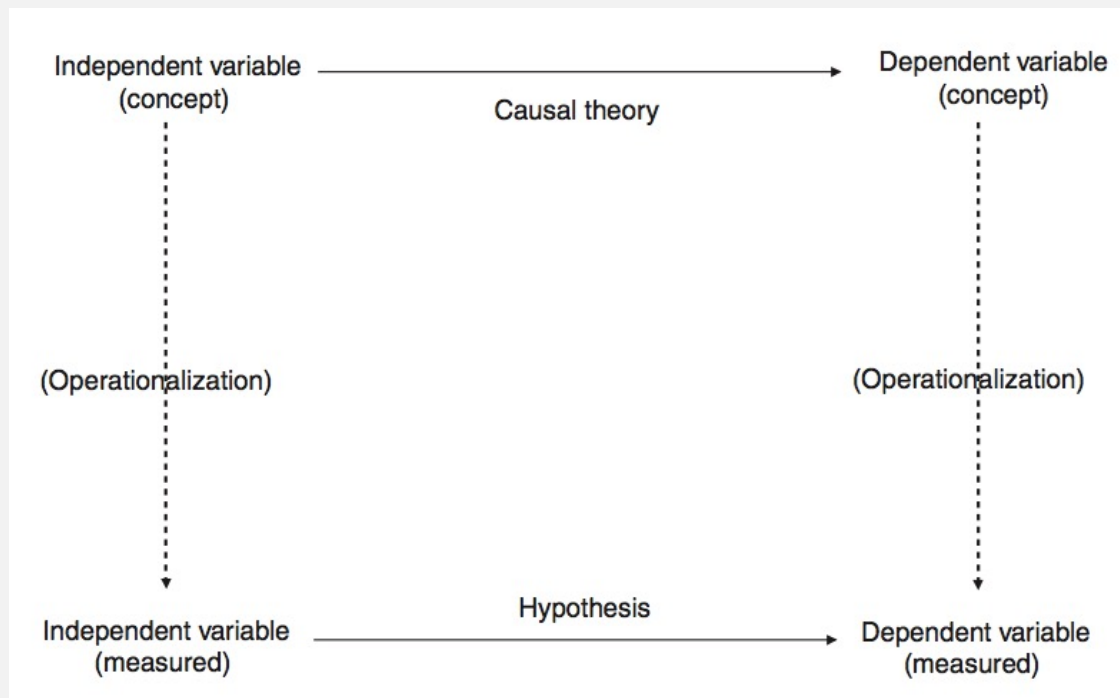# SCALING (AND MEASUREMENT)

SICSS-Edinburgh

June 2023
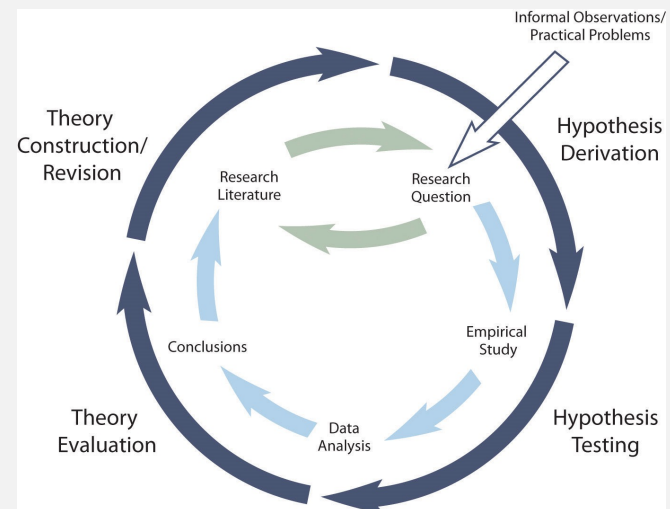
Dr. Ugur Ozdemir

FROM THEORY TO HYPOTHESIS

RESEARCH STAGES

- Deductive stages
  - Read existing theory
  - Formulate research question
  - Conceptualisation
  - Formulate hypotheses
  - Operationalisation
  - Data collection
  - Test hypotheses
  - Reflect back on theory
  - Publish results
  - Repeat!
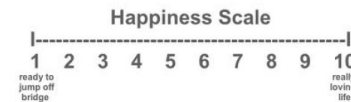
MEASUREMENT IN SOCIAL SCIENCES

How? [instruments]

What? Grades Emotions Practices Attitudes Social Class Behaviour ...

Happiness Scale

HAIR AND ELECTORAL CHANCES

- Does baldness affect electoral success?

  - Lee Sigelman et al (1990) "Hair Loss and Electability: The Bald Truth" *Journal of Nonverbal Behavior* 14(4): NO

  - Susan Banducci et al (2003) "Candidate Appearance Cues in Low-Information Elections" *APSA Conference Paper*: YES

	Operationalising baldness	**Operationalising electoral success**
General	Are they bald? Yes-No Assess the following: total bald, partially bald, full head of hair	Did they win? Yes – No If they lost, did they win back their deposit? Yes-No
Specific	% of total head area covered by hair	What % votes did they receive?

RISK AND SUPPORT FOR INDEPENDENCE

- Do risk attitudes affect support for independence?
  - Blais et al YES
  - Clarke et al NO

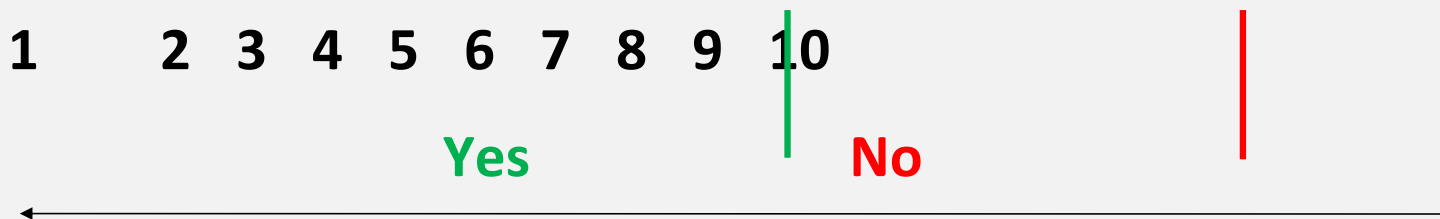|	Operationalising risk	Operationalising support for independence
General	Do you like taking risks? Yes-No	Do you support independence>? Yes-No
Specific	How much do you agree with the following statement: I like taking risks? Agree strongly, agree, neither agree nor disagree, disagree, disagree strongly	

OR …

KEY: AGGREGATION

- Many interesting social scientific measures involve some kind of aggregation and or scaling.
  - How to measure democracy?
  - Can you measure authoritarian values from multiple Likert scale research questions?
  - Can you deduce locations from pairwise distances?
  - How can we estimate ideologies of MPs from their voting histories?
- Other conventional methods? PCA? Factor analysis?

LAND OF THE FREE

Brasil

USA
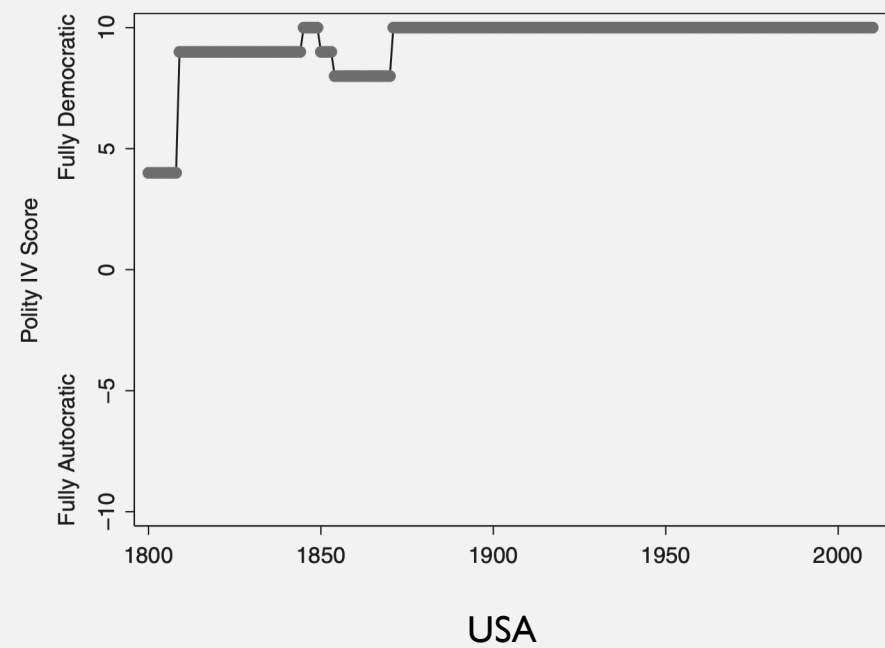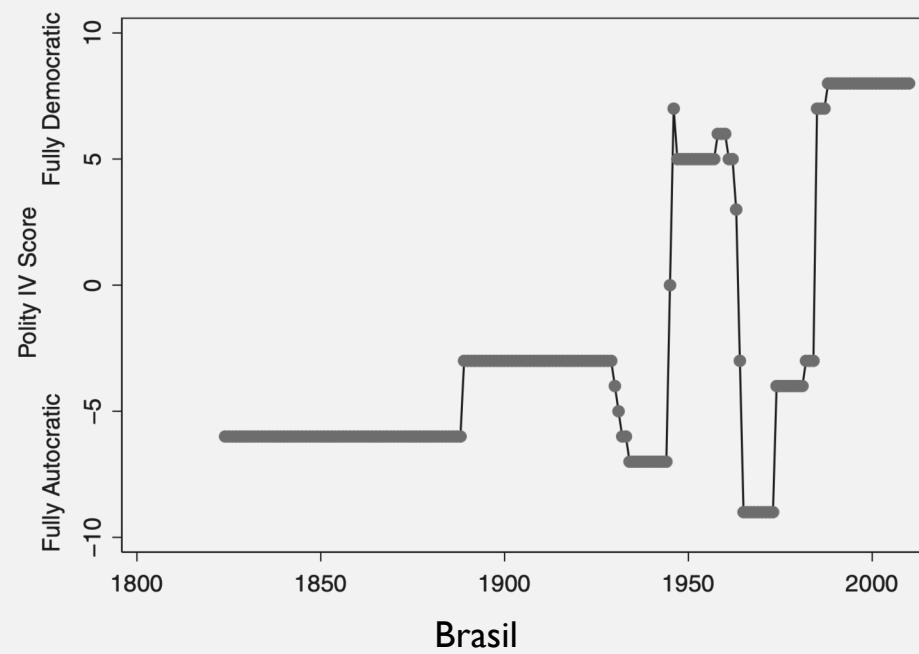
OUTLINE OF THE METHODS

1. Issue Scaling
    i. Aldrich-McKelvey Scaling
    ii. Basic Space Scaling
2. Multidimensional Scaling: SMACOF
3. Unfolding of Binary Data: Optimal Classification

A-M SCALING

A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections

JOHN H. ALDRICH
Michigan State University

RICHARD D. MCKELVEY
Carnegie-Mellon University

Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions

Christopher Hare   University of Georgia
David A. Armstrong II   University of Wisconsin-Milwaukee
Ryan Bakker   University of Georgia
Royce Carroll   Rice University
Keith T. Poole   University of Georgia

Aldrich and McKelvey's (1977) pathbreaking solution to the problem of DIF is to treat raw placements as linear distortions of the "true" positions of the stimuli (e.g., political parties and candidates). By estimating

In brief, differential item functioning (DIF) occurs when groups (such as defined by gender, ethnicity, age, or education) have different probabilities of endorsing a given item on a multi-item scale after controlling for overall scale scores.

A-M SCALING

	Self-Placement	Stimuli 1	Stimuli 2	.	.	Stimuli K				Scaled Ideals
Ind 1	2	3	6	.	.	4		Ind 1		0.7
Ind 2	5	7	2	.	.	6		Ind 2		0.5
Ind 3	6	4	5	.	.	3		Ind 3		2.4
.	.	.	.	.	.	.		.		.
.	.	.	.	.	.	.		.		.
Ind N	7	6	3	.	.	2		Ind N		1.5

A-M Scaling
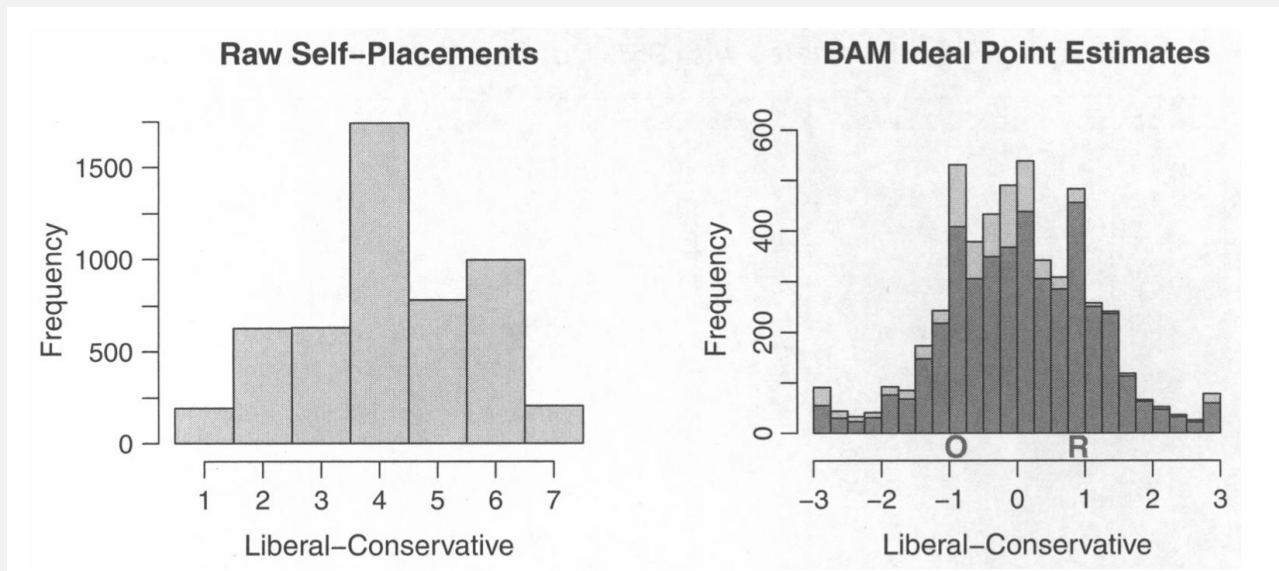
A-M SCALING

A-M SCALING

The `aldmck()` function requires five arguments: the matrix to be analyzed (`franceEES2009`), the column number for respondent self-placements (1), the column number for a stimuli to be placed on the left side of the dimension (2 for the Extreme Left party), a vector of missing value codes, and a logical argument (`TRUE/FALSE`) that specifies whether verbose output is desired as the function is executed. The left stimulus requirement is a function of the recovered space being defined only up to a rotation. By convention, `aldmck()` places left-leaning stimuli on the left end of the scale by assigning them negative scores.

BASIC SPACE SCALING

	Issue 1	Issue 2	Issue 3	.	.	Issue K
Ind 1	2	3	6	.	.	4
Ind 2	5	7	2	.	.	6
Ind 3	6	4	5	.	.	3
.	.	.	.	.	.	.
.	.	.	.	.	.	.
Ind N	7	6	3	.	.	2

Basic Space Scaling

→

(K > M)

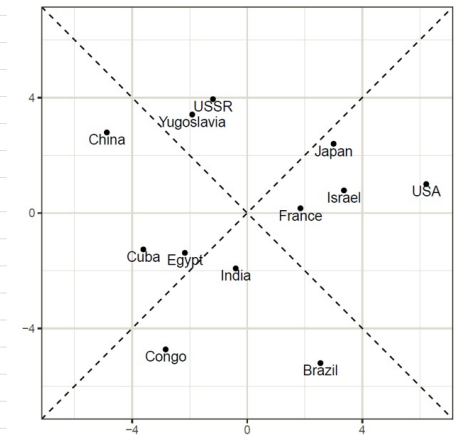|	Dimension 1	Dimension 2	Dimension 3	.	.	Dimension M
Ind 1	0.7	-1.2	2.1	.	.	0.9
Ind 2	0.5	-0.3	-0.8	.	.	1.3
Ind 3	2.4	1.3	-2.6	.	.	0.4
.	.	.	.	.	.	.
.	.	.	.	.	.	.
Ind N	1.5	0.5	-0.7	.	.	-1.4

BASIC SPACE SCALING

Generally, the `blackbox()` function is used to scale individuals from preference data. For example, survey respondents state their preferred policy outcome on multiple issue scales, and the `blackbox()` procedure recovers the ideal points of these individuals in an s-dimensional basic space. Conversely, the `blackbox_transpose()` function is used to estimate the latent coordinates of stimuli that are rated by individuals (i.e., based on perceptual data). For example, a set of party experts rank the positions of European political parties across a set of issue scales. The `blackbox_transpose()` function transposes the matrix, placing the stimuli on the rows and the individuals on the columns—since in these cases we want to scale the stimuli. Despite this distinction, both functions are applications of the same underlying method, which we detail below.

MULTIDIMENSIONAL SCALING

	Brazil	Congo	Cuba	Egypt	France	India	Israel	Japan	China	USSR	USA	Yugoslavia
Brazil	9.00	4.83	5.28	3.44	4.72	4.50	3.83	3.50	2.39	3.06	5.39	3.17
Congo	4.83	9.00	4.56	5.00	4.00	4.83	3.33	3.39	4.00	3.39	2.39	3.50
Cuba	5.28	4.56	9.00	5.17	4.11	4.00	3.61	2.94	5.50	5.44	3.17	5.11
Egypt	3.44	5.00	5.17	9.00	4.78	5.83	4.67	3.83	4.39	4.39	3.33	4.28
France	4.72	4.00	4.11	4.78	9.00	3.44	4.00	4.22	3.67	5.06	5.94	4.72
India	4.50	4.83	4.00	5.83	3.44	9.00	4.11	4.50	4.11	4.50	4.28	4.00
Israel	3.83	3.33	3.61	4.67	4.00	4.11	9.00	4.83	3.00	4.17	5.94	4.44
Japan	3.50	3.39	2.94	3.83	4.22	4.50	4.83	9.00	4.17	4.61	6.06	4.28
China	2.39	4.00	5.50	4.39	3.67	4.11	3.00	4.17	9.00	5.72	2.56	5.06
USSR	3.06	3.39	5.44	4.39	5.06	4.50	4.17	4.61	5.72	9.00	5.00	6.67
USA	5.39	2.39	3.17	3.33	5.94	4.28	5.94	6.06	2.56	5.00	9.00	3.56
Yugoslavia	3.17	3.50	5.11	4.28	4.72	4.00	4.44	4.28	5.06	6.67	3.56	9.00

MDS

PROBLEMS WITH TRADITIONAL MEASUREMENT

- Since item facilities (proportions) are bounded by zero and one, they cannot form a scale at the interval level of measurement.

- Since items are not at the interval level of measurement, parametric statistics should not be used: means, standard deviations and statistics that depend on them.

- A total test score is obtained by simple summation: for example, Likert summation. This may not be justified.

WHY USE MORE SOPHISTICATED METHODS?

- Produces scales at the interval level of measurement.
- Assesses the dimensionality of scales.
- Measures the error associated with each case.
- Measures the consistency of the pattern of responses for each case.
- Enables persons to be measured using different sets of items.
- Handles cases with missing data.