

Heterogeneity in the Pathways to Suicidal Ideation

Ross Jacobucci
University of Notre Dame



Heterogeneity

Challenge: Heterogeneity

Current suicide theory not only have limitations with respect to *time*, but also for *who*.

Relatively little research has been done assessing heterogeneity. And the research that has, mainly focused on:

- static group differences w/ respect to prevalence
- Cross-sectional or retrospective assessment

Little to no research has paired dynamics w/ heterogeneity:
Heterogeneity in pathways or risk.

What is Heterogeneity?

Nunes et al. (2020): "We define heterogeneity as the degree to which a system diverges from a state of perfect conformity."

Molenaar (2004) discusses heterogeneity in two realms:

- Idiographic vs. Nomothetic (Ergodicity)
- Psychological process across time (Stationarity)

Most of psychiatry and clinical psychology seems to be concerned with heterogeneity in symptom/diagnostic profiles. If using indexical items (i.e., Kendler 2007), can use combinatorics, if constitutive, then requires more complex factor analysis.

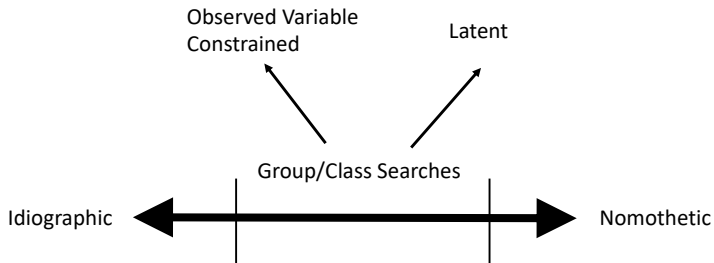
Heterogeneity As Pathways

Baltes and Nesselroade (1979): analyses of causes (determinants) of interindividual differences in intraindividual change.

Multifinality and equifinality refer to diversity in the origins, processes, and outcomes in development (Cicchetti & Rogosch, 1996).

Statistics

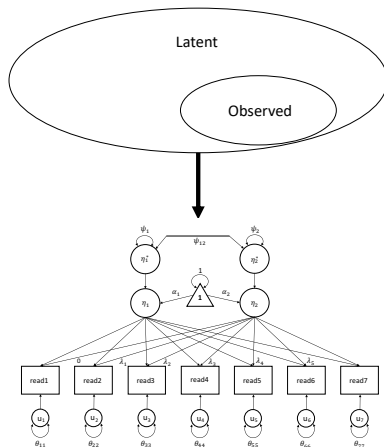
Nomothetic-Idiographic Continuum



Among these approaches, they can roughly be broken down into two groups:

- Latent
- Observed Variable Constrained

Latent Versus Observed



Latent

Latent approaches to identifying groups/classes include a large number of methods:

- Mixture Models
- Clustering
- Latent Dirichlet Allocation
- GIMME

The defining feature is that the latent variables are categorical, and observed variables do not place constraints on the categories.

Observed Variable Constrained

This encompasses the general tree-based methods, such as:

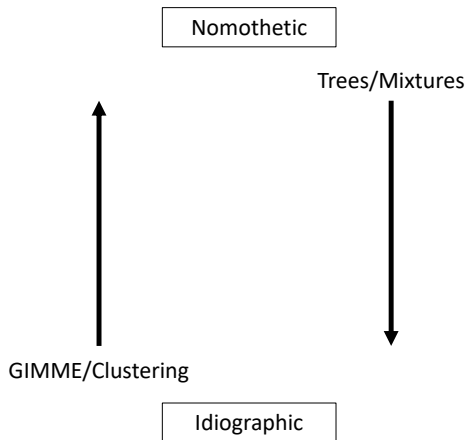
- SEM Trees
- Model Based Partitioning
- and many other derivatives.

A number of different options in R:

- `semtree`
- `longRPart`
- `longRPart2`
- `glmertree`
- `REEMtree`
- `MplusTrees`

These approaches will always fit worse than latent approaches, but with increasing numbers of covariates, particularly continuous, the discrepancy in fit will be minimized.

Bottom-Up vs. Top-Down



Study

Data Overview

A sample of 40 participants were recruited from online forums surrounding the topics of mental health and suicidality (i.e., Reddit Forums; Facebook Groups). All participants were required to be 18 years or older, have a past-year history of suicidal thoughts or behaviors.

The final sample, after requiring each person to have at least twenty responses, included data from 35 participants. Further, for the analyses, we excluded any responses that were more than 12 hours apart, resulting in a total of 2,030 unique responses. During the study period 27 unique participants reported suicidal thinking (> 1), and non-zero suicidal thinking was reported on 30.6% ($n = 520$) of prompts.

Goals

Goal: Identify heterogeneity in the lagged effects in predicting Active SI

This involves integrating two types of data:

- EMA (4x per day)
- Baseline data

EMA Data

Variables:

- Active SI – Sum of two Likert items used in prior research
- Negative Affect – Sum from PANAS-X
- Positive Affect – Sum from PANAS-X
- Belongingness – Sum of Loneliness Q and Thwarted Belongingness
- Burdensomeness – Sum of Perceived Burdensomeness and Feelings of Uselessness

Baseline Data

The following questions:

- Current Mental Health Treatment
- Been Hospitalized for Mental Health
- Gender – 6 Male, 18 Female, 4 Other
- Race
- Ethnicity
- PHQ Total

	Race
American Indian or Alaskan Native	1
Asian	4
Black or African American	2
Native Hawaiian or Pacific Islander	0
White	19
More than 1 race	1
Other	1
Prefer not to answer	0

Software 1: glmertree

The notable difference between `longRPart2` and `glmertree` is that `glmertree` estimates a single set of random effects parameters (i.e., Φ and σ_e^2) for all nodes and `longRPart2` estimates the full mixed-effects model within each node.

Used the `glmertree` package.

Following Settings:

- Maximum Depth = 2
- Minimum Node Size = 120

Software 2: longRPart2

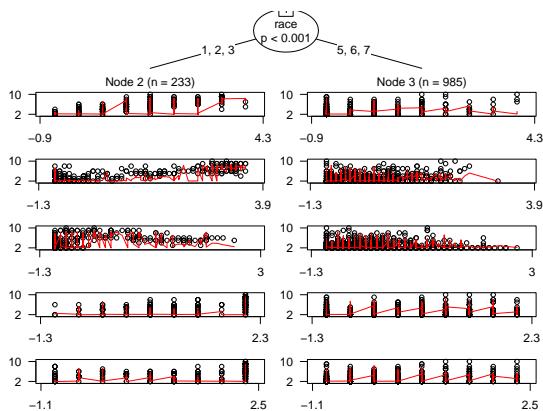
longRPart2 pairs nlme with rpart, examining the percentage change in the deviance. By default it accepts a split if more than a 1% improvement in root node deviance. Up to the user to determine this criterion.

10 parameters in the base model, so chi-square criterion is 18.3.

Following Settings:

- Maximum Depth = 2
- Minimum Node Size = 120

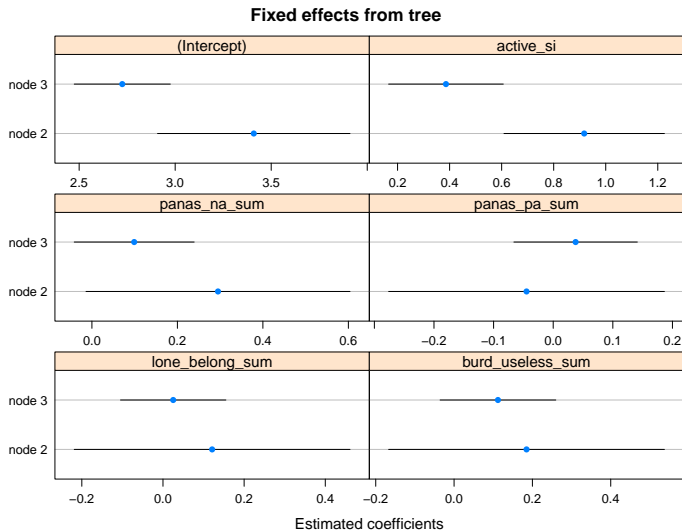
glmer tree Results 1



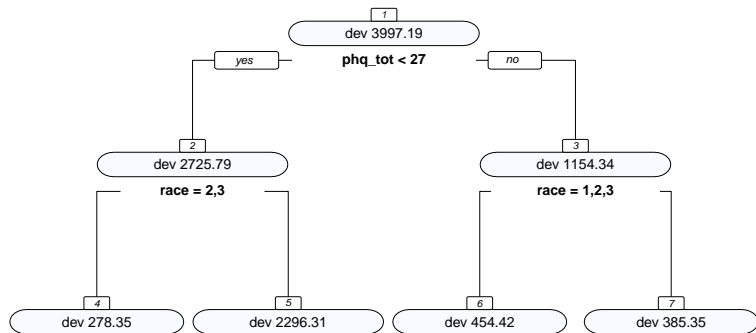
Node 2: American Indian, Black, Asian

Node 3: White, Multiple, Other

glmer tree Results 2



longRPart2 Results



longRPart2 Parameter Estimates

```
> summary(lrp.out)
$importance
      race  phq_tot  gender
2535.4606 2057.1264 158.8672

$parameters
      node  b0i      auto      na_b      pa_b  belong_b  burd_b  resid
1      2  3.007959 -0.2394245  0.096190722 -0.31913576  1.10770181 -0.17677276  0.9541036
2      3  2.637629  1.1309912 -0.172812077  0.03134407  0.14775248  0.10130314  0.7177409
3      4  2.966637  0.3071224  0.162180940  0.02943144  0.06841113  0.14395191  1.2271207
4      5  2.798430  0.8657663  0.226106823  0.09463952  0.01518529  0.06826428  0.1623336
5      6  3.935244  0.8022912  0.162584989 -0.27929310  0.18636253  0.37017899  1.1699389
6      7  2.393291  0.4919139  0.002662265  0.06046089 -0.09695662  0.06563788  0.4531702
```

Node 2: Lower Depression

Node 3: Higher Depression

Node 4: Race = American Indian, Asian or Black, Lower Depression

Node 5: Race = White, Lower Depression

Node 6: Race = American Indian, Asian or Black, Higher Depression

Node 7: Race = White, Higher Depression

Conclusion

Results Summary

- Lower Depression = belongingness stronger link
- Higher Depression = more consistency
- Non-White and Higher Depression = More varied cross-lag

Limitations

Can't discuss heterogeneity in risk or pathways without clarifying the role of time. Ultimately dynamics and heterogeneity are intertwined.

- Individual's processes may be happening at different time scales, or varying functional forms

Tree algorithms: Under-evaluated.

Study Design for Heterogeneity

In most methods for identifying heterogeneity, need a sufficient sample size for each potential grouping.

This means heterogeneity requires larger sample sizes than nomothetic or idiographic.

Limitations with P or T has significant consequences.

Thank You!

More info: rjacobucci.com