Psychology 405: Psychometric Theory

William Revelle
Northwestern University

Spring, 2003

http://pmc.psych.nwu.edu/revelle/syllabi/405.html

Psychometric Theory

- 'The character which shapes our conduct is a definite and durable 'something', and therefore… it is reasonable to attempt to measure it. (Galton, 1884)
- "Whatever exists at all exists in some amount. To know it thoroughly involves knowing its quantity as well as its quality" (E.L. Thorndike, 1918)

Psychometric Theory: Goals

1. To acquire the fundamental vocabulary and logic of psychometric theory.
2. To develop your capacity for critical judgment of the adequacy of measures purported to assess psychological constructs.
3. To acquaint you with some of the relevant literature in personality assessment, psychometric theory and practice, and methods of observing and measuring affect, behavior, cognition and motivation.

Psychometric Theory: Goals II

4. To instill an appreciation of and an interest in the principles and methods of psychometric theory.
5. This course is not designed to make you into an accomplished psychometist (one who gives tests) nor is it designed to make you a skilled psychometrician (one who constructs tests), nor will it give you "hands on" experience with psychometric computer programs.

Psychometric Theory: Requirements

- Objective Midterm exam
- Objective Final exam
- Final paper applying principles from the course to a problem of interest to you.
- Sporadic applied homework and data sets

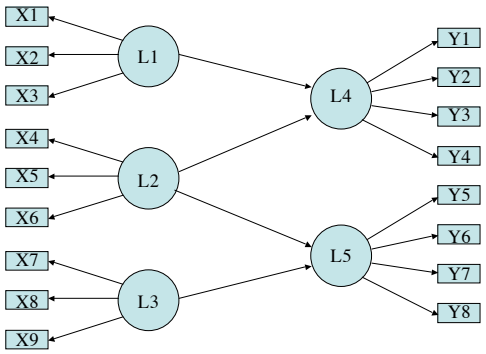
Text and Syllabus

- Nunnally, Jum & Bernstein, Ira (1994) *Psychometric Theory* New York: McGraw Hill, 3rd ed.(required: available at Norris)
- Loehlin, John (1998) *Latent Variable Models: an introduction to factor, path, and structural analysis* . Hillsdale, N.J.: LEA. (recommended)

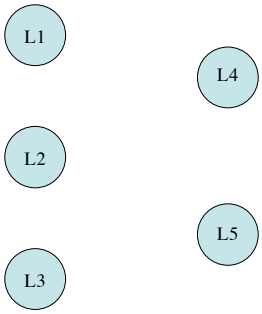
Syllabus: Overview

I. Individual Differences and Experimental Psychology
II. Models of measurement
III. Test theory
  A. Reliability
  B. Validity (predictive and construct)
  C. Structural Models
  D. Test Construction
IV. Assessment of traits
V. Methods of observation of behavior

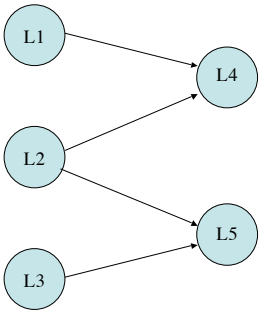
Psychometric Theory: A conceptual Syllabus



Constructs/Latent Variables



Theory as organization of constructs



Theory development and testing

- Theories as organizations of observables
- Constructs, latent variables and observables
  - Observables
    - Multiple levels of description and abstraction
    - Multiple levels of inference about observables
  - Latent Variables
    - Latent variables as the common theme of a set of observables
    - Central tendency across time, space, people, situations
  - Constructs as organizations of latent variables and observed variables

Theories as metaphors and analogies-1

- Physics
  - Planetary motion
    - Ptolemy
    - Galileo
    - Einstein
  - Springs, pendulums, and electrical circuits
  - The Bohr atom
- Biology
  - Evolutionary theory
  - Genetic transmission

Theories as metaphors and analogies-2

- Business competition and evolutionary theory
  - Business niche
  - Adaptation to change in niches
- Learning, memory, and cognitive psychology
  - Telephone as an example of wiring of connections
  - Digital computer as information processor
  - Parallel processes as distributed information processor

Examples of psychological constructs and their operationalization as observables

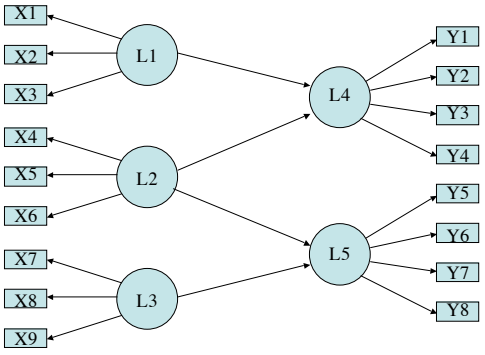
- Anxiety
  - Trait
  - State
- Love
- Conformity
- Intelligence
- Learning and memory
  - Procedural - memory for how
  - Episodic -- memory for what
    - Implicit
    - explicit

Models and theory

- Formal models
  - Mathematical models
  - Dynamic models - simulations
- Conceptual models
  - As guides to new research
  - As ways of telling a story
    - Organizational devices
    - Shared set of assumptions

Observables/measured variables

Psychometric Theory: A conceptual Syllabus



A Theory of Data: What can be measured

X1

What is measured?

Individuals

Objects

What kind of measures are taken?

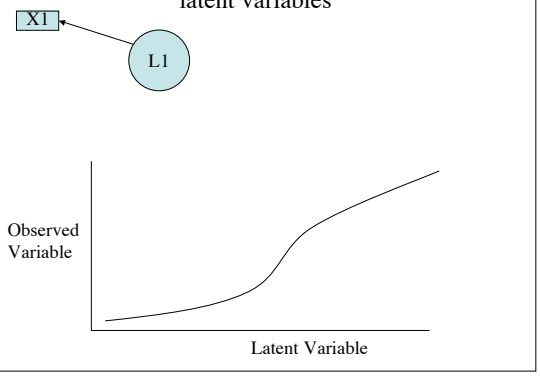
proximity

order

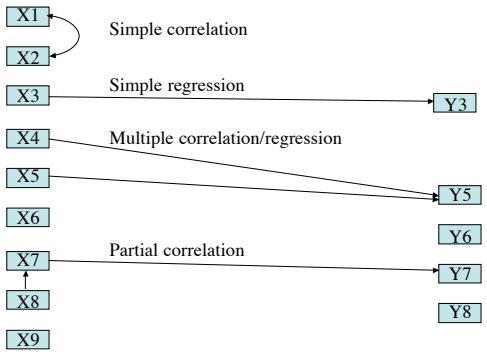
Comparisons are made on:

Single Dyads or Pairs of Dyads

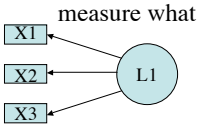
Scaling: the mapping between observed and latent variables



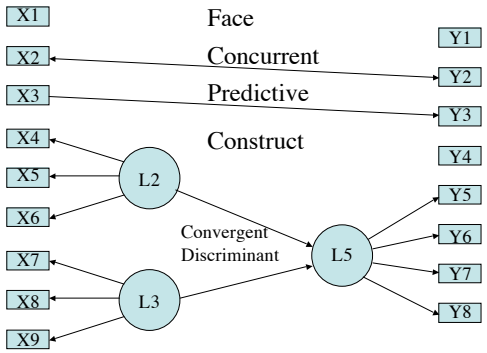
Variance, Covariance, and Correlation



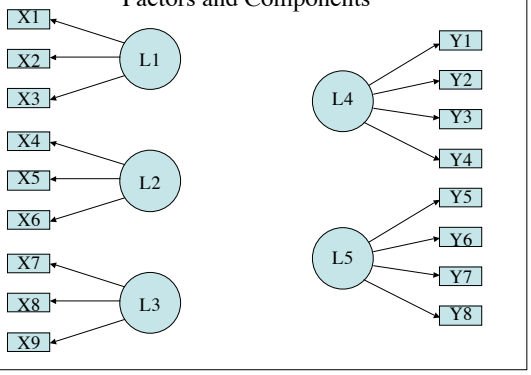
Classic Reliability Theory: How well do we measure what ever we are measuring



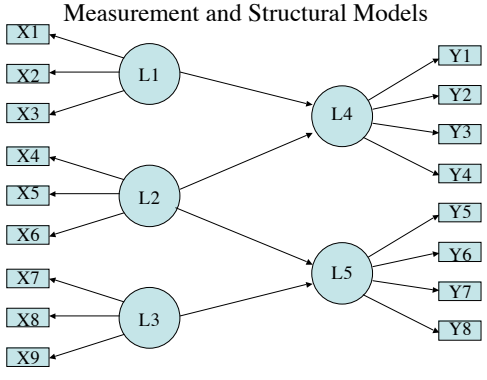
Types of Validity: What are we measuring



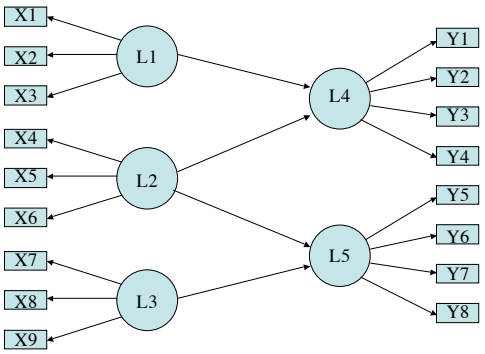
Techniques of Data Reduction: Factors and Components



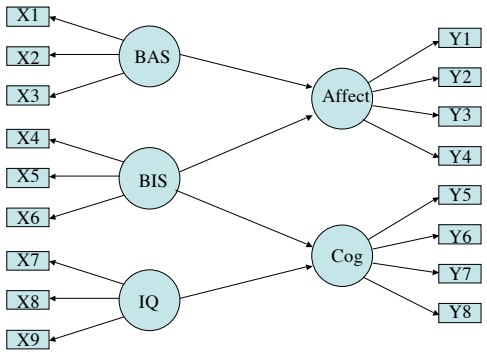
Structural Equation Modeling: Combining Measurement and Structural Models



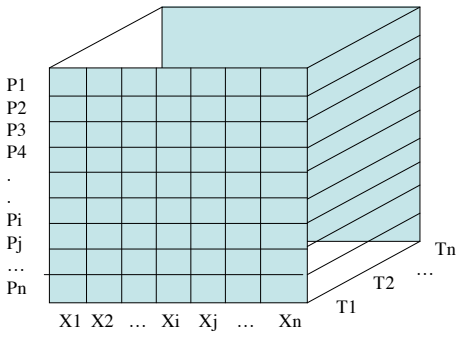
Scale Construction: practical and theoretical



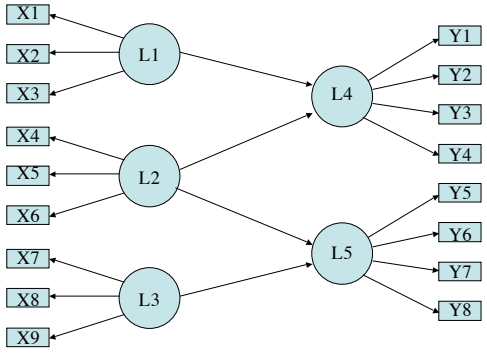
Traits and States: What is measured?



The data box: measurement across time, situations, items, and people



Psychometric Theory: A conceptual Syllabus



Syllabus: Overview

I. Individual Differences and Experimental Psychology
   - A. Two historic approaches to the study of psychology
   - B. Individual differences and general laws
   - C. The two disciplines reconsidered

II. Models of measurement
   - A. Theory of Data
   - B. Issues in scaling
   - C. Variance, Covariance, and Correlation

III. Test theory
   - A. Reliability
   - B. Validity (predictive and construct)
   - C. Structural Models
   - D. Test Construction

IV. Assessment of traits

V. Methods of observation of behavior

Two Disciplines of Psychological Research

(Cronbach, 1957, 1975; Eysenck, 1966, 1997)

B=f(Personality)	B=f(P*E)	B=f(Environment)
	Darwin	
Galton		Weber, Fechner
Binet, Terman		Watson, Thorndike
Allport, Burt	Lewin	Hull, Tolman
Cattell	Atkinson, Eysenck	Spence, Skinner
Epstein		Mischel

Two Disciplines of Psychological Research

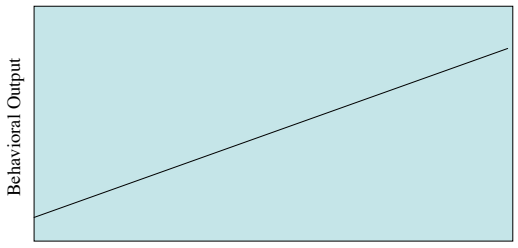
	B=f(Person)	B=f(Environment)
Method/ Model	Correlational Observational Biological/field	Experimental Causal Physical/lab
Statistics	Variance Dispersion Correlation/ Covariance	Mean Central Tendency t-test, F test
Effects	Individuals Individual Differences	Situations General Laws
	B=f(P,E) Effect of individual in an environment Multivariate Experimental Psychology	

Types of Relationships

(Vale and Vale, 1969)

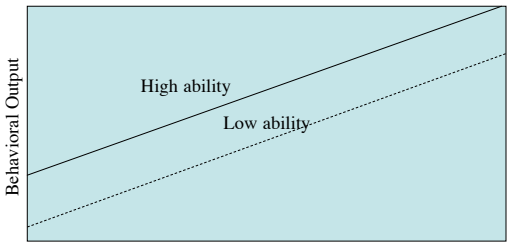
- Behavior = f(Situation)
- Behavior = f_1(Situation)+ f_2(Personality)
- Behavior = f_1(Situation)+f_2(Personality)+ f_3(Situation*Personality)
- Behavior = f_1(Situation * Personality)
- Behavior = idiosyncratic

Types of Relationships: Behavior = f(Situation)



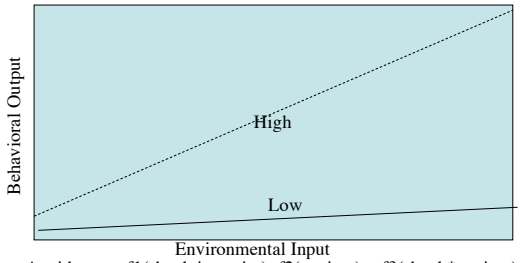
Neuronal excitation = f(light intensity)

Types of Relationships: Behavior = f_1(Situation)+f_2(Person)



Probability of college = f_1(income) + f_2(ability)

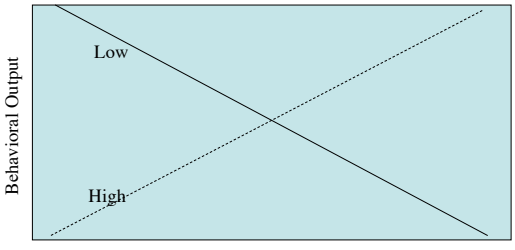
Types of Relationships: Behavior = f1(Situation)+f2(Personality)+ f3(Situation*Personality)



Avoidance = f1(shock intensity)+f2(anxiety) + f3(shock*anxiety)

Reading = f1(sesame street) = f2(ability) + f3(ss * ability)

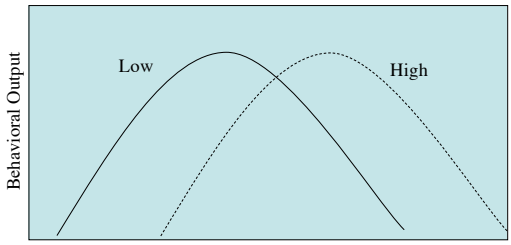
Types of Relationships: Behavior = f(Situation*Person)



Eating = f(preload * restraint)

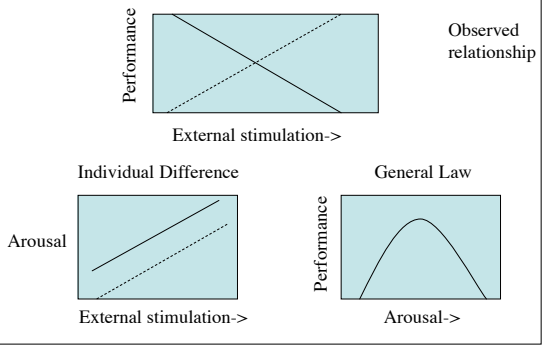
GRE = f(caffeine * impulsivity)

Types of Relationships: Behavior = f(Situation*Person)

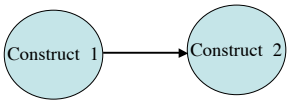


GRE = f(caffeine * impulsivity)

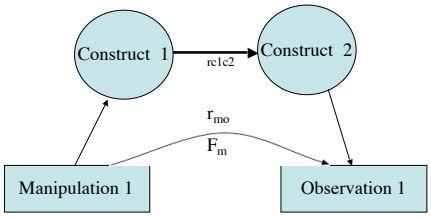
Persons, Situations, and Theory



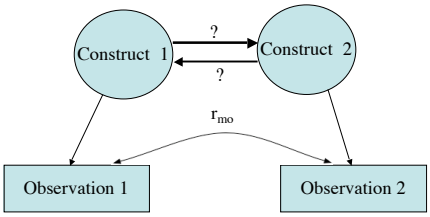
Theory and Theory Testing I: Theory



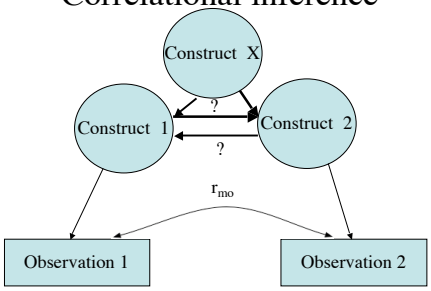
Theory and Theory Testing II: Experimental manipulation



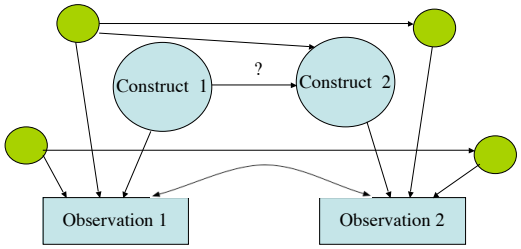
Theory and Theory Testing III: Correlational inference



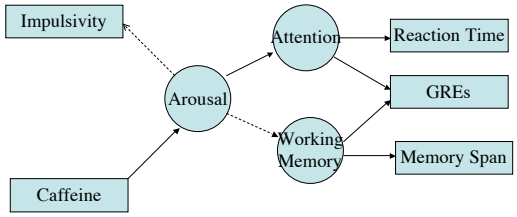
Theory and Theory Testing IV: Correlational inference



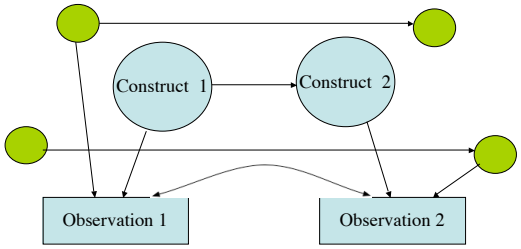
Theory and Theory Testing V: Alternative Explanations



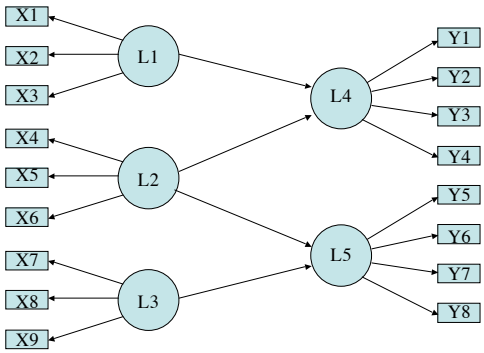
Individual differences and general laws



Theory and Theory Testing VI: Eliminate Alternative Explanations



Psychometric Theory: A conceptual Syllabus



A Theory of Data: What can be measured

X1

What is measured?

Individuals

Objects

What kind of measures are taken?

proximity

order

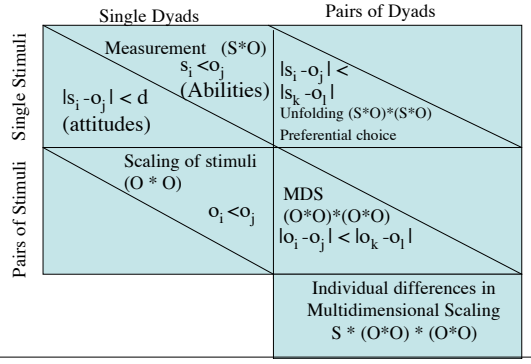
Comparisons are made on:

Single Dyads or Pairs of Dyads

Coombs: A theory of Data

- O = {Stimulus Objects} S={Subjects}
- $O = \{o_1, o_2, \ldots, o_i, \ldots, o_n\}$
- $S = \{s_1, s_2, \ldots, s_i, \ldots, s_m\}$
- $S \times O = \{(s_1, o_1), (s_i, o_j), \ldots, (s_m, o_n)\}$
- $O \times O = \{(o_1, o_1), (o_i, o_j), \ldots, (o_n, o_n)\}$
- Types of Comparisons:
  - Order $s_i < o_j$ (aptitudes or amounts)
  - Proximities $|s_i - o_j| < d$ (preferences)

Coombs typology of data



Metric spaces and the axioms of a distance measure

- A metric space is a set of points with a distance function, D, which meets the following properties
- Distance is symmetric, positive definite, and satisfies the triangle inequality:
  - $D(X,Y) >= 0$ (non negativity)
  - $D(X,Y) = 0$ iff $X=Y$ ($D(X,X)=0$ reflexive)
  - $D(X,Y) = D(Y,X)$ (symmetric)
  - $D(X,Y) + D(Y,Z) => D(X,Z)$ (triangle inequality)

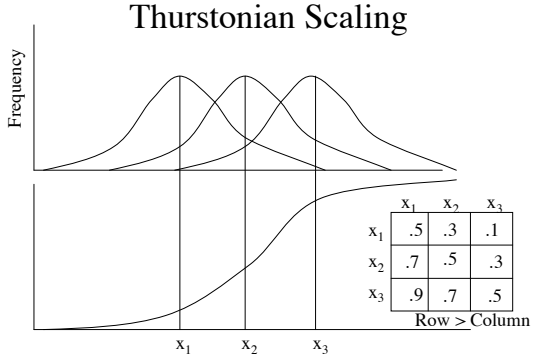
Scaling of Stimuli (O*O)

- Finding a distance metric for a set of stimuli
  - Sports teams (wins and losses)
  - Severity of crimes (judgments of severity)
  - Quality of merchandise (judgments)
  - Political orientations of judges (history of decisions -- voting with or against majority)

Thurstonian Scaling of Stimuli

- What is scale location of objects I and J on an attribute dimension D?
- Assume that object I has mean value m_i with some variability.
- Assume that object J has a mean value m_j
- Assume equal and normal variability (Thurston case 5)
  - Less restrictive assumptions are cases 1-4)
- Observe frequency of ($o_i < o_j$)
- Convert relative frequencies to normal equivalents
- Result is an interval scale with arbitrary 0 point

Thurstonian Scaling



	x_1	x_2	x_3
x_1	.5	.3	.1
x_2	.7	.5	.3
x_3	.9	.7	.5

Row > Column

Preferential Choice and Unfolding (S * O) * (S*O)

Comparison of the distance of subject to an item versus another subject to another item:

$$|s_i - o_j| < |s_k - o_l|$$

Do you like broccoli more than I like spinach?

Or more typically: do you like broccoli more than you like spinach?

Preferential choice Unfolding (S*O)*(S*O)

Preferential Choice: I scales

- Question asked an individual:
  - Do you prefer object j to object k?
- Model of answer:
  - Something is preferred to something else if if it "closer" in the attribute space or on a particular attribute dimension
  - Individual has an "Ideal point" on the attribute.
  - Objects have locations along the same attribute
  - $|s_i - o_j| < |s_i - o_k|$
  - The I scale is the individual's rank ordering of preferences

Preferential Choice: J scales

- Individual preferences can give information about object to object distances that are true for multiple people
- Locate people in terms of their I scales along a common J scale.

Preferential Choice: free choice

- If you had complete freedom of choice, how many children would you like to have? _X_
- If you could not have that many, what would your second choice be? _Y_
- Third choice? _Z_
- Fourth choice? -W-
- Fifth choice? _V_

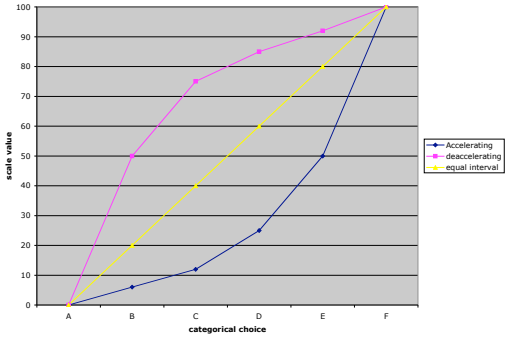
Preferential Choice: forced choice

1. If you had complete freedom of choice, how many children would you like to have? _X_
2. If you could not have X, would you rather have X+1 or X-1 (Y).
3. If could not have X or Y, would you rather have (min(X,Y)-1) or max (X,Y)+1. (Z)
4. If you could have X, Y or Z, would you rather have min(X,Y,Z)-1 or max (X,Y, Z)+1
5. Repeat (4) until either 0 or 5

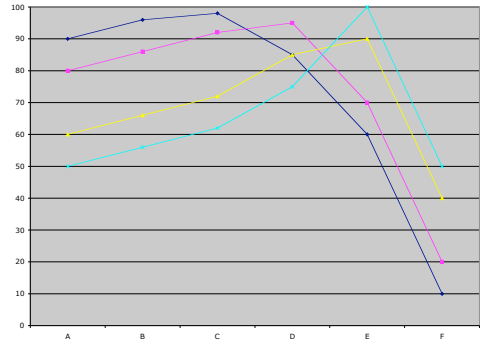
Preferential choice- underlying model

- On a scale from 0 to 100, if 0 means having 0 children, and 100 means having 5 children, please assign the relative location of 1, 2, 3, and 4 children.
- On this same scale, please give your preferences for having 0, 1, 2, 3, 4, or 5 children.

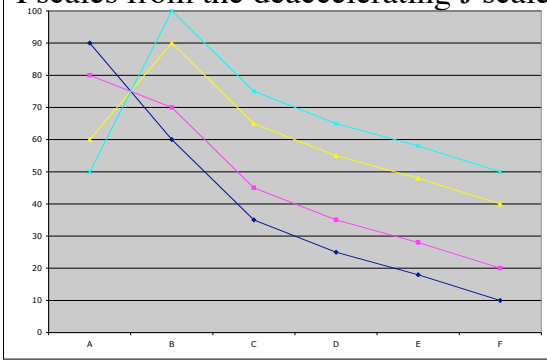
Alternative J scales



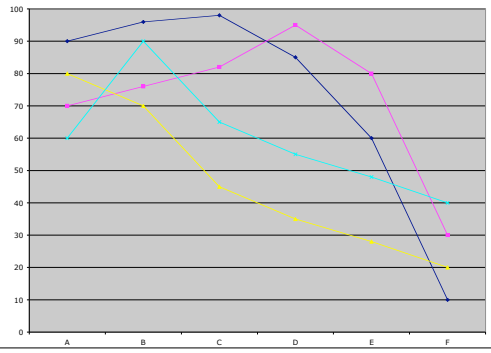
4 I scales from the accelerating J scale



I scales from the deaccelerating J scale



I scales from two J scales



Unfolding of Preferences

- Consider the I scale 234105
- What information has this person given us?
- Unfold to give J scale
- Ideal point is closest to 2, furthest from 5.
- J scale of
- 0 1 2 3 4 5
- Critical information: 2|3 occurs after 1|4

Joint scales, Points and Midpoints

- 0 1 2 3 4 5
- 0|1 0|2 0|3 0|4 0|5
- 1|2 1|3 **1|4** 1|5
- **2|3** 2|4 **2|5**
- **3|4** 3|5
- 4|5

I scales and midpoints example 1

- **Preference Orders: (Individual Scales)** **Midpoints crossed**

0 1 2 3 4	
1 0 2 3 4	01
1 2 0 3 4	01 02
1 2 3 0 4	01 02 03
1 2 3 4 0	01 02 03 04
2 1 3 4 0	01 02 03 04 12
2 3 1 4 0	01 02 03 04 12 13
2 3 4 1 0	01 02 03 04 12 13 14
3 2 4 1 0	01 02 03 04 12 13 14 23
3 4 2 1 0	01 02 03 04 12 13 14 23 24
4 3 2 1 0	01 02 03 04 12 13 14 23 24 34

I scales and Midpoints: Example 2

- **Preference Orders: (Individual Scales)** — **Midpoints crossed**

0 1 2 3 4										
1 0 2 3 4	01									
1 2 0 3 4	01	02								
2 1 0 3 4	01	02	12							
2 1 3 0 4	01	02	12	03						
2 3 1 0 4	01	02	12	03	13					
3 2 1 0 4	01	02	12	03	13	23				
3 2 1 4 0	01	02	12	03	13	23	04			
3 2 4 1 0	01	02	12	03	13	23	04	14		
3 4 2 1 0	01	02	12	03	13	23	04	14	24	
4 3 2 1 0	01	02	12	03	13	23	04	14	24	34

Distance information from midpoints

- Consider:
- 1 1|4 4
- 2 2|3 3 vs
- 2 2|3 3
- Midpoint orders imply distance information
- If 2|3<1|4 then (12) < (34)
- If 2|3 >1|4 then (12)> (34)

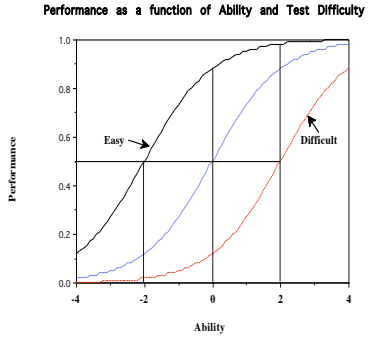
From midpoints to partial orders

- **Data example 1**
- 0|3 < 1|2 <=> (01) > (23)
- 0|4 < 1|2 <=> (01) > (24)
- 0|4 < 1|3 <=> (01) > (34)
- 0|4 < 2|3 <=> (02) > (34)
- 1|4 < 2|3 <=> (12) > (34)
  - **Partial Orders of distances**
- (04) > (03) > (02) > (12) > (34)
- (04) > (03) > (02) > (01) > (24) > (34)
- (04) > (03) > (02) > (01) > (24) > (23)

Measurement (S * O)

- Ordering of abilities: $s_i < o_j$
- Proximity of attitudes $|s_i - o_j| < d$

Measurement: Objects and Subjects



Latent and Observed Scores -- The problem of scale

Much of our research is concerned with making inferences about latent (unobservable) scores based upon observed measures. Typically, the relationship between observed and latent scores is monotonic, but not necessarily (and probably rarely) linear. This leads to many problems of inference. The following examples are abstracted from real studies. The names have been changed to protect the guilty.

Quality of school affects writing

- A leading research team in motivational and educational psychology was interested in the effect that different teaching techniques at various colleges and universities have upon their students. They were particularly interested in the effect upon writing performance of attending a very selective university, a less selective university, or a two year junior college. A writing test was given to the entering students at three institutions in the Boston area. After one year, a similar writing test was given again. Although there was some attrition from each sample, the researchers report data only for those who finished one year. The pre and post test scores as well as the change scores were:

College and Writing

	Pretest	Posttest	Change
Junior College	1	5	4
Non-selective university	5	27	22
Selective university	27	73	45

From these data, the researchers concluded that the quality of teaching at the very selective university was much better and that the students there learned a great deal more. They proposed to study the techniques used there in order to apply them to the other institutions.

Are their conclusions justified? Can you think of several reasons that their conclusions could be incorrect?

School Quality and Mathematics

- Another research team in motivational and educational psychology was interested in the effect that different teaching techniques at various colleges and universities have upon their students. They were particularly interested in the effect upon mathematics performance of attending a very selective university, a less selective university, or a two year junior college. A math test was given to the entering students at three institutions in the Boston area. After one year, a similar math test was given again. Although there was some attrition from each sample, the researchers report data only for those who finished one year. The pre and post test scores as well as the change scores were

College Quality and Mathematics

	Pretest	Posttest	Change
Junior College	27	73	45
Non-selective university	73	95	22
Selective university	95	99	4

- From these data, the researchers concluded that the quality of teaching at the very selective university was much worse and that the students there learned a great deal less than the other universities. They proposed to study the techniques used at these other institutions in order to apply them to the very selective university.
- Are their conclusions justified? Can you think of several reasons that their conclusions could be incorrect?