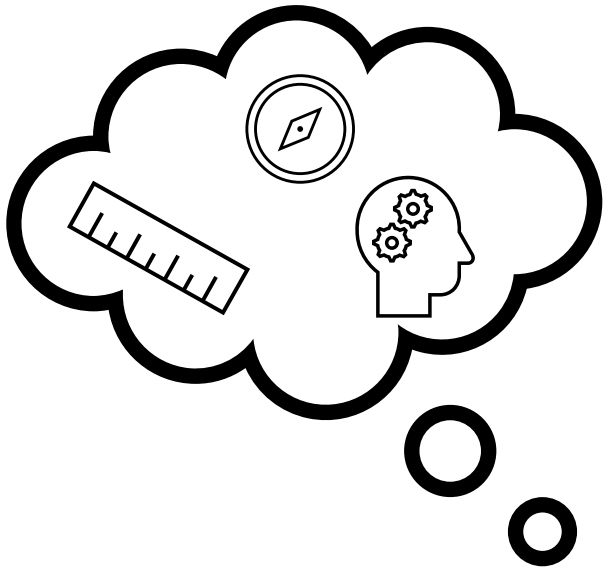
$R^2 = 0.52$
t = 8.1s

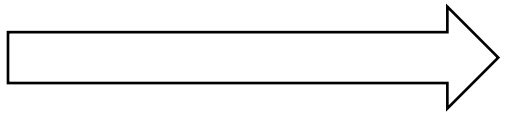


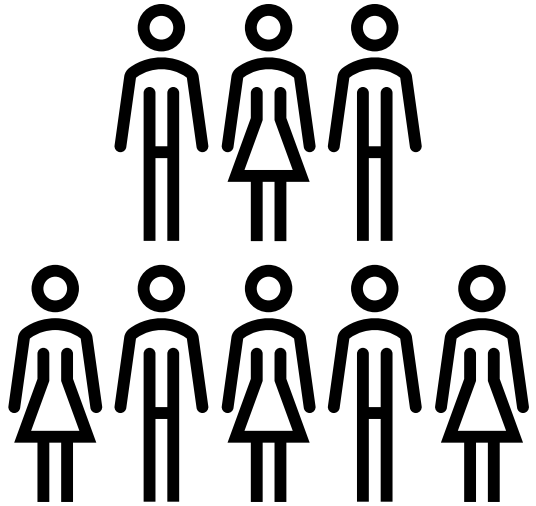


$R^2 = 0.91$
t = 2.3s

Scenario 3: Latent construct violates linearity







Latent variable

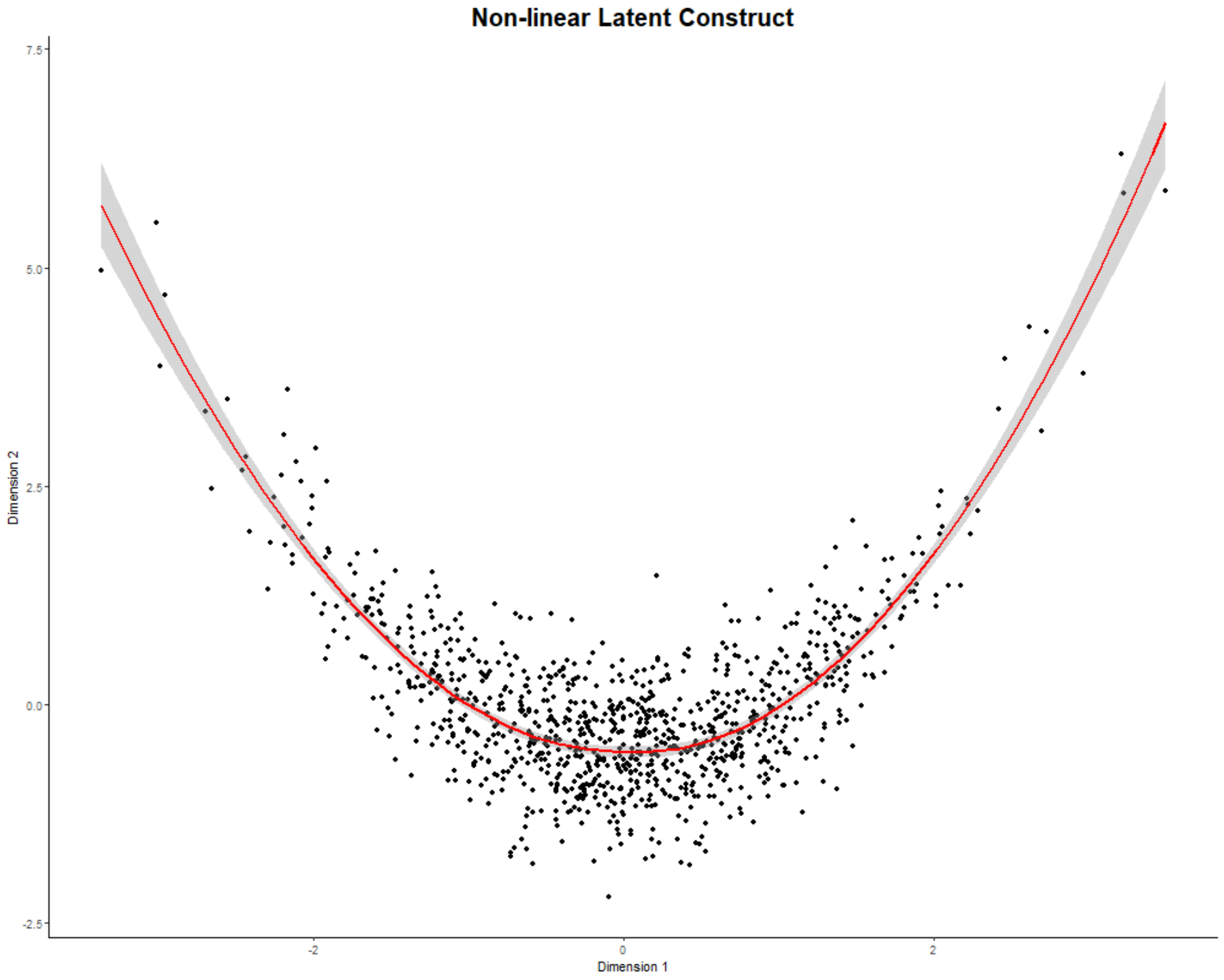
- Continuous
- **Non-linear**

LGDP

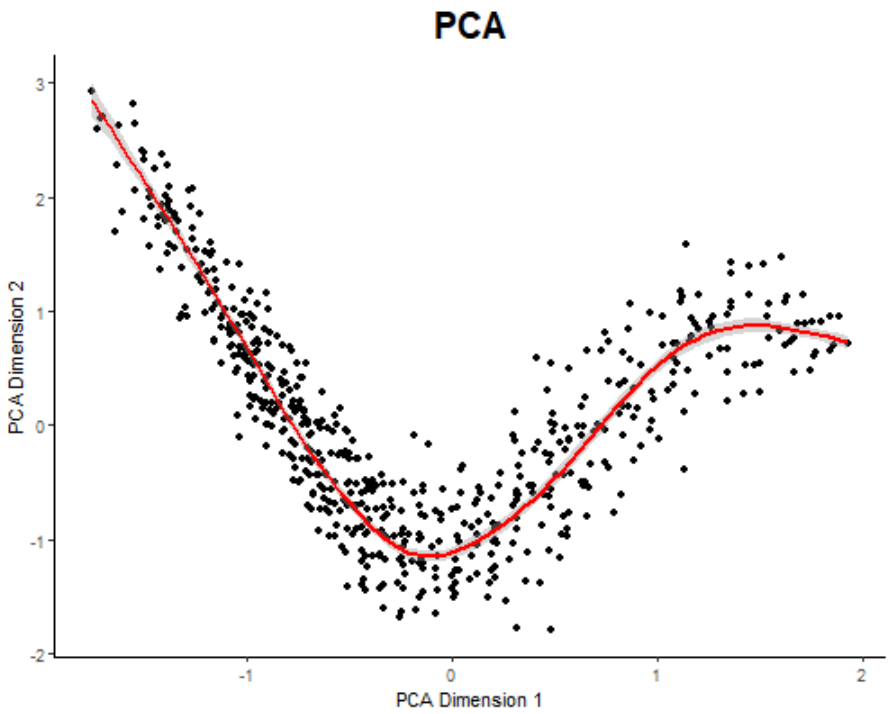
- 20% random measurement error
- Linear and non-linear components fused together

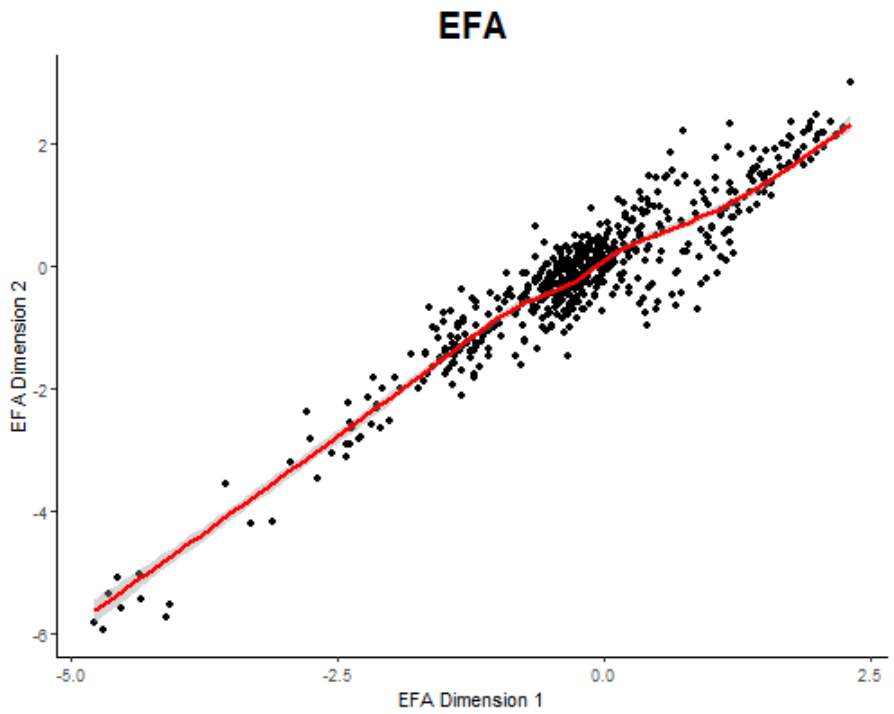
Manifest data

- 12 5-point Likert scale items
- **Items inherit 40% linear & 40% non-linear component**


Non-linear Latent Construct
7.5
5.0
2.5
0.0
-2.5
Dimension 2
-2
0
2
Dimension 1

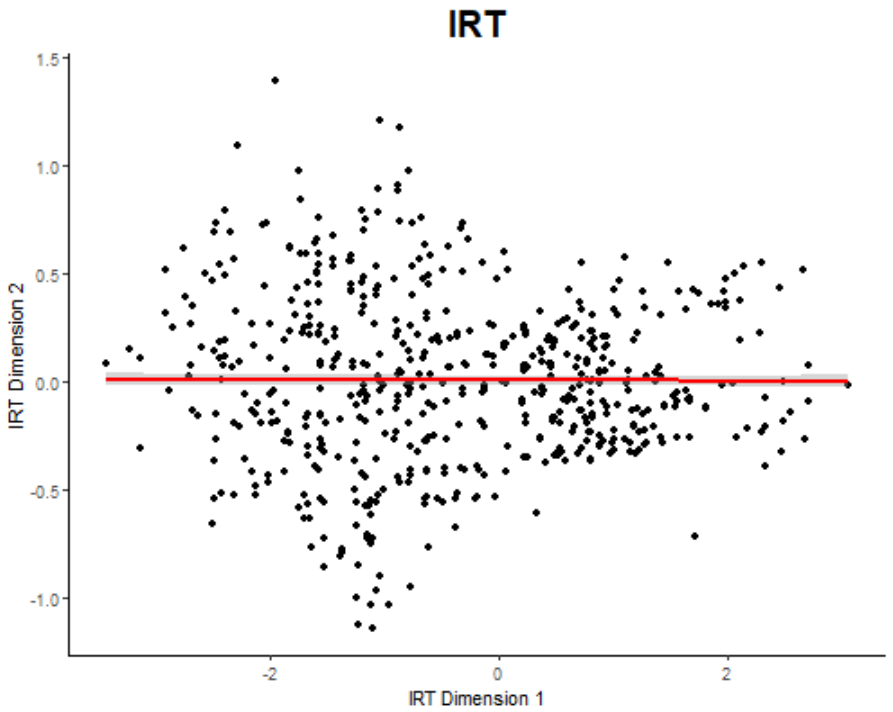
$R^2 = 0.43$
$t = 0.3s$

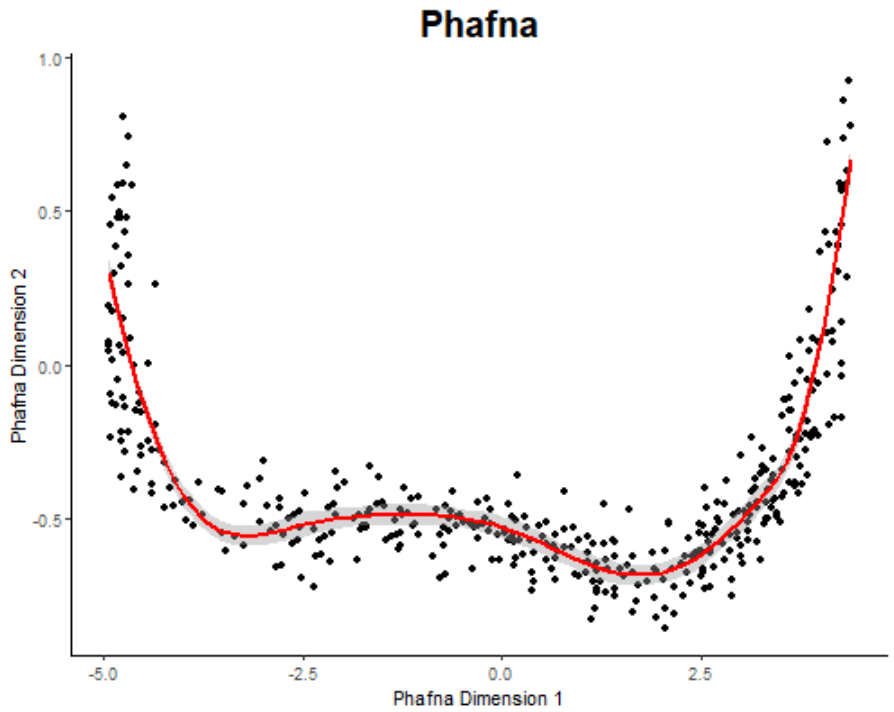




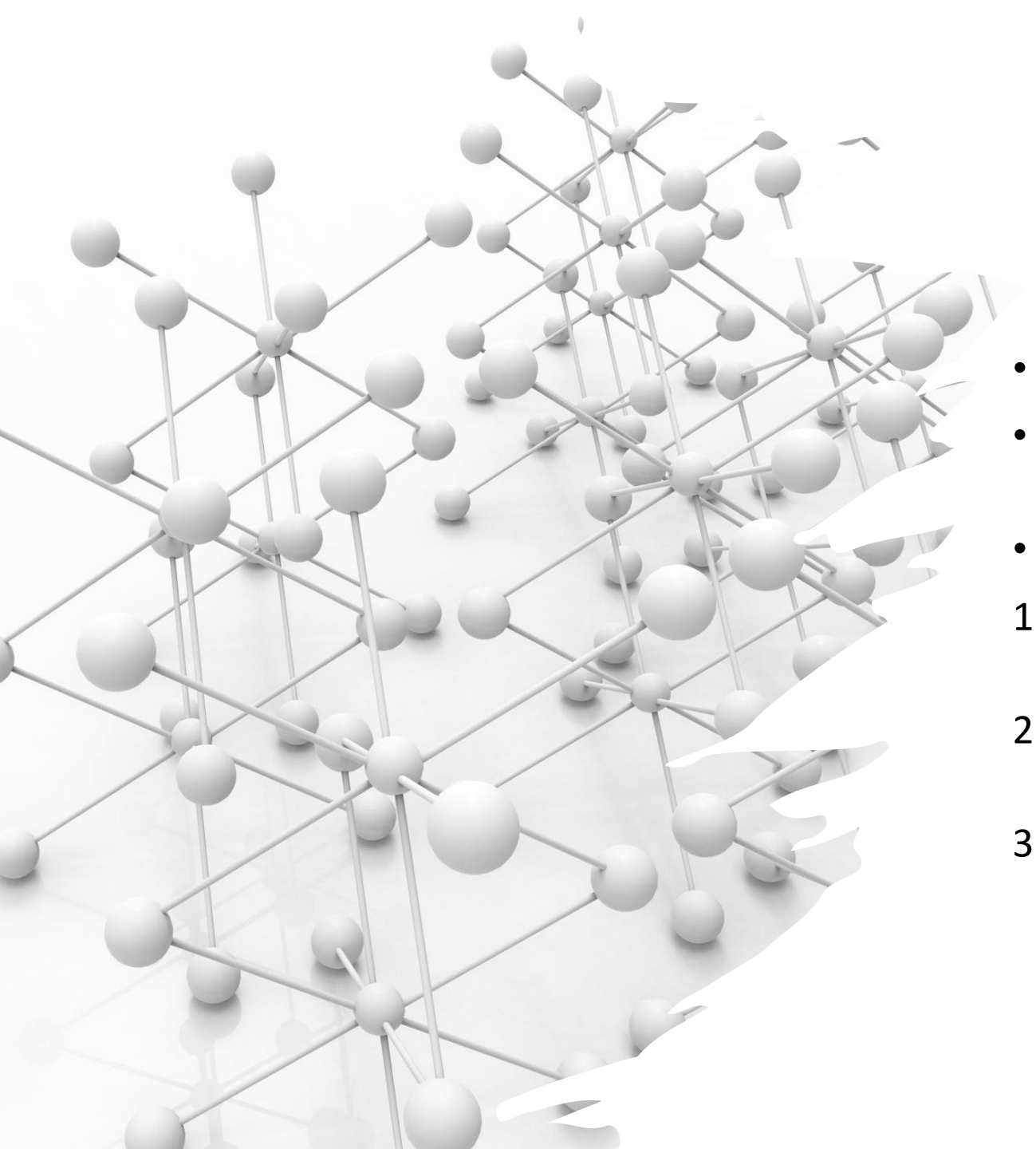
$R^2 = 0.32$
$t = 3.4s$

$R^2 = 0.4$
$t = 8.1s$





$R^2 = 0.86$
$t = 2.3s$



Introducing PhaFNA

- Physics-aided, Factor Network Analysis
- Combines elements from factor analysis, IRT models, and statistical belief-network analysis
- Three basic ideas:

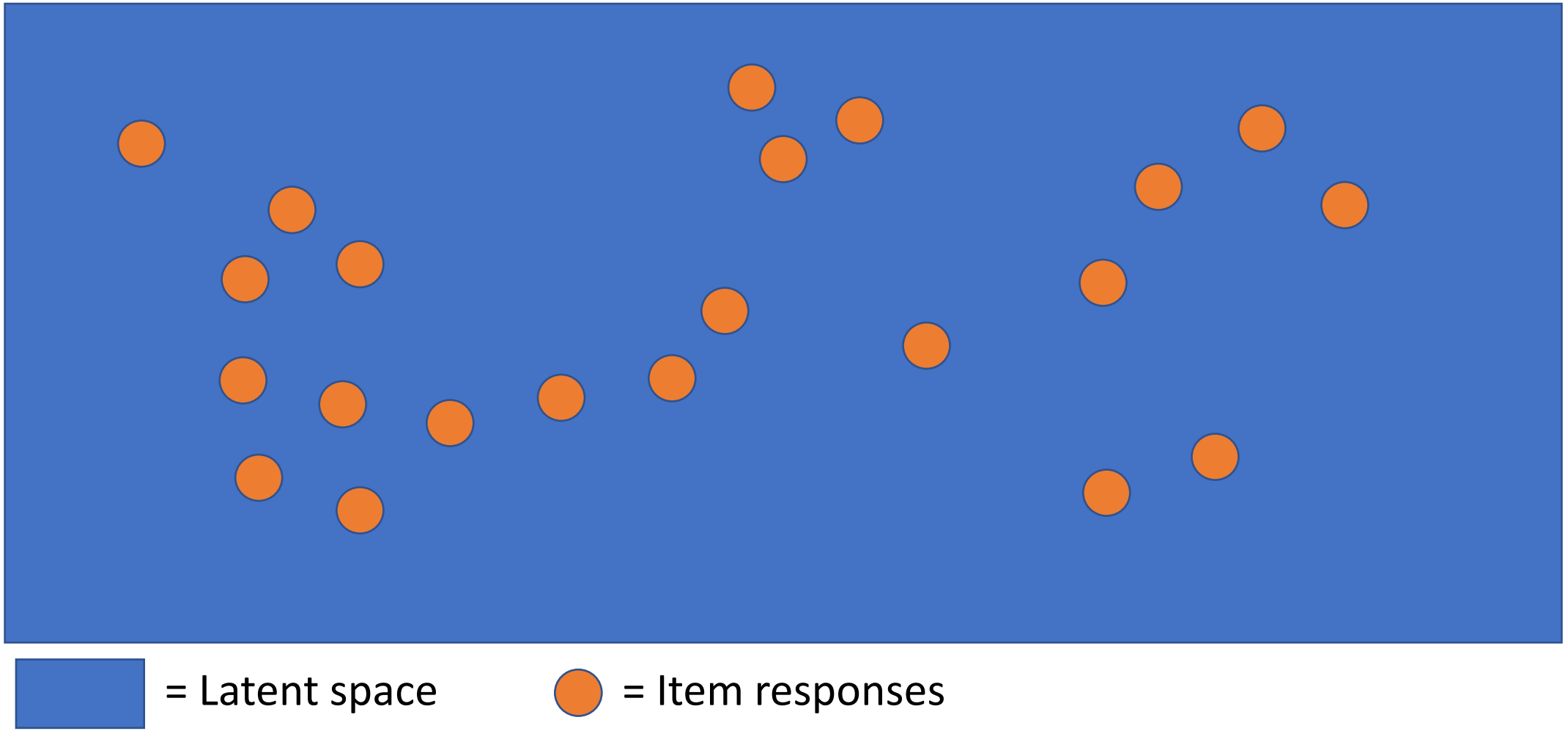
1. Treat item responses as physical particles that interact in latent space
2. Network communities of items physically anchor factors in latent space
3. Treats latent variables (i.e. factors) separate from latent space

Difference between latent factors and latent space

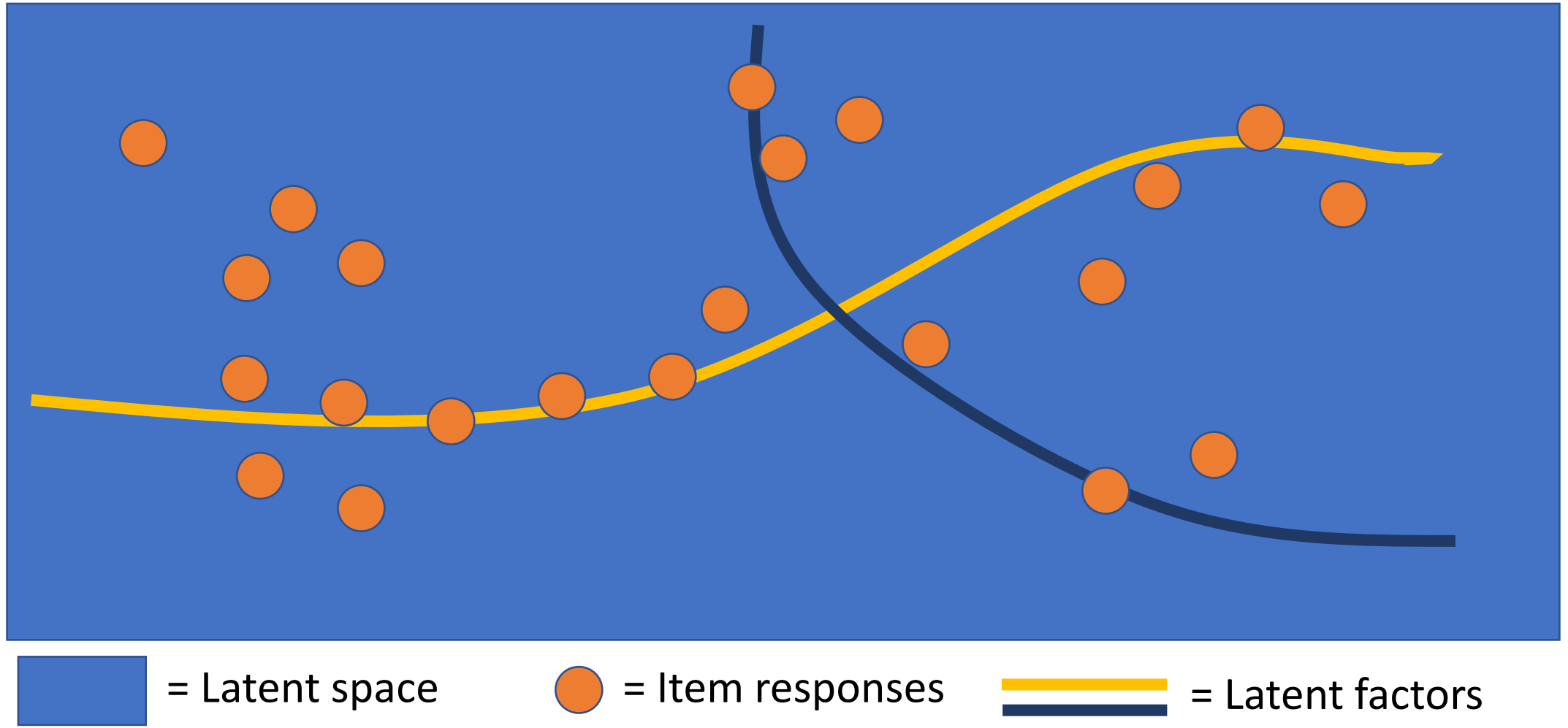
- Researchers commonly assume that latent factors live somewhere in a (wider) latent space
- Latent factors **span** the latent space in all conventional latent variable models
  - Assumptions about linearity and dimensionality **baked into** latent space
- **PhaFNA models differ** in that
  - Latent space and latent factors are conceptually separate
  - Latent space “just is” (predefined space)
  - Latent factors cut across the latent space in any possible direction

Latent space in PhaFNA

Latent space in PhaFNA



Latent space in PhaFNA



The four forces in PhaFNA

The four forces in PhaFNA

- **Spring-force**: positively correlated item-responses attract each other (increases with distance)

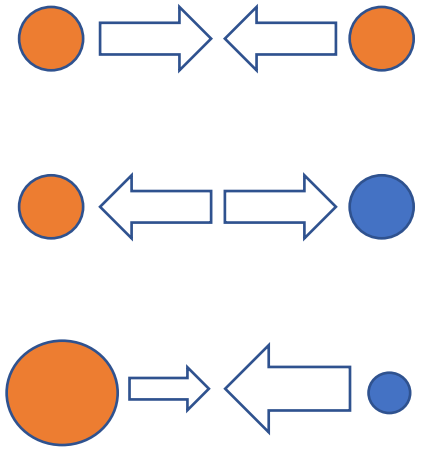
The four forces in PhaFNA

- **Spring-force**: positively correlated item-responses attract each other (increases with distance)

- **Electromagnetism**: negatively correlated item-responses repulse each other (decreases with distance)

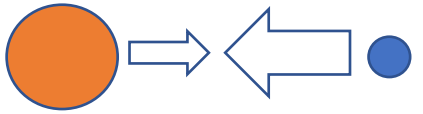
The four forces in PhaFNA

- **Spring-force**: positively correlated item-responses attract each other (increases with distance)
- **Electromagnetism**: negatively correlated item-responses repulse each other (decreases with distance)
- **Gravity:** more popular item responses are more massive (decreases with square of distance)



The four forces in PhaFNA

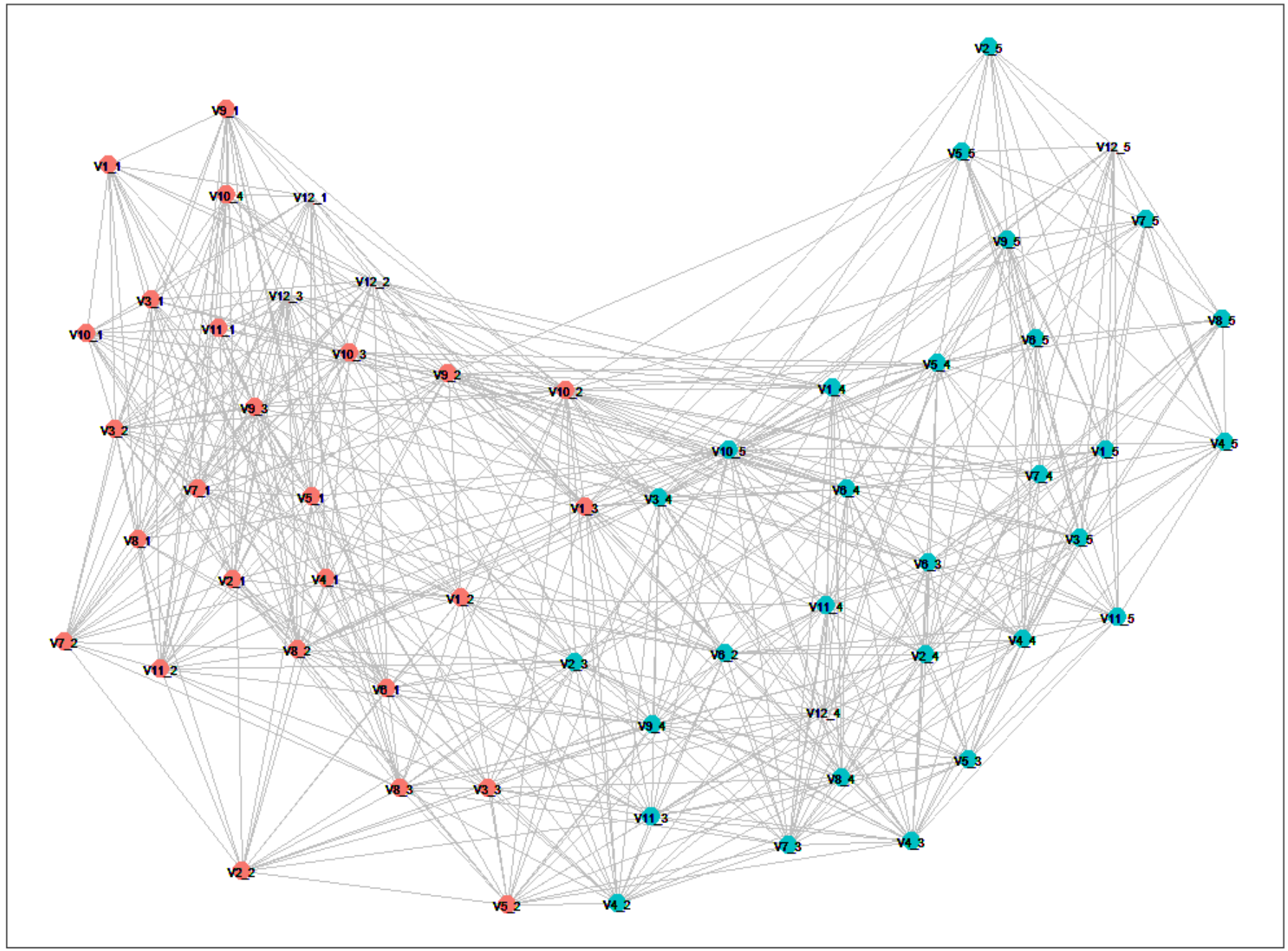
- **Spring-force**: positively correlated item-responses attract each other (increases with distance)
- **Electromagnetism**: negatively correlated item-responses repulse each other (decreases with distance)
- **Gravity:** more popular item responses are more massive (decreases with square of distance)
- **Drag:** All item-response particles lose momentum over time (ensures faster model convergence)

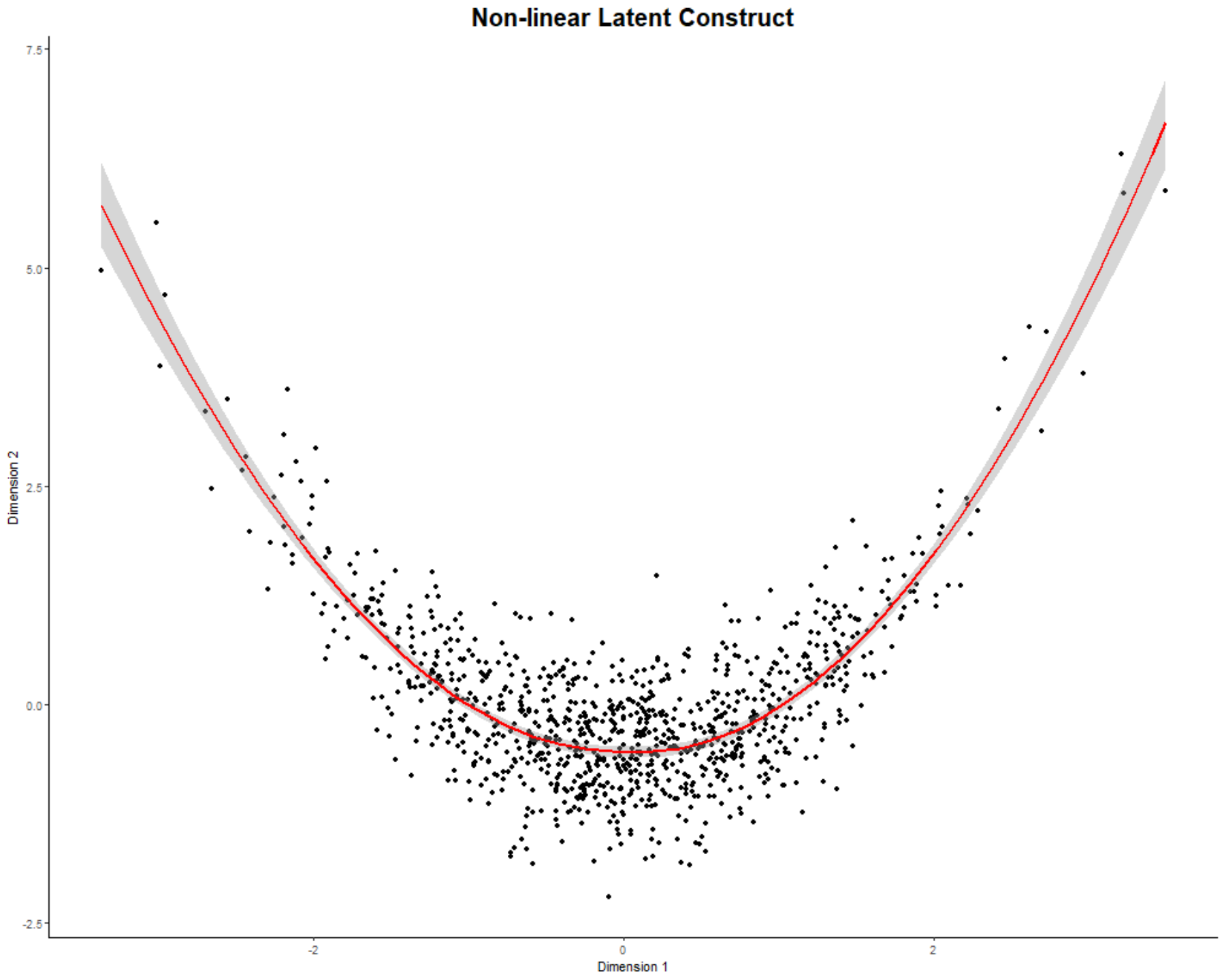


Demonstration

https://philip-warncke.netlify.app/files/demonstration.html

PhaFNA item plot

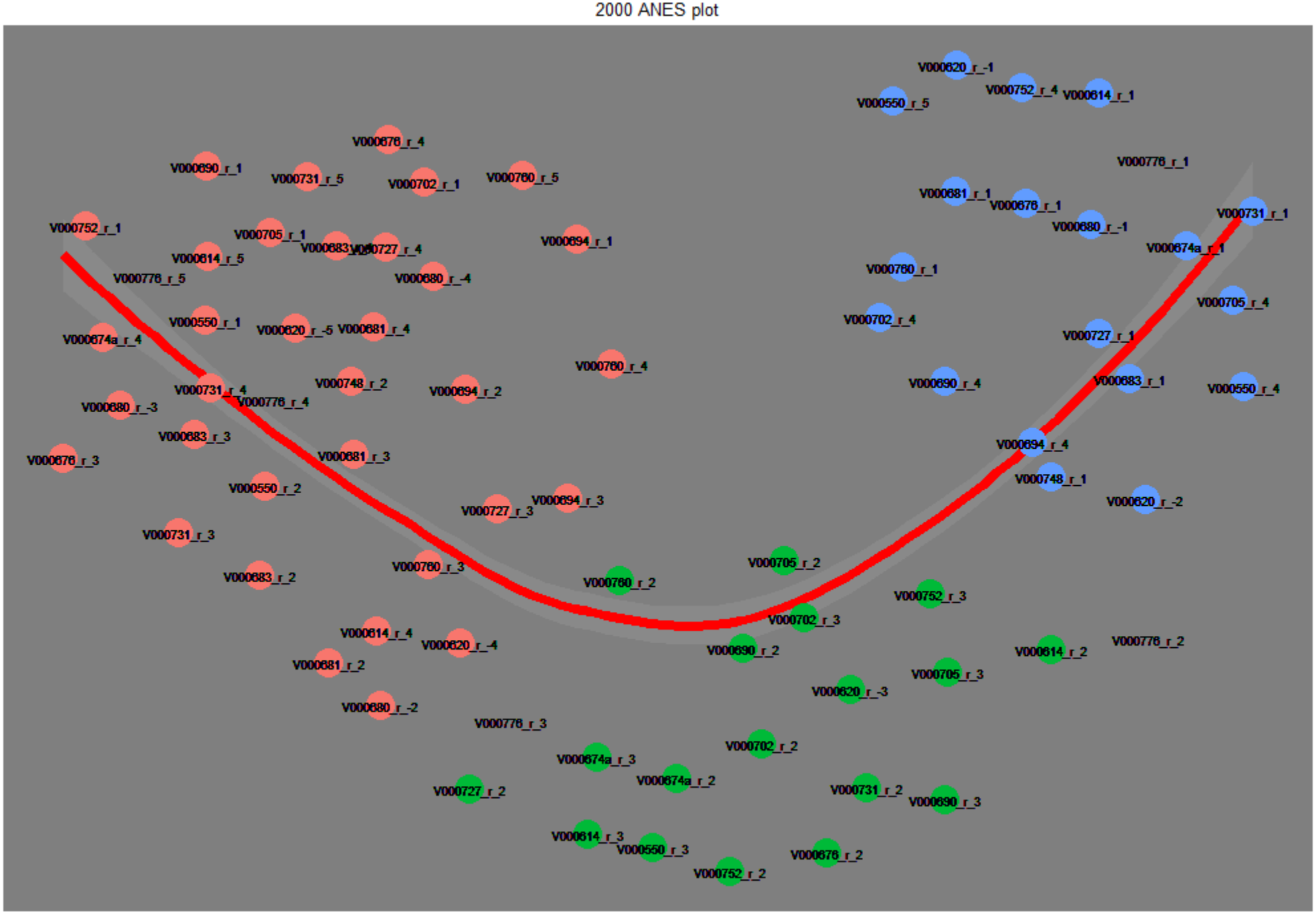



Non-linear Latent Construct
7.5
5.0
2.5
0.0
-2.5
Dimension 2
-2
0
2
Dimension 1

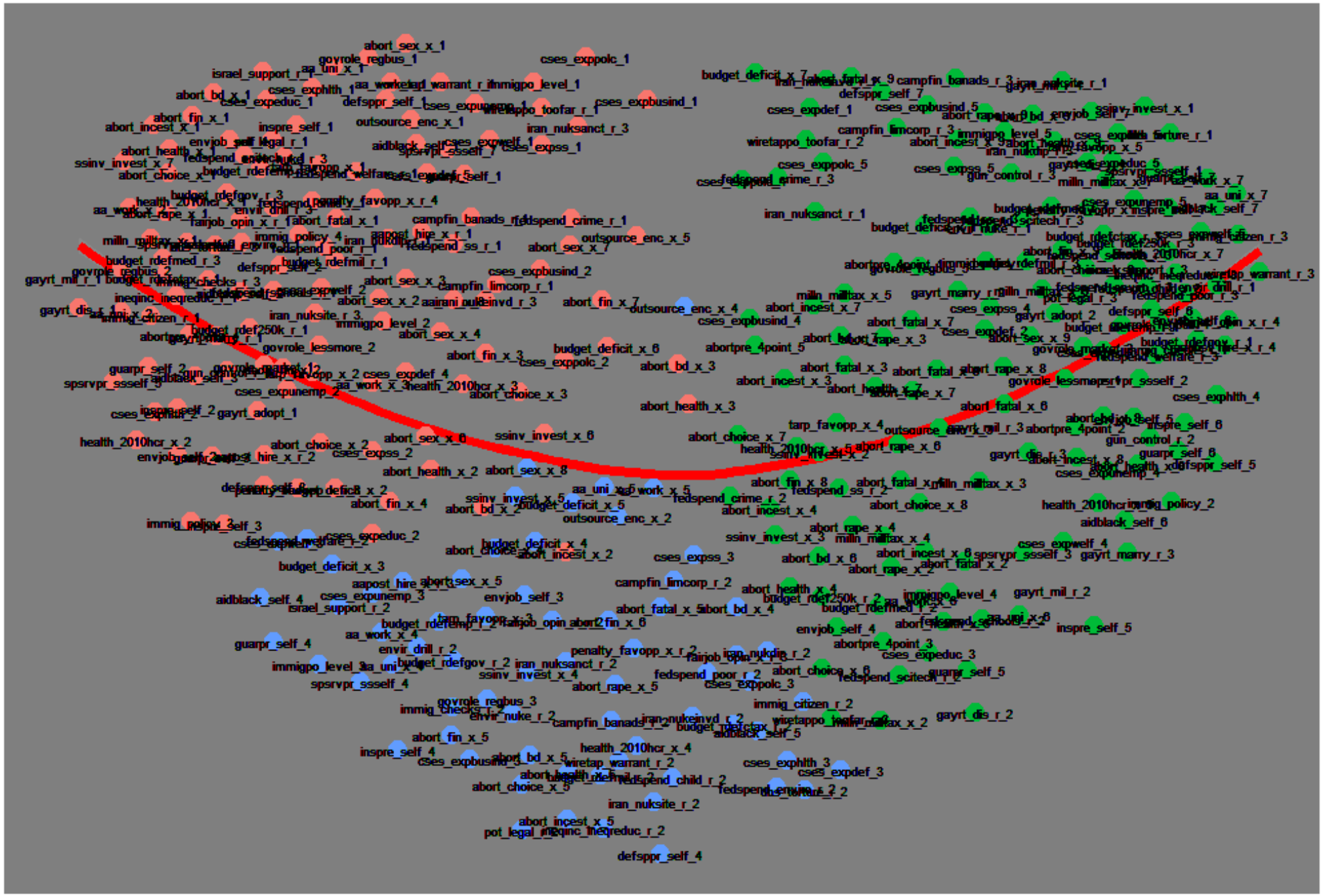
Applications

- Mapping the ideological space
- Tracing polarization across time

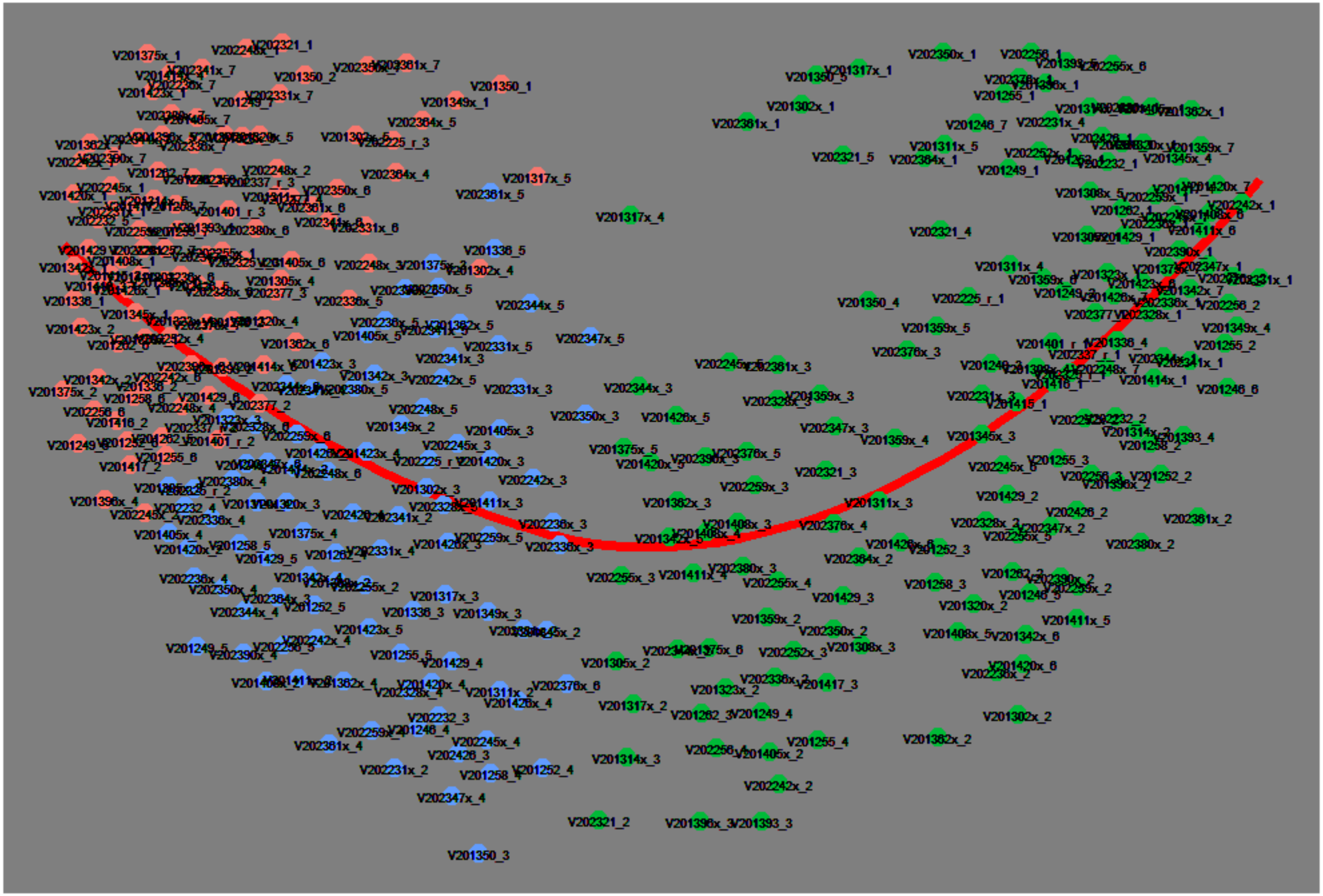
ANES 2000


2000 ANES plot
V000620_r_-1
V000752_r_4
V000614_r_1
V000550_r_5
V000676_r_4
V000690_r_1
V000731_r_5
V000702_r_1
V000760_r_5
V000776_r_1
V000681_r_1
V000676_r_1
V000731_r_1
V000752_r_1
V000705_r_1
V000680_r_-1
V000674a_r_1
V000614_r_5
V000727_r_4
V000694_r_1
V000760_r_1
V000776_r_5
V000680_r_-4
V000705_r_4
V000550_r_1
V000620_r_-5
V000681_r_4
V000702_r_4
V000727_r_1
V000674a_r_4
V000760_r_4
V000731_r_4
V000748_r_2
V000694_r_2
V000690_r_4
V000683_r_1
V000550_r_4
V000776_r_4
V000680_r_-3
V000683_r_3
V000694_r_4
V000676_r_3
V000681_r_3
V000748_r_1
V000550_r_2
V000694_r_3
V000620_r_-2
V000727_r_3
V000731_r_3
V000705_r_2
V000760_r_3
V000683_r_2
V000760_r_2
V000752_r_3
V000702_r_3
V000614_r_4
V000620_r_-4
V000690_r_2
V000614_r_2
V000776_r_2
V000681_r_2
V000705_r_3
V000620_r_-3
V000680_r_-2
V000776_r_3
V000702_r_2
V000674a_r_3
V000674a_r_2
V000727_r_2
V000731_r_2
V000690_r_3
V000614_r_3
V000550_r_3
V000676_r_2
V000752_r_2

ANES 2012

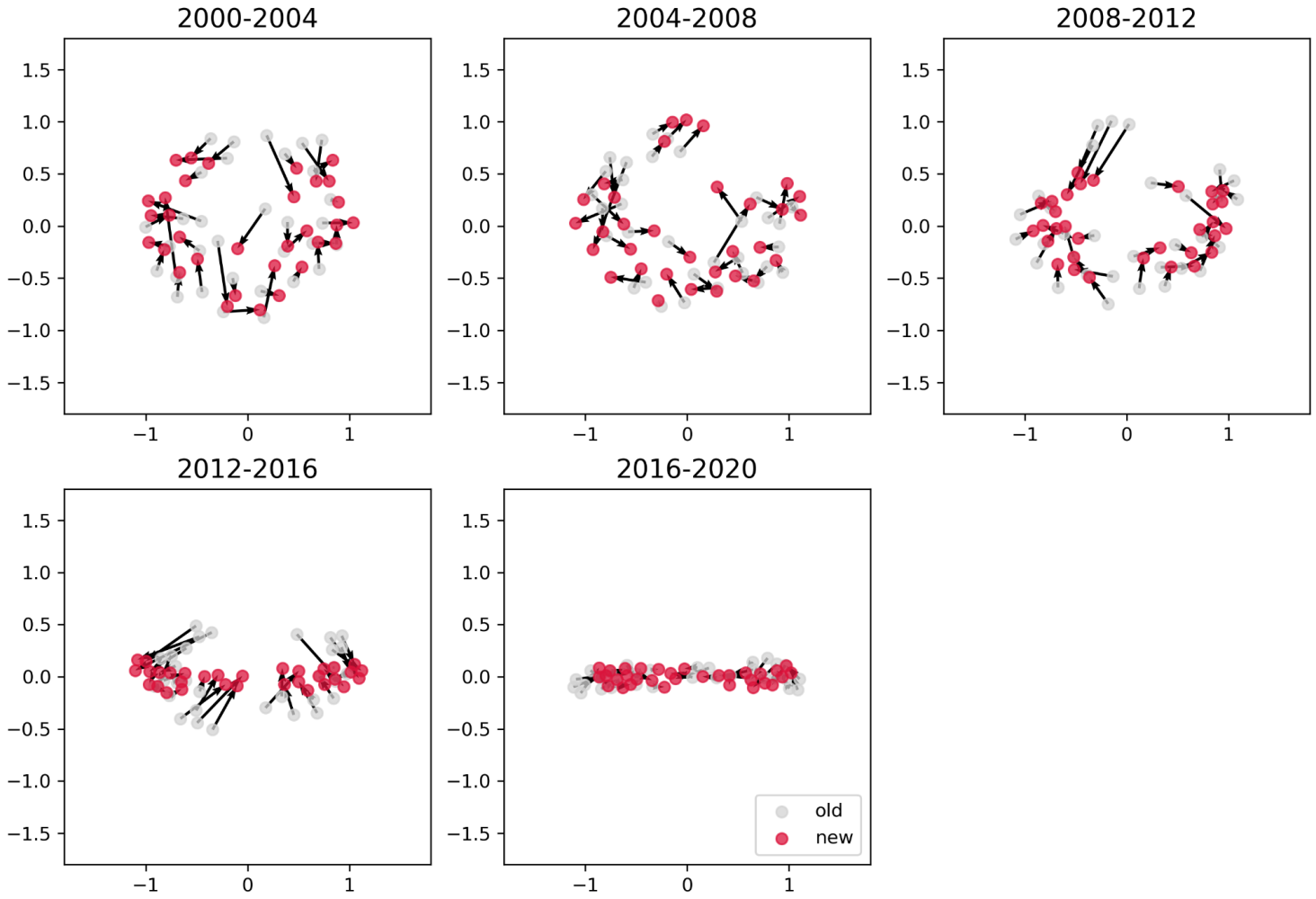


ANES 2020



Mapping Polarization with Physics

Co-authored with Dino Carpentras (ETH Zurich), Yijing Chen (CEU), Bart de Bruin (Leiden), and Anne Speer (BIGSSS Bremen)



Conclusions

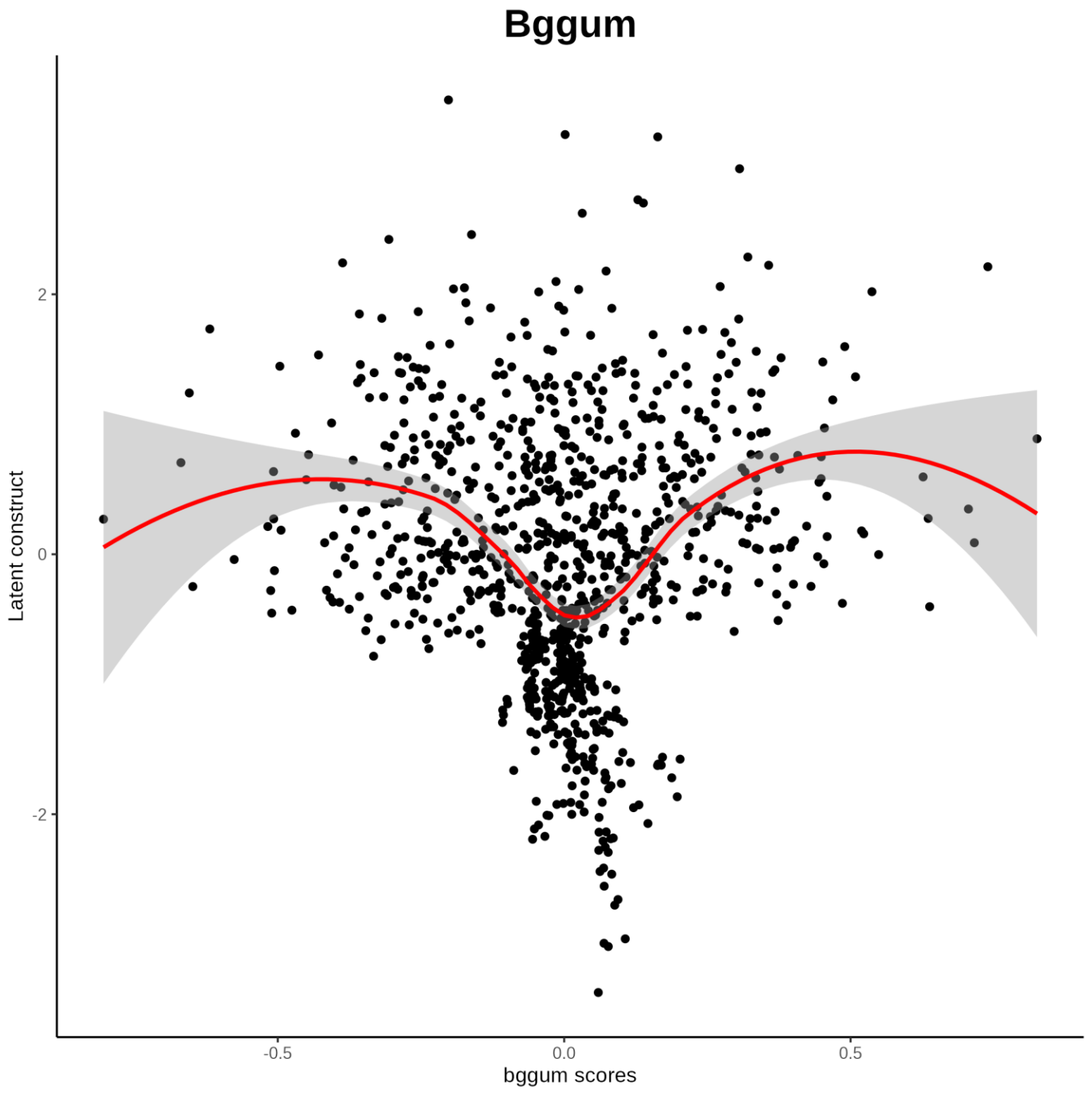
- Conventional latent variable models (LVM's) struggle when
  - LDGP's introduce non-linearities
  - Latent variables are non-linear
- PhaFNA is a novel latent variable modeling approach that simulates interactions between item responses in a physical latent space
- PhaFNA, more so than conventional LVM's is capable of
  - Restoring linear latent variables if LDGP's introduce non-linearities
  - Estimating non-linear latent variables

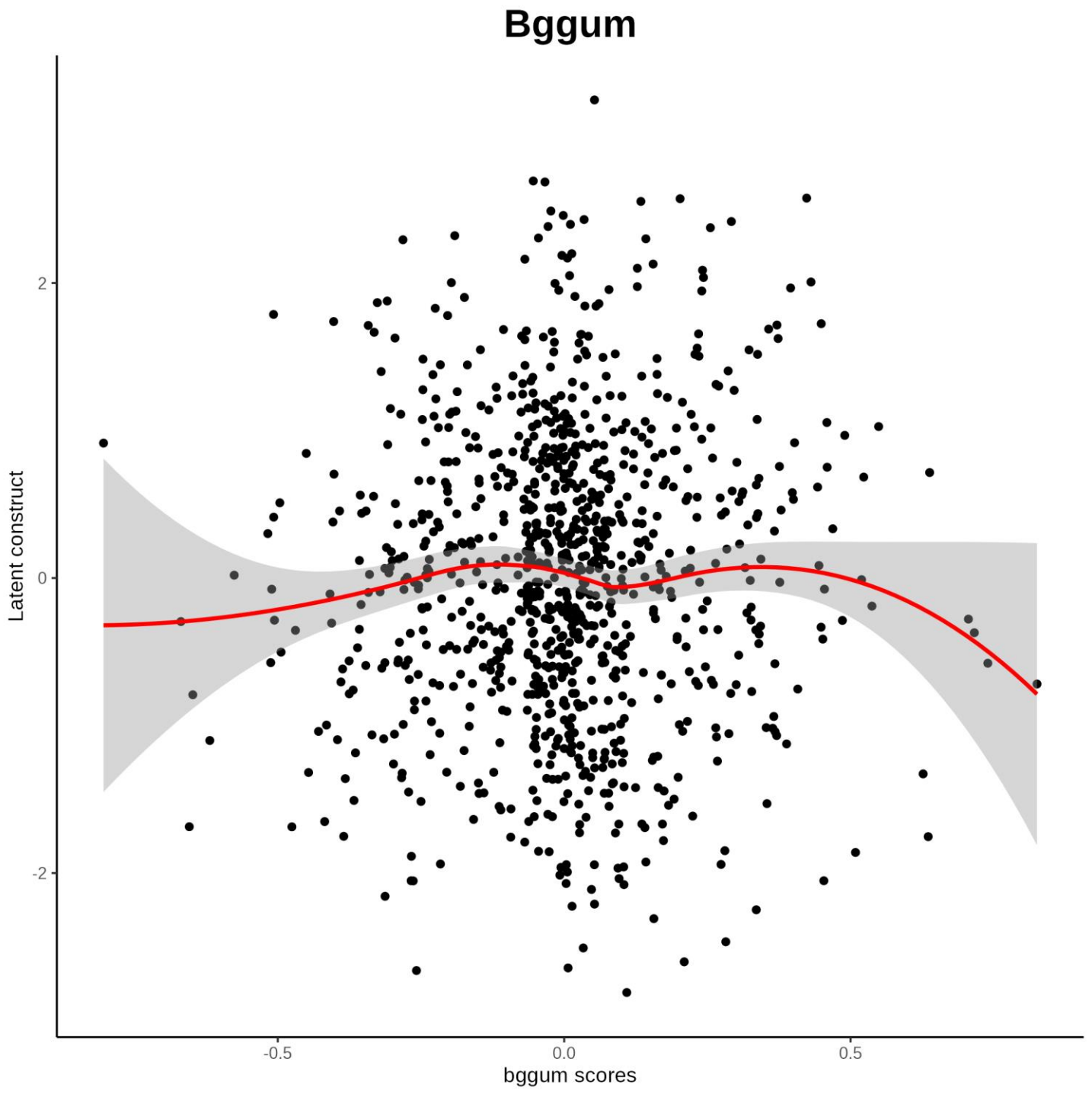
Next steps

- Developing diagnostic tools (non-linearities, multi-dimensionalities)
- Improving performance (e.g. implementation in Unity engine with GPU support)
- Implementation in higher dimensions

Please get in touch with collaboration ideas

- Philip Warncke (UNC Chapel Hill) pwarncke@live.unc.edu
- Dino Capentras (ETA Zurich) dino.carpentras@gmail.com
- Adrian Lüders (University of Hohenheim) Adrian.Lueders@uni-hohenheim.de
- Mike Quale (University of Limerick) mike.quayle@ul.ie


Bggum
Latent construct
2
0
-2
-0.5
0.0
0.5
bggum scores


Bggum
Latent construct
2
0
-2
-0.5
0.0
0.5
bggum scores