Psychology 405: Psychometric Theory

William Revelle
Northwestern University

Spring, 2013

http://personality-project.org/revelle/syllabi/405.syllabus.html
and
http://personality-project.org/revelle/syllabi/405.old.syllabus.html

Lord Kelvin's dictum

In physical science a first essential step in the direction of learning any subject is to find principles of numerical reckoning and methods for practicably measuring some quality connected with it. I often say that when you can measure what you are speaking about and express it in numbers you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science, whatever the matter may be. (Thomsom, 1891)

Taken from Michell (2003) in his critique of psychometrics:
Michell, J. The Quantitative Imperative: Positivism, Naïve Realism and the Place of Qualitative Methods in Psychology, Theory & Psychology, Vol. 13, No. 1, 5-31 (2003)

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)

Psychology and the need for measurement

- “The history of science is the history of measurement” (J. M. Cattell, 1893)
- “We hardly recognize a subject as scientific if measurement is not one of its tools” (Boring, 1929)
- “There is yet another [method] so vital that, if lacking it, any study is thought ... not be scientific in the full sense of the word. This further an crucial method is that of measurement.” (Spearman, 1937)
- “One’s knowledge of science begins when he can measure what he is speaking about and express in numbers” (Eysenck, 1973)

Michell, J. The Quantitative Imperative: Positivism, Naïve Realism and the Place of Qualitative Methods in Psychology, Theory & Psychology, Vol. 13, No. 1, 5-31 (2003)

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

1. To instill an appreciation of and an interest in the principles and methods of psychometric theory.
2. 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)
3. It will give you limited experience with psychometric computer programs (although all of the examples will use R, it not necessary to learn R).

Psychometric Theory: Requirements

- Asking questions!
- 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.(very highly recommended)
- Loehlin, John (2004) *Latent Variable Models: an introduction to factor, path, and structural analysis (4th edition*. Hillsdale, N.J.: LEA. (highly recommended)
- Revelle, W. *An introduction to psychometric theory with applications in R* (under development - see web)
  - http://personality-project.org/r/book
- web guide to class:
  - http://personality-project.org/revelle/syllabi/405.syllabus.html

Syllabus: Overview

I. Individual Differences and Experimental Psychology

II. Models of measurement

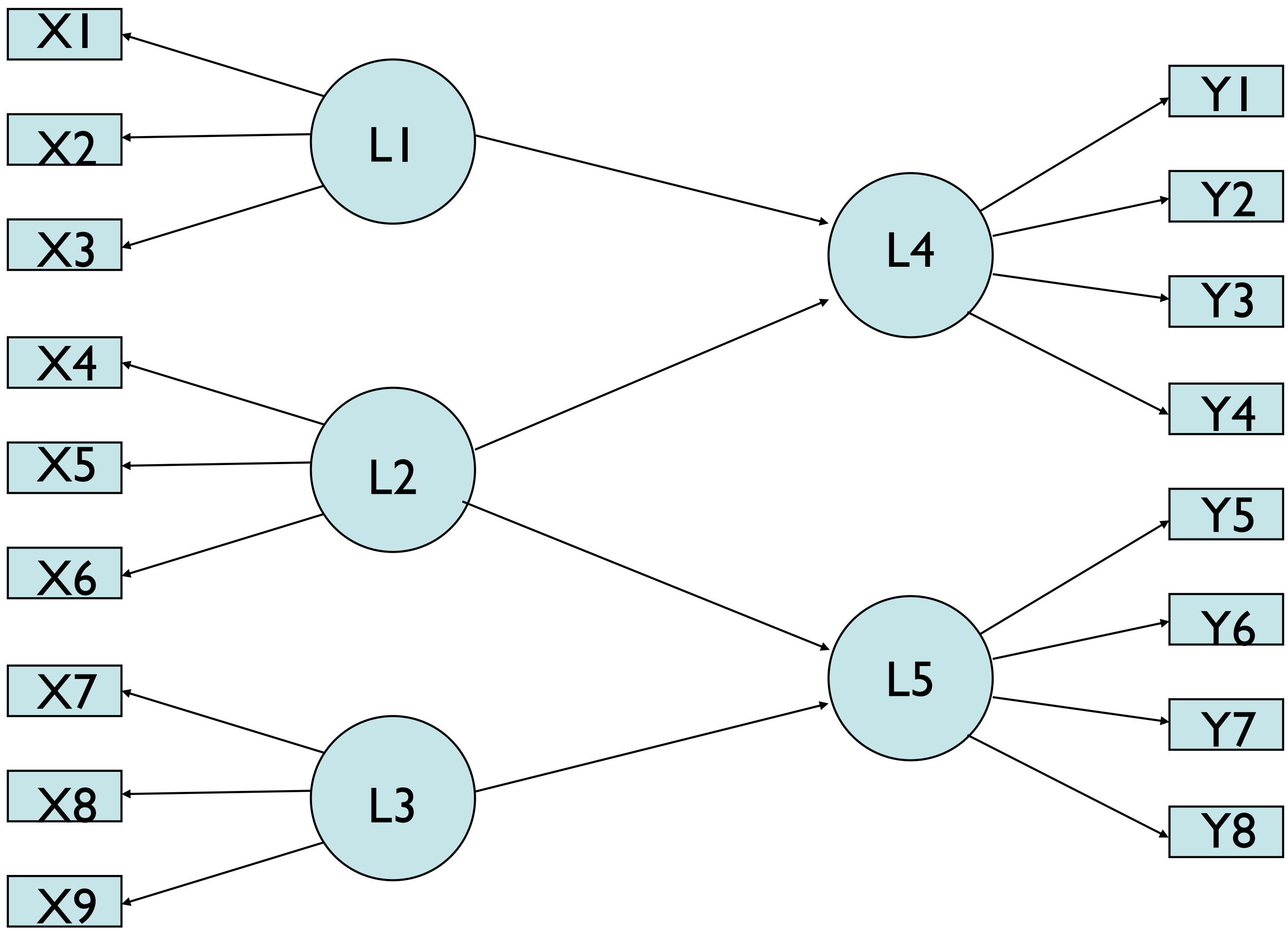
III. Test theory

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

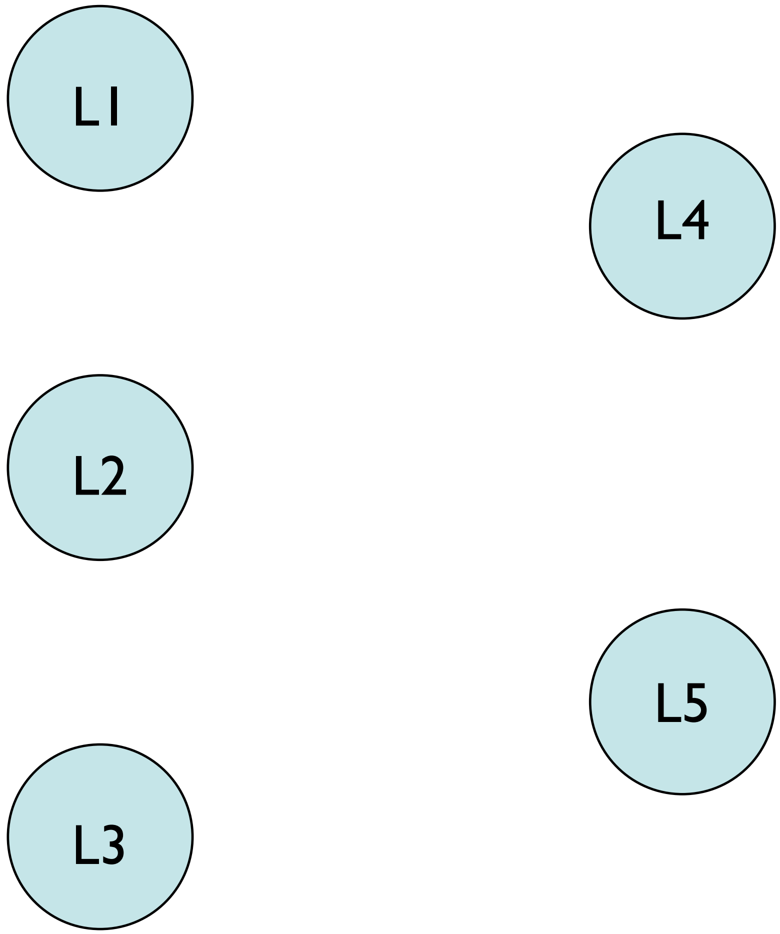
IV. Assessment of traits

V. Methods of observation of behavior

Psychometric Theory: A conceptual Syllabus



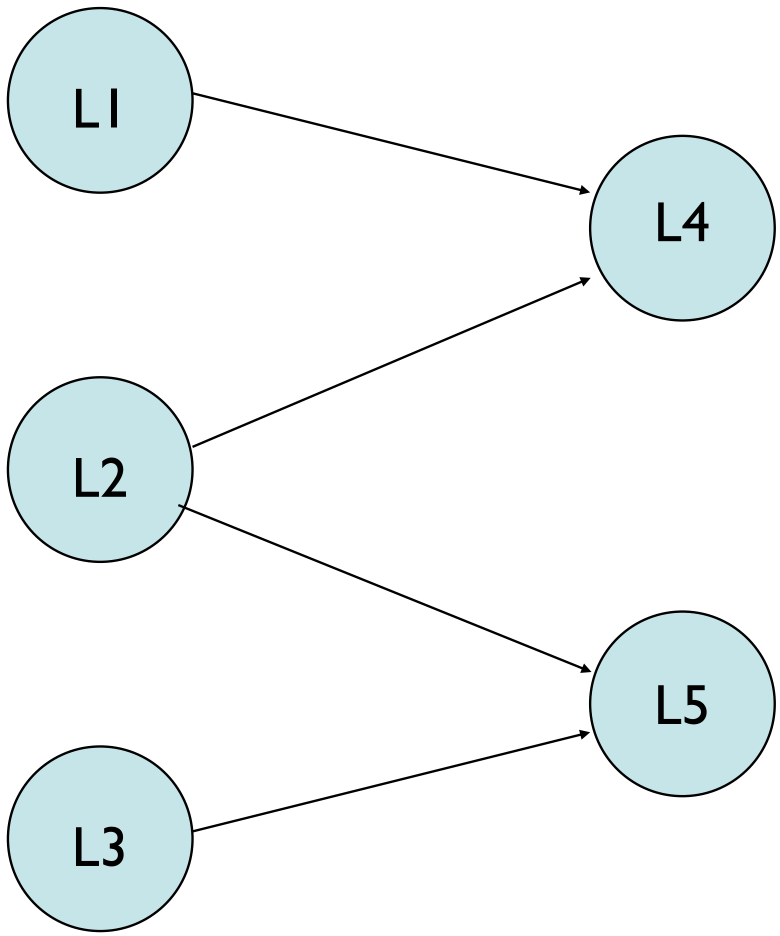
Constructs/Latent Variables



Examples of psychological constructs

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

Theory as organization of constructs



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

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

Observable or measured variables

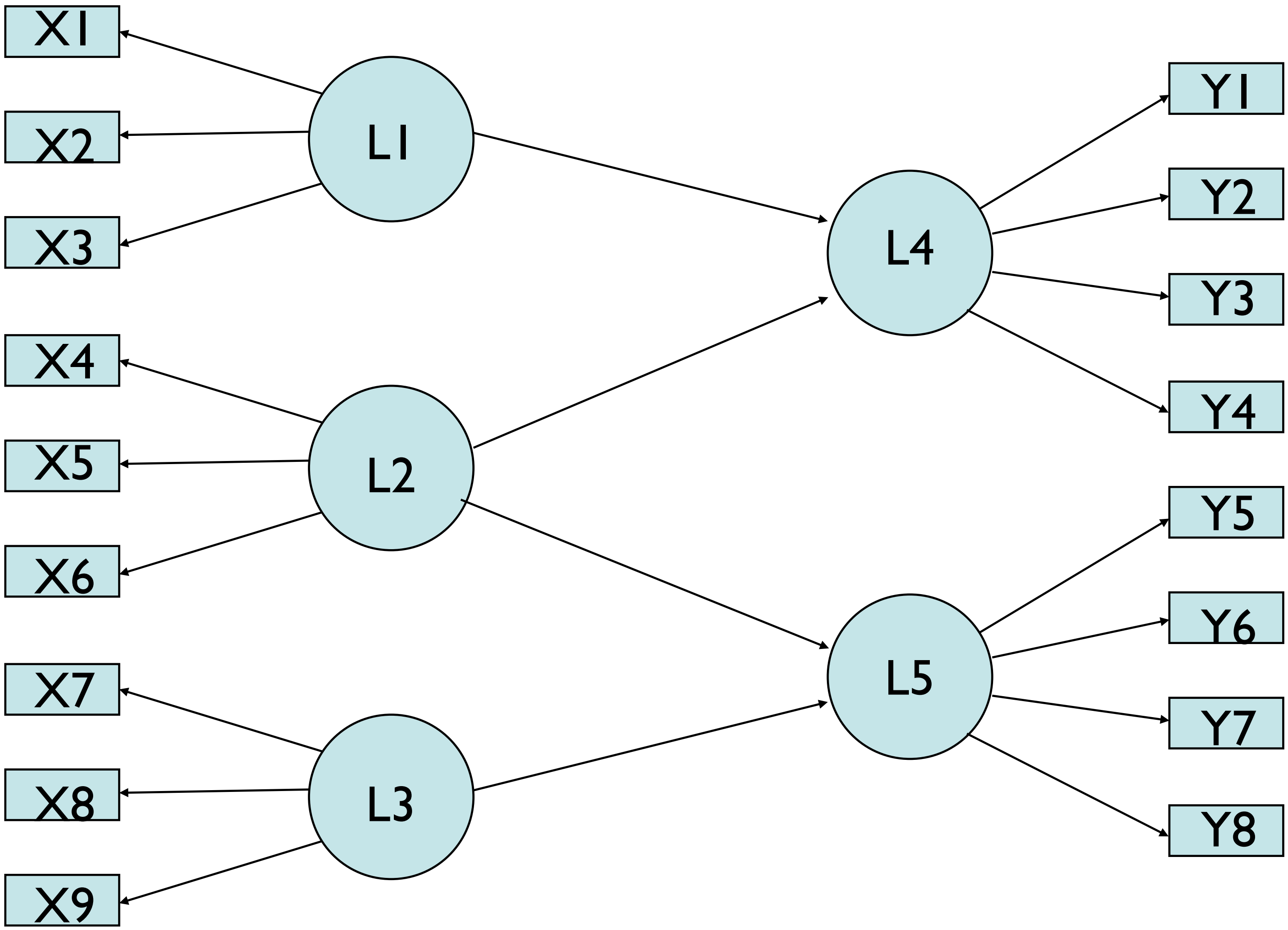
Observed Variables

- Item Endorsement
- Reaction time
- Choice/Preference
- Blood Oxygen Level Dependent Response
- Skin Conductance
- Archival measures

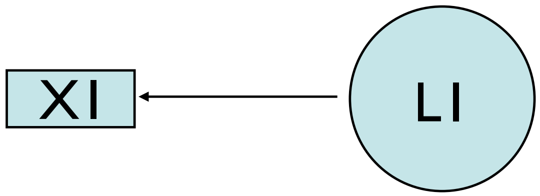
Theory development and testing

- Theories as organizations of observable variables
- Constructs, latent variables and observed variables
  - Observable variables
    - Multiple levels of description and abstraction
    - Multiple levels of inference about observed variables
  - 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

Psychometric Theory: A conceptual Syllabus



A Theory of Data: What can be measured

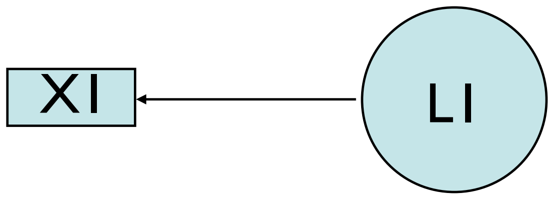


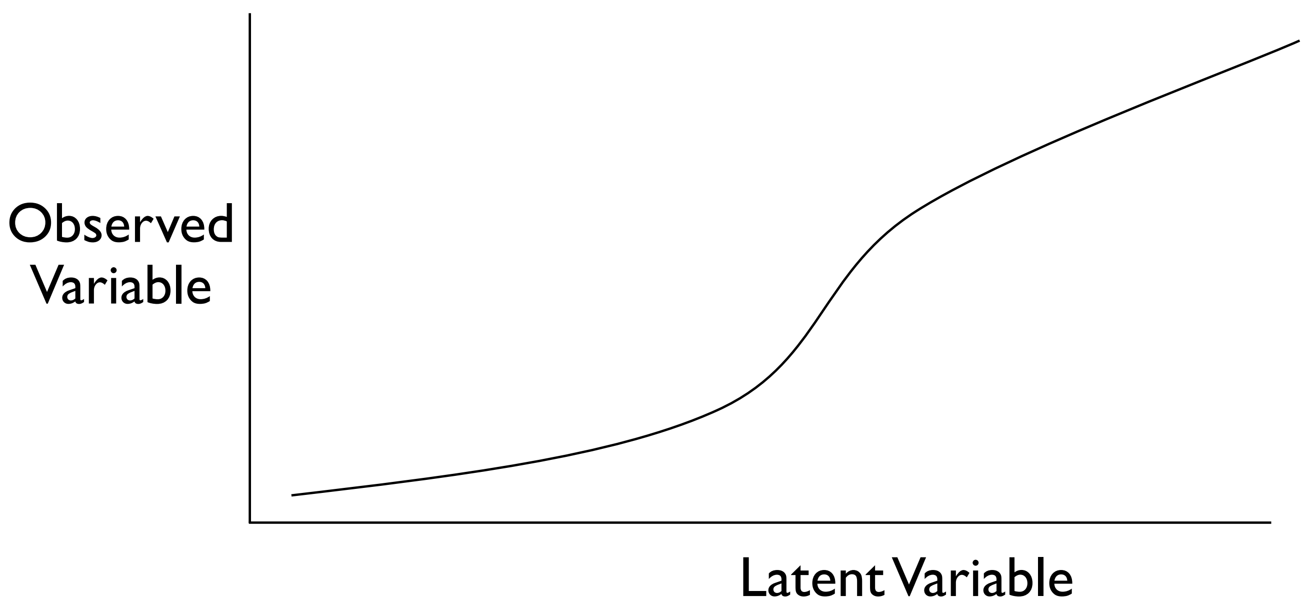
What is measured?
- Objects
- Individuals

What kind of measures are taken?
- Order
- Proximity

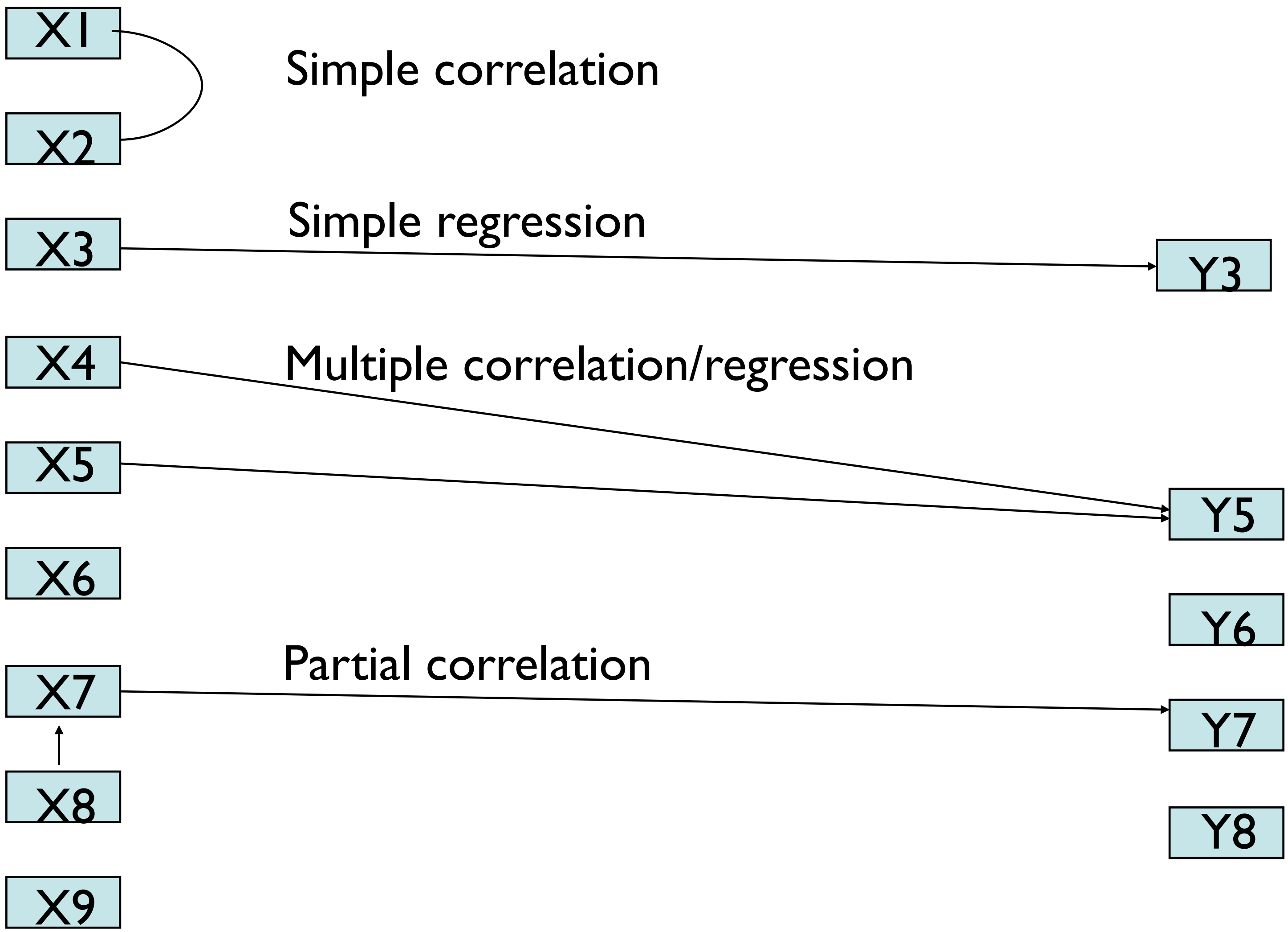
What kind of comparisons are made?
- Single Dyads
- Pairs of Dyads

Scaling:
the mapping between observed and latent variables

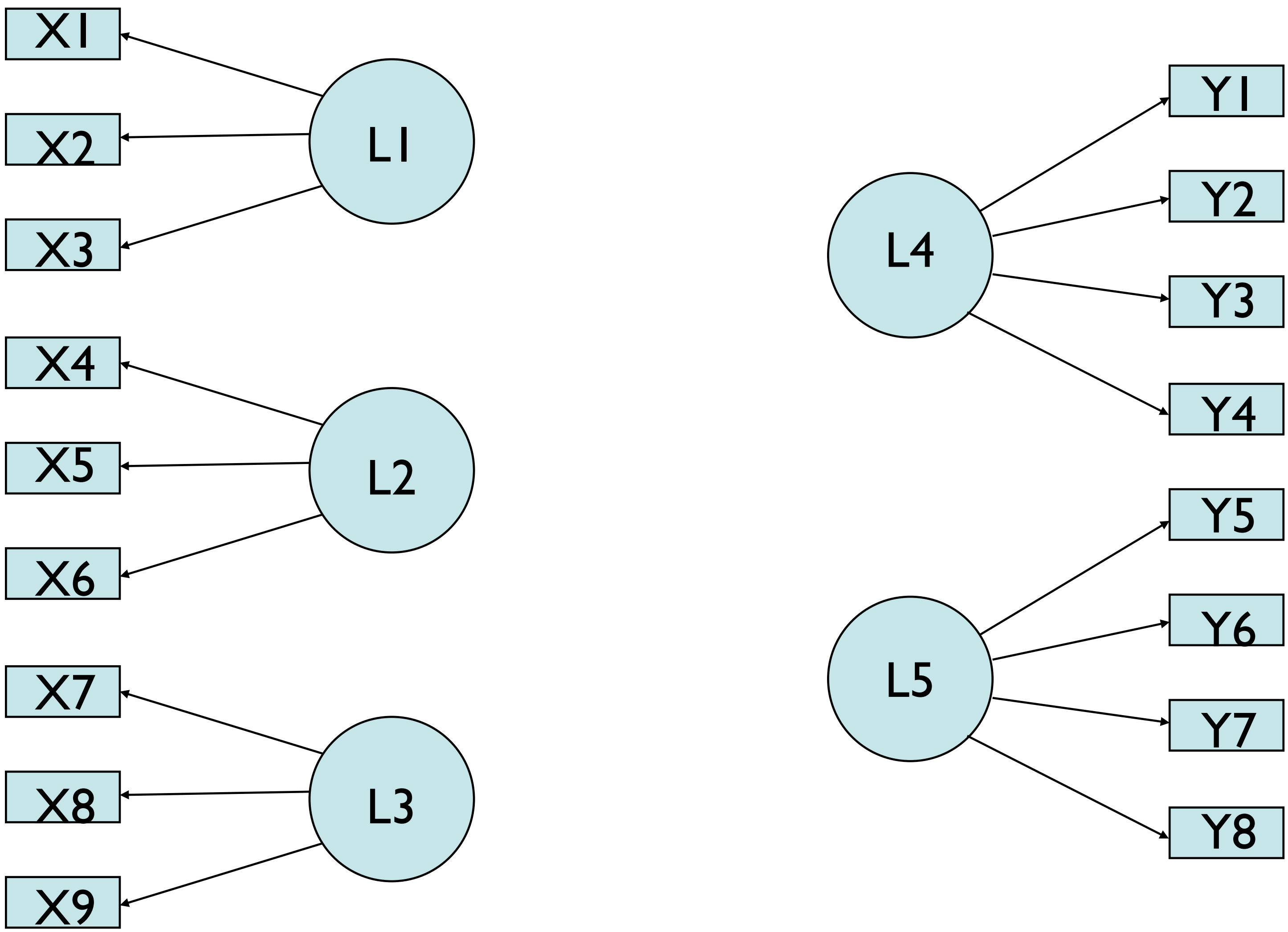




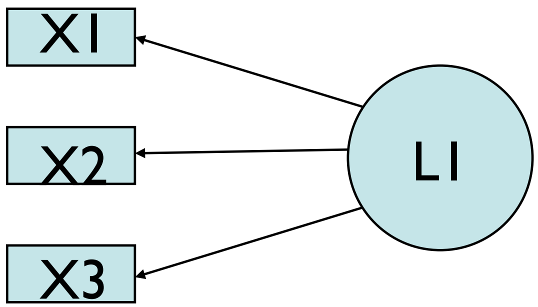
Variance, Covariance, and Correlation



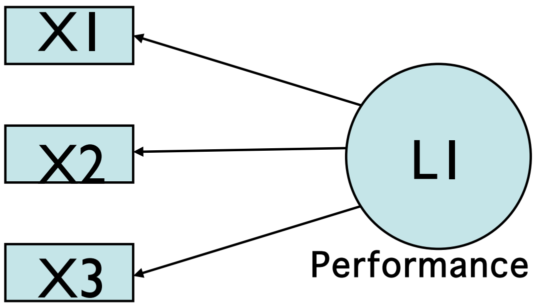
Techniques of Data Reduction: Factor and Components Analysis



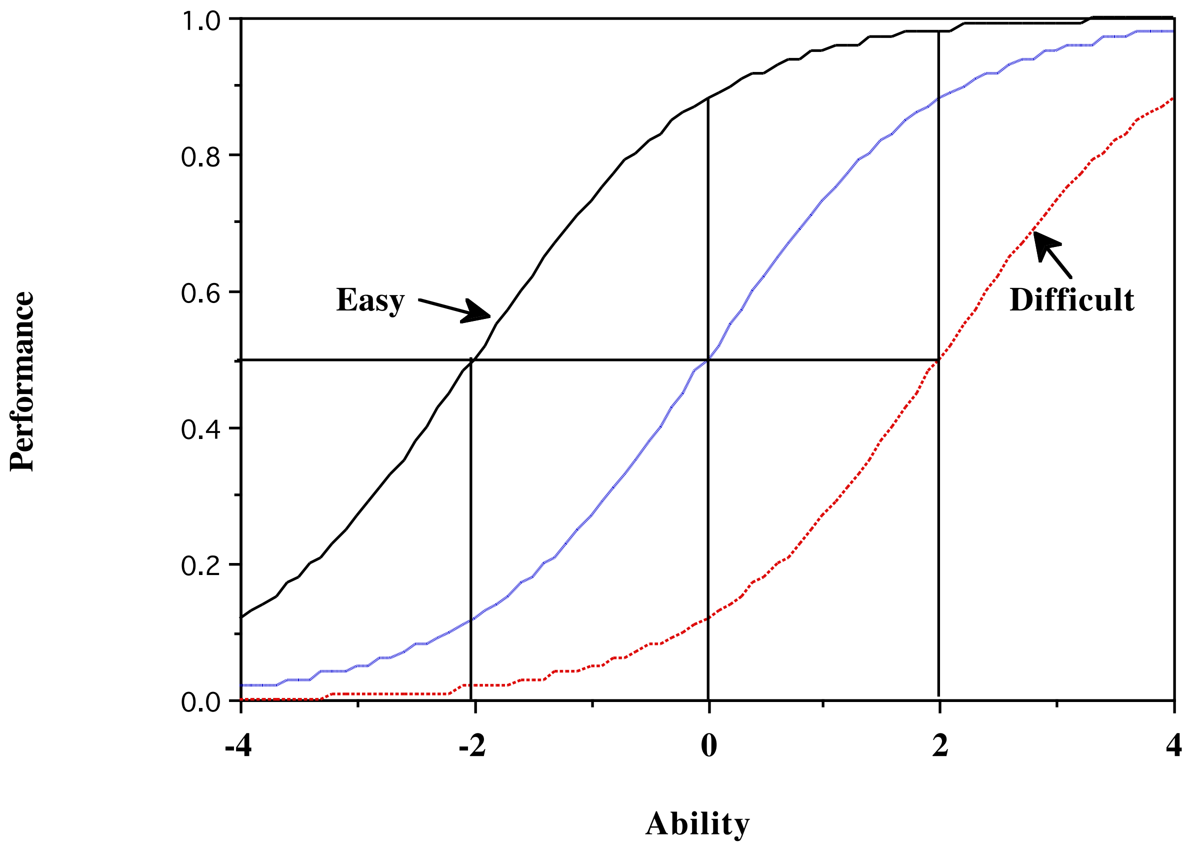
Classic Reliability Theory: How well do we measure what ever we are measuring



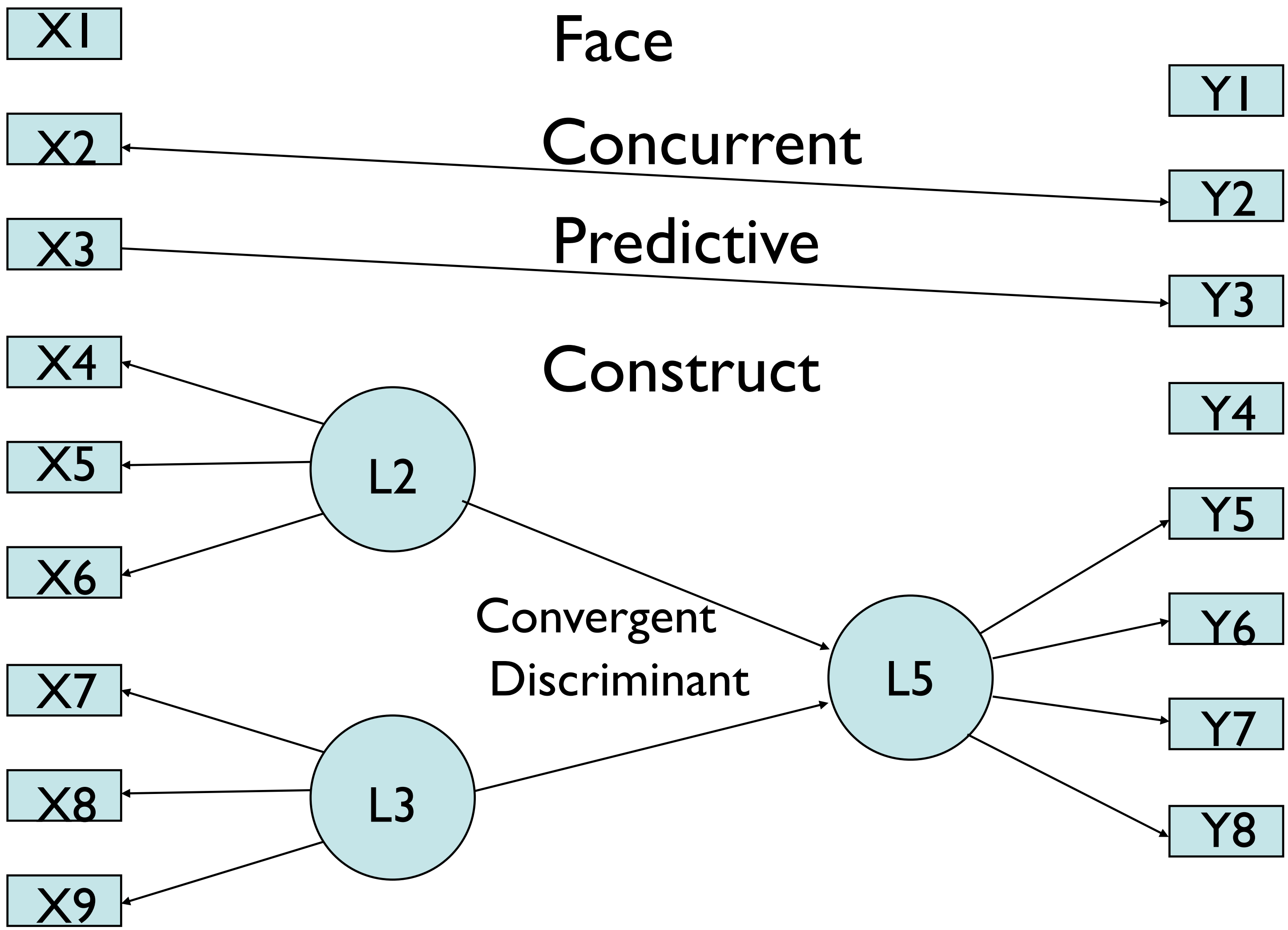
Modern Reliability Theory: Item Response Theory
How well do we measure what ever we are measuring



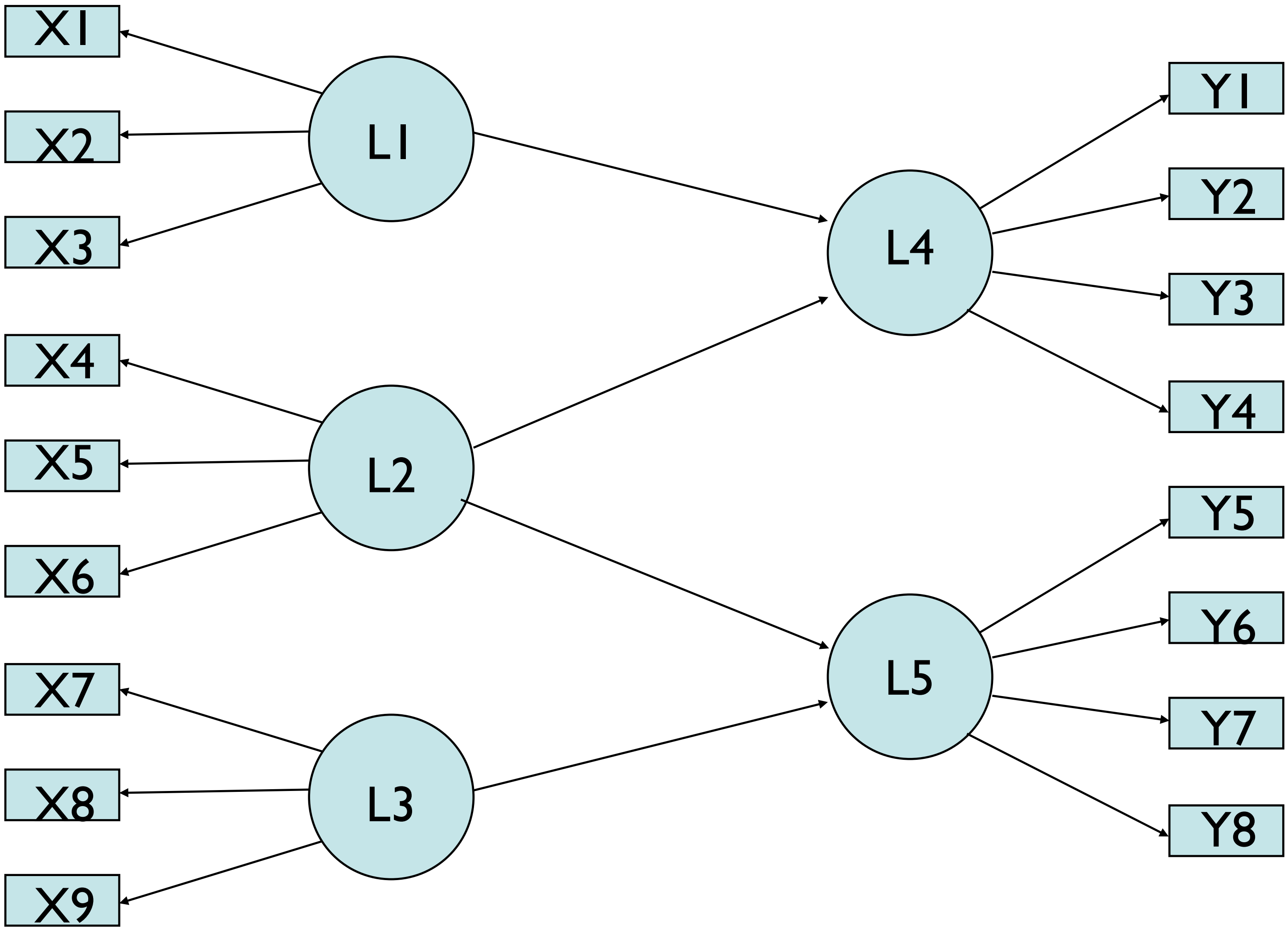
Performance as a function of Ability and Test Difficulty



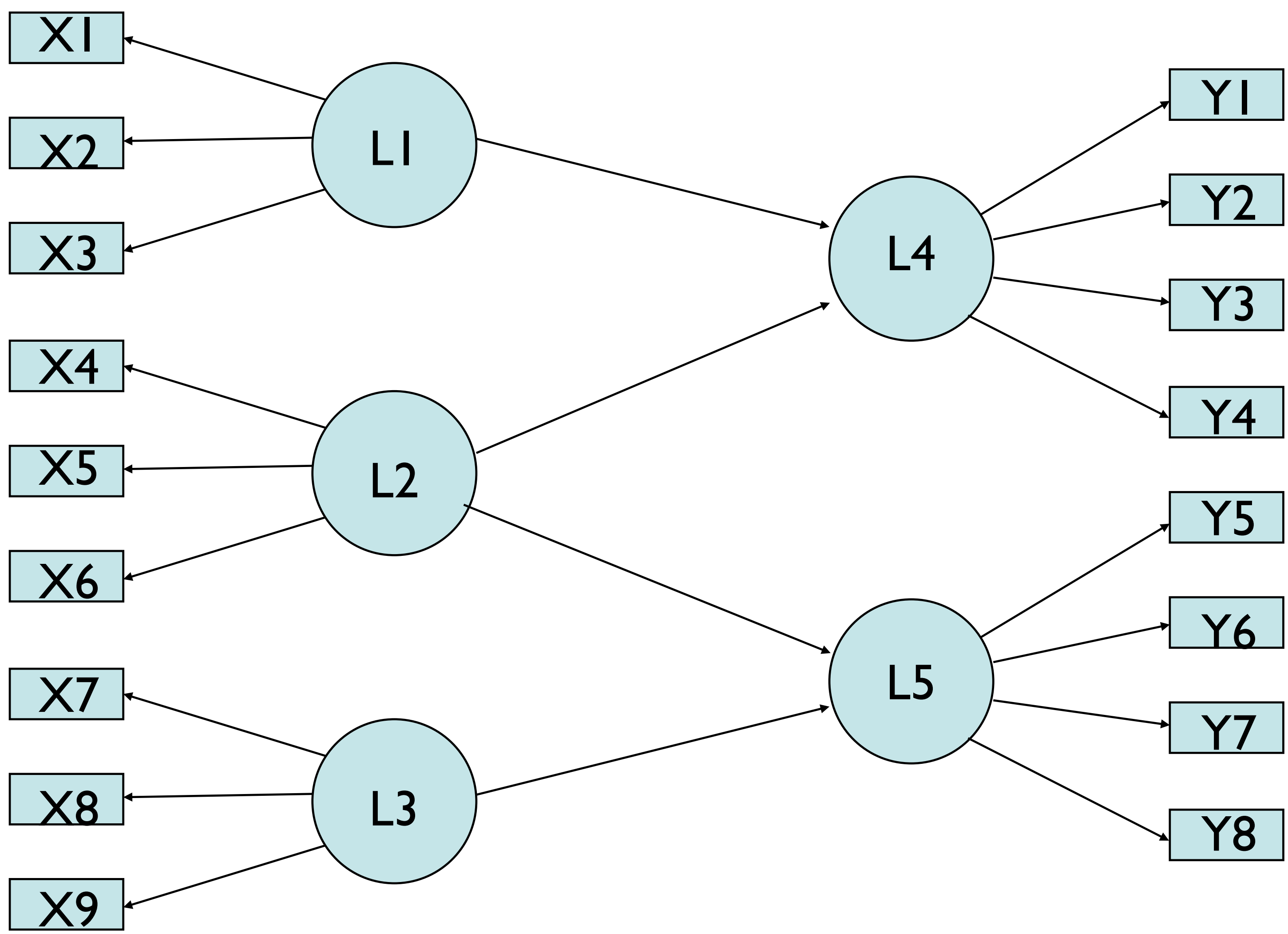
Types of Validity: What are we measuring



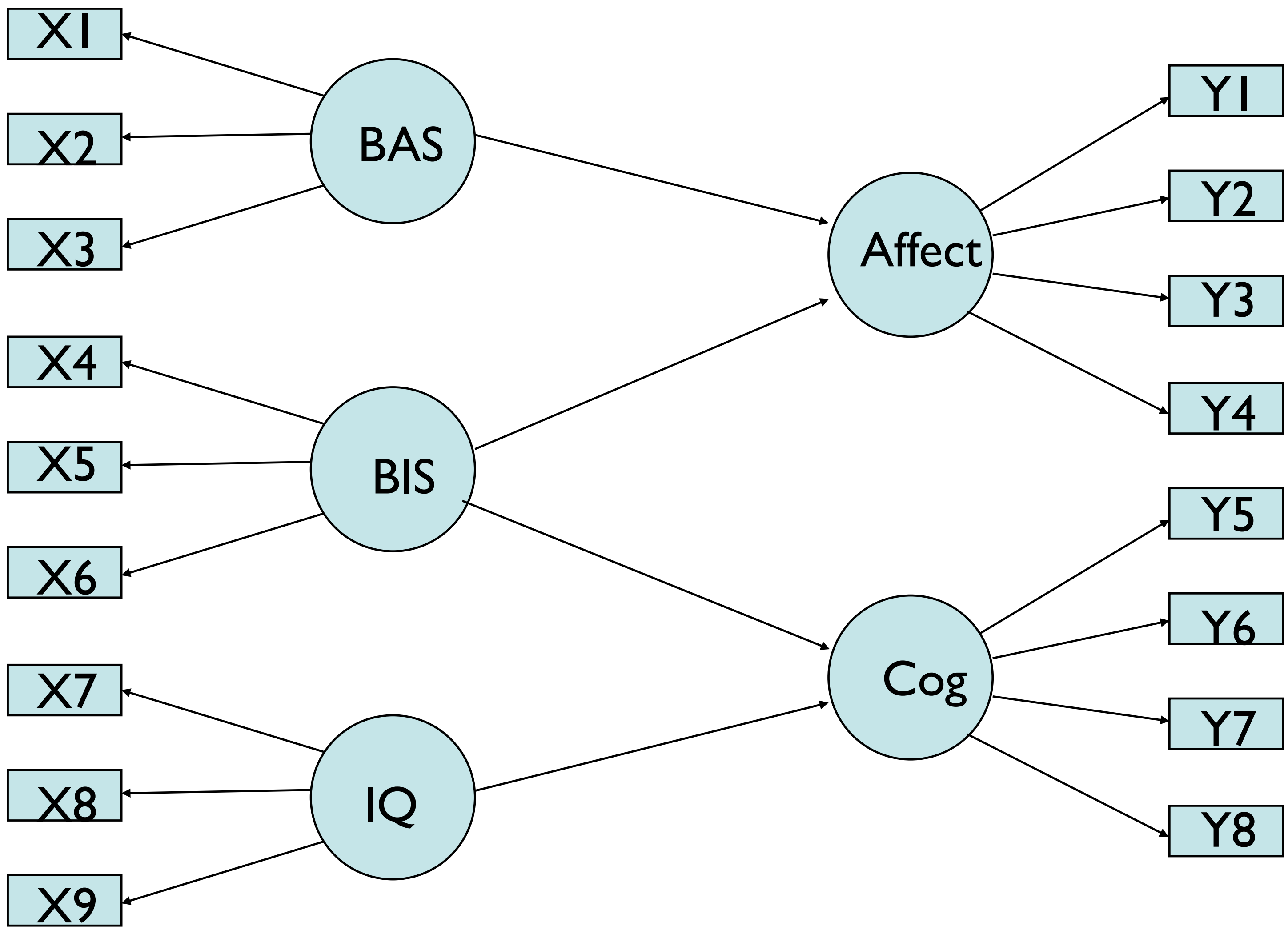
Structural Equation Modeling: Combining Measurement and Structural Models



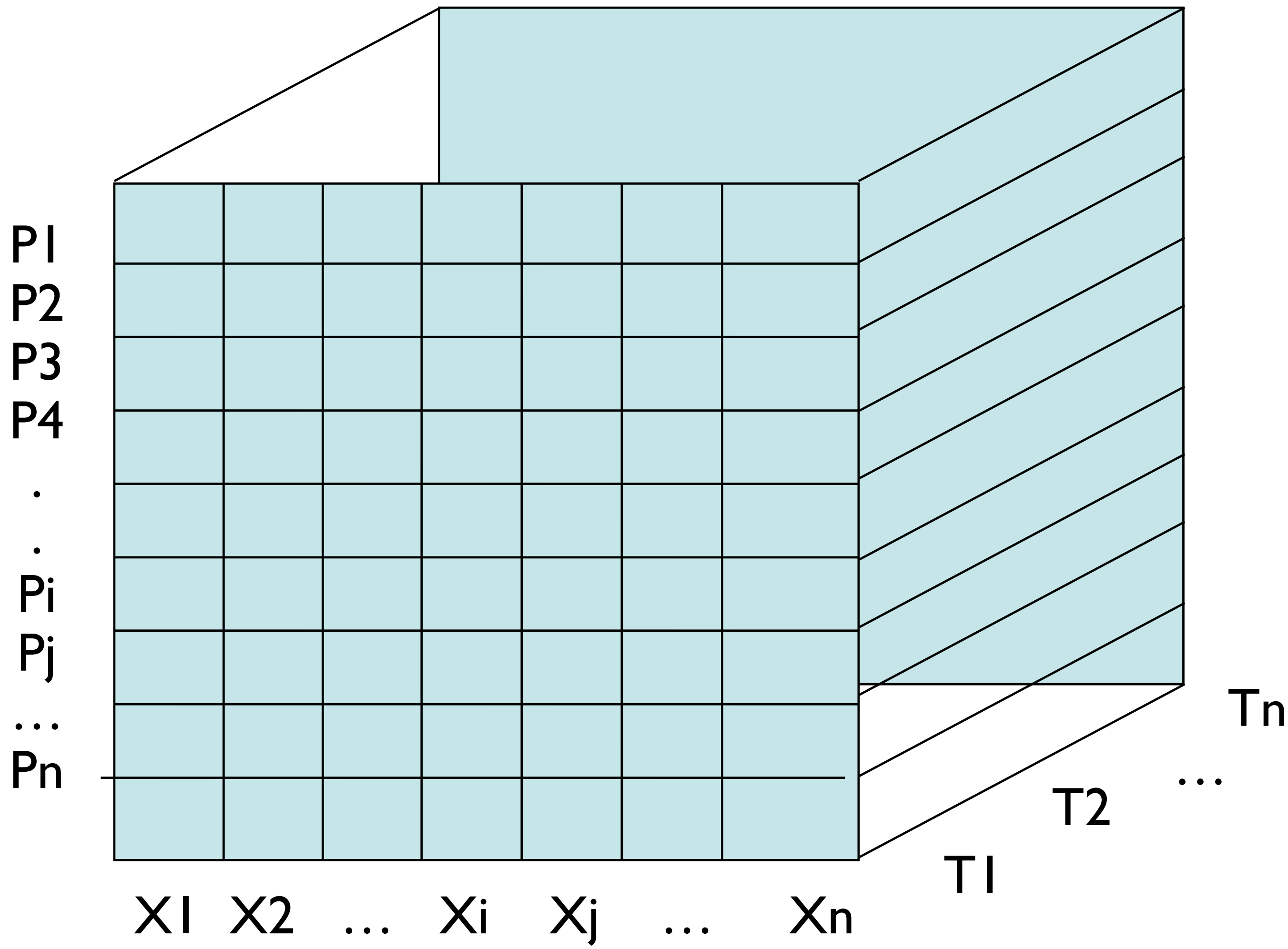
Scale Construction: practical and theoretical



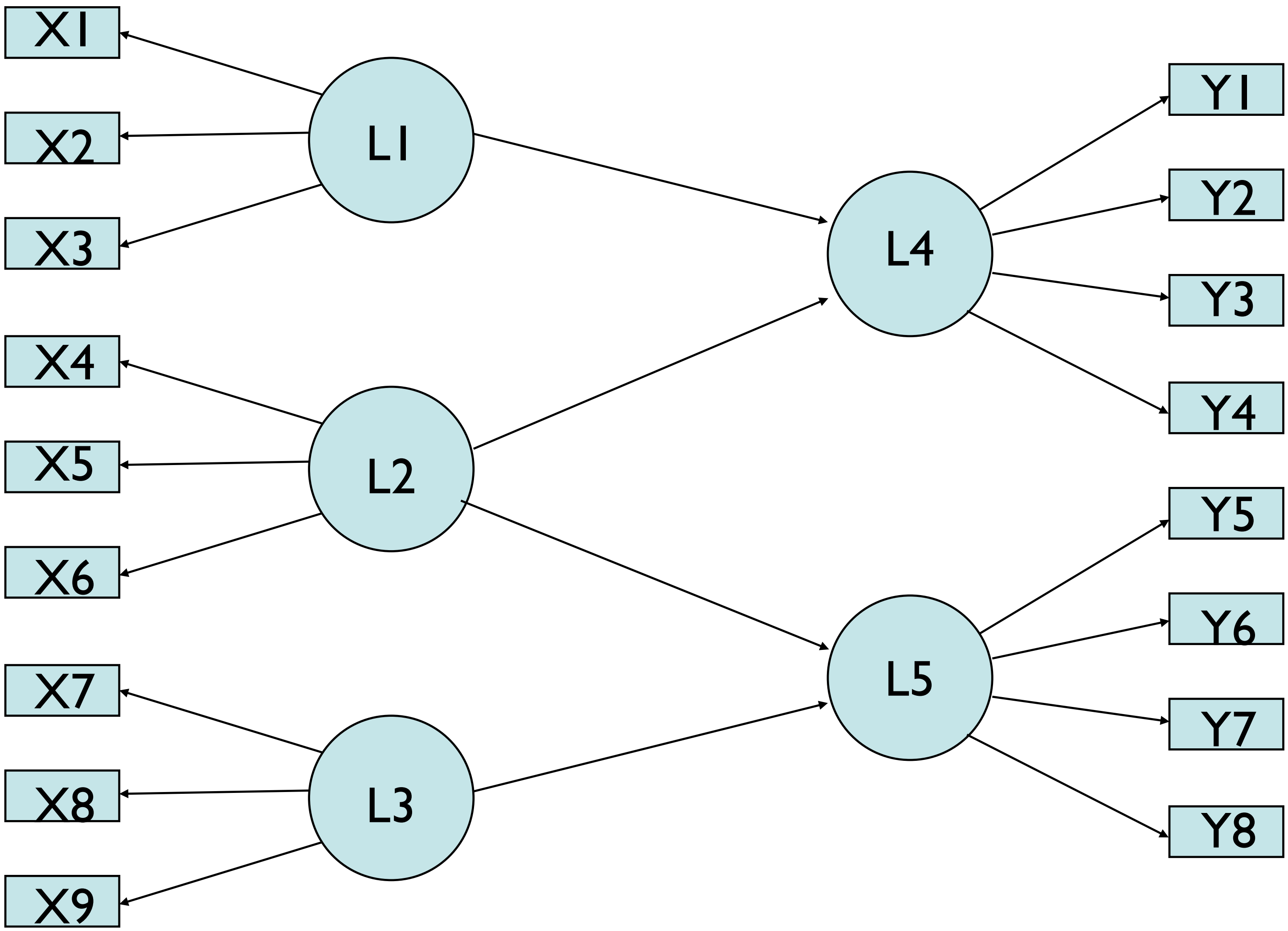
Traits and States: What is measured?



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
   D. Dimension reduction: Factor, Component and Cluster analysis
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

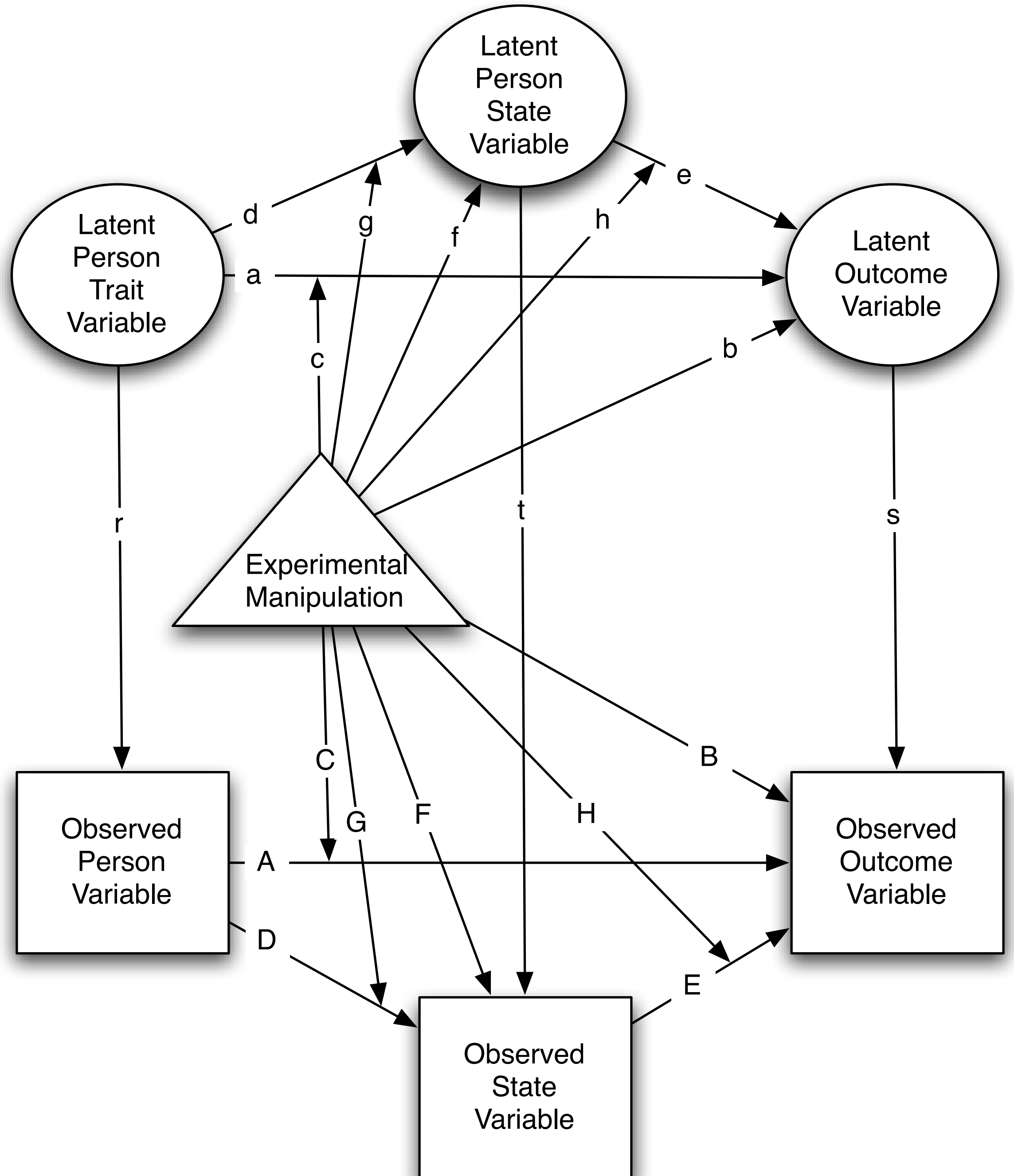
Cronbach, (1957, 1975); Eysenck (1966, 1997), Revelle & Oehlberg (2009)

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

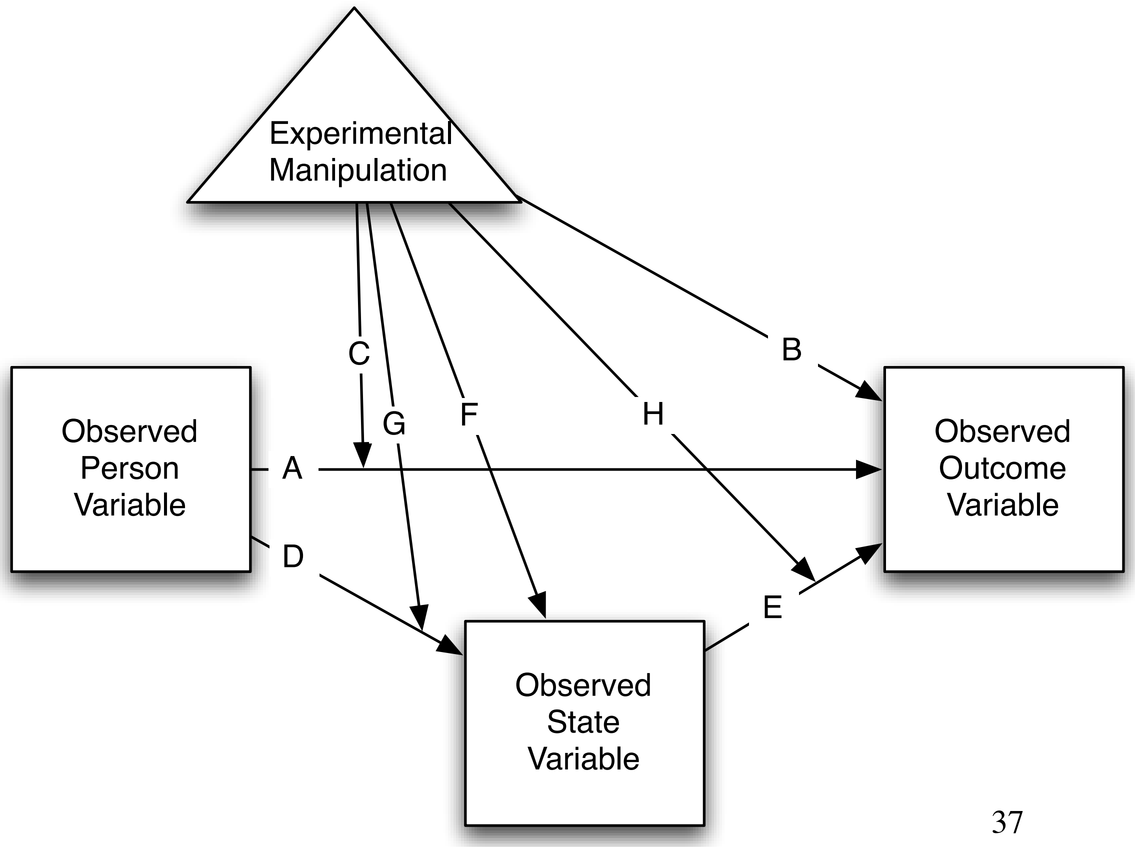
Two Disciplines of Psychological Research

	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	

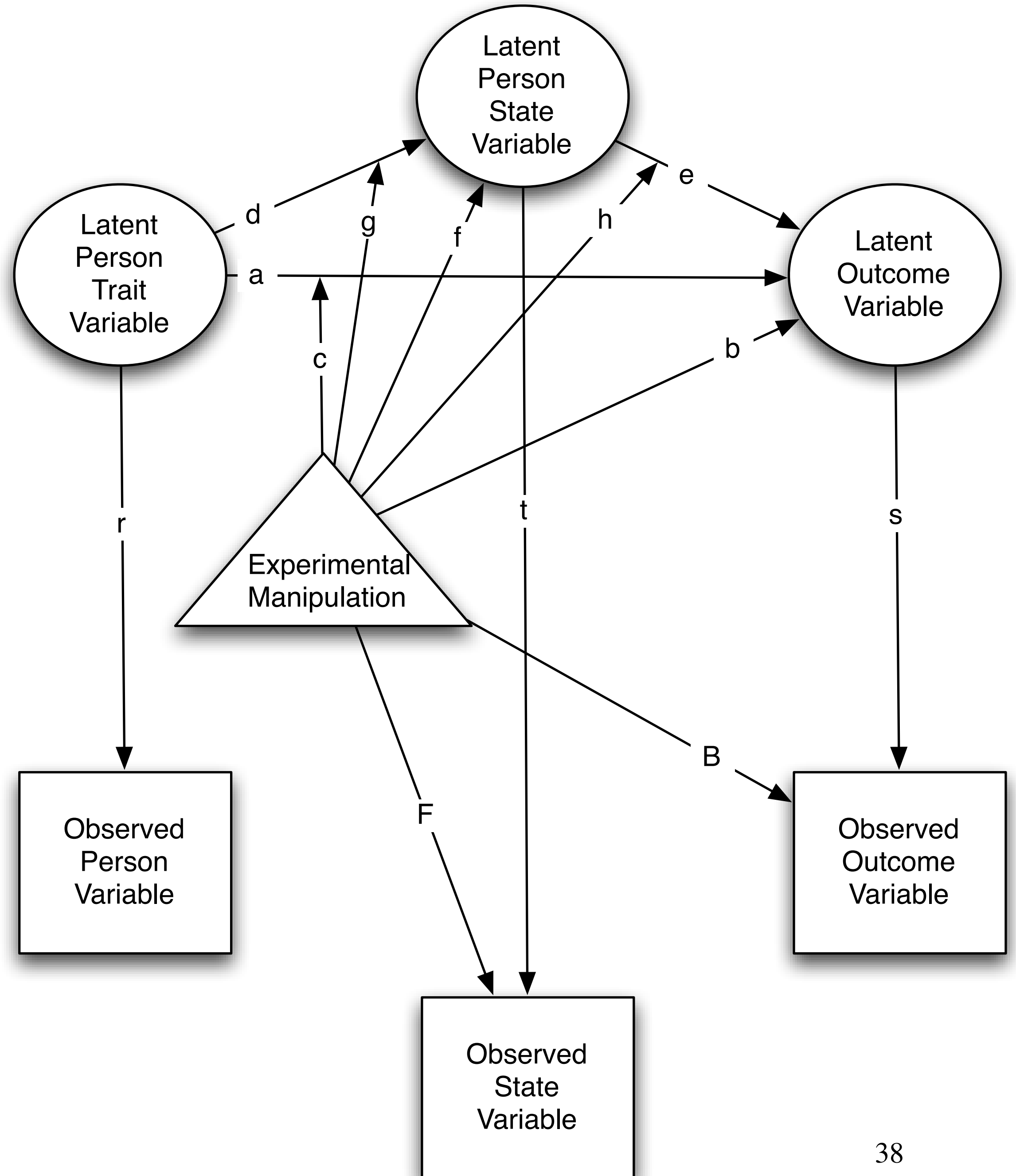
Experimental Personality Research involves theory, measurement and experimental technique



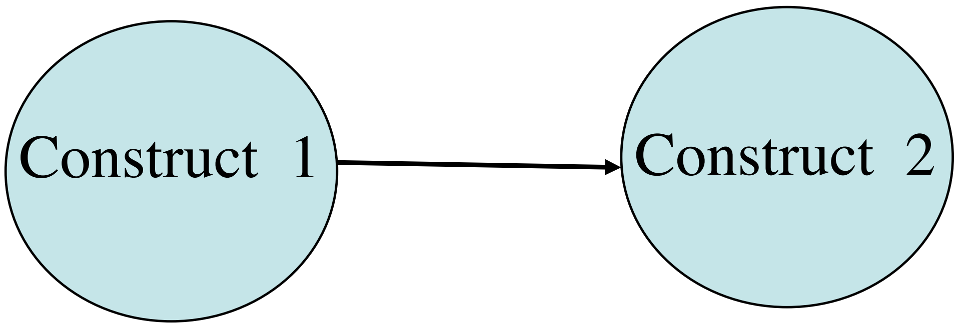
Experimental Personality Research involves theory, measurement and experimental technique



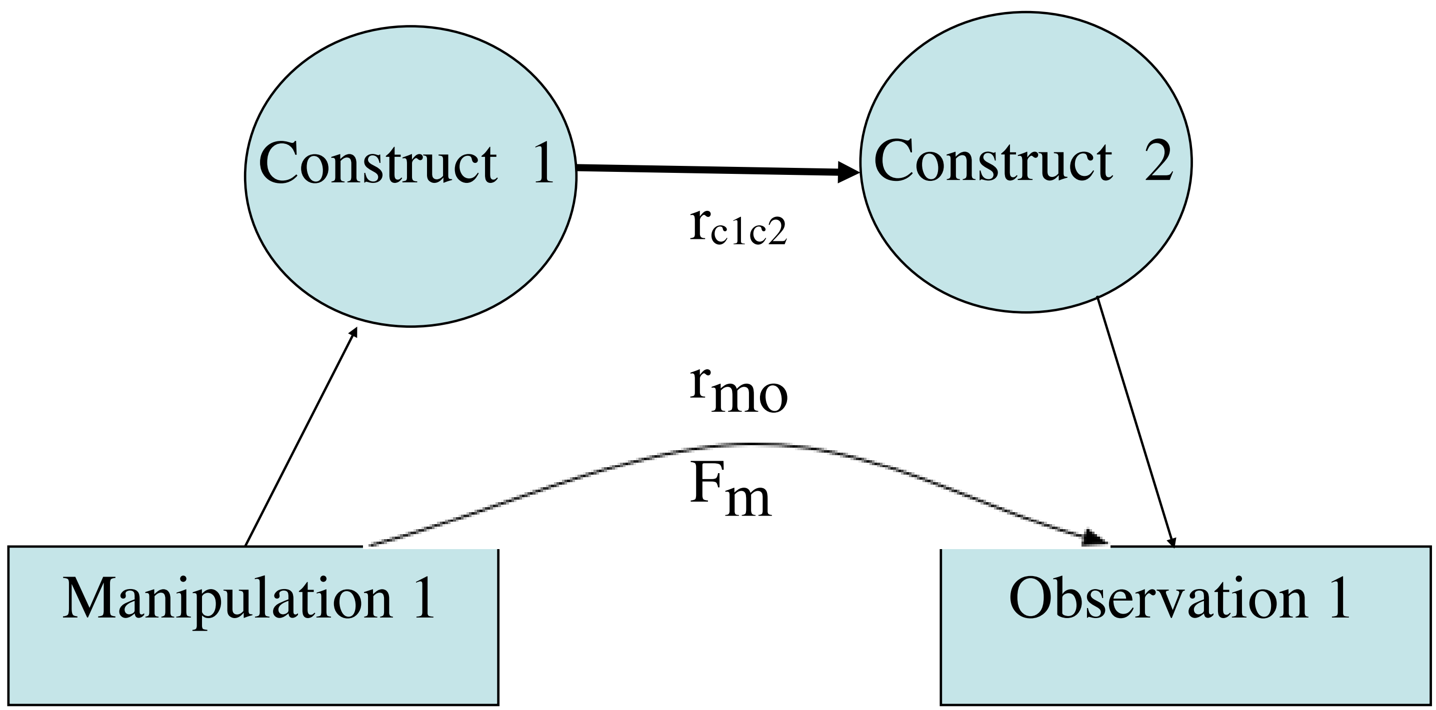
Experimental Personality Research involves theory, measurement and experimental technique



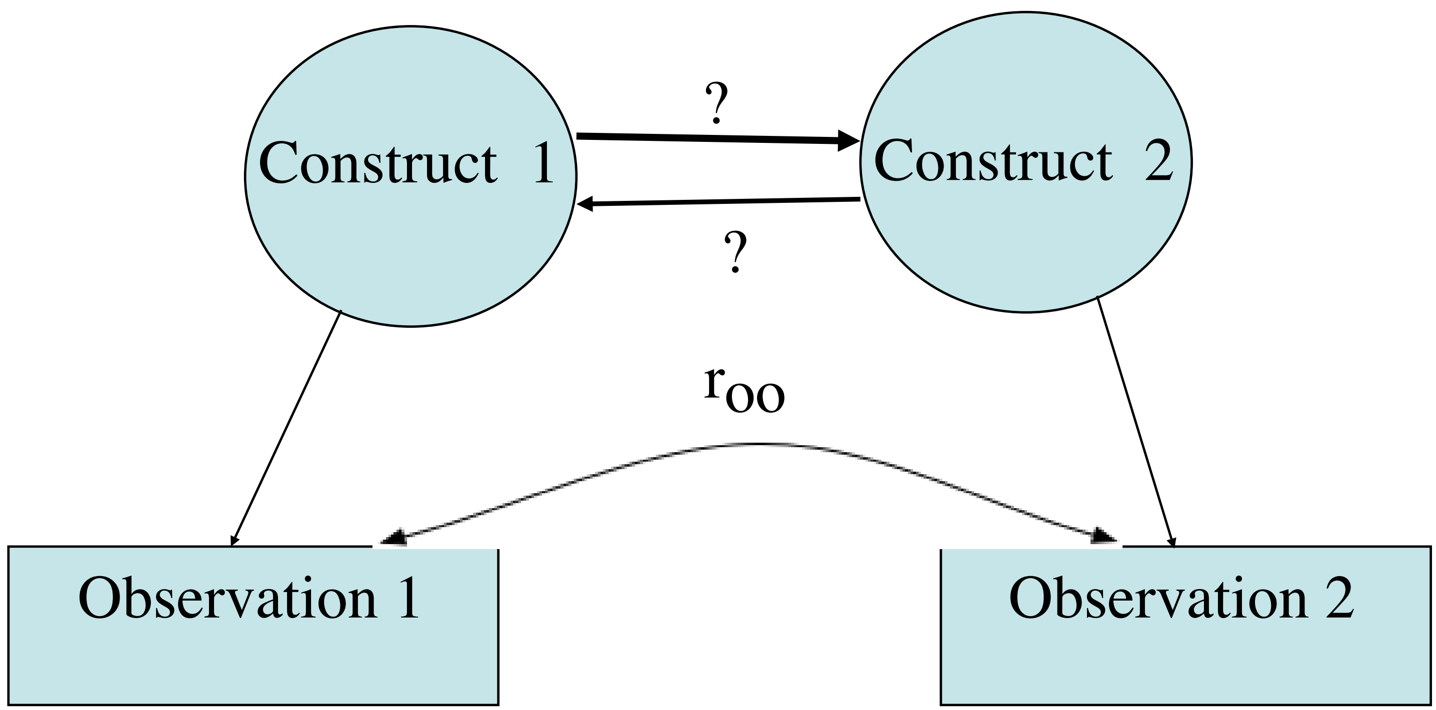
Theory and Theory Testing I: Theory



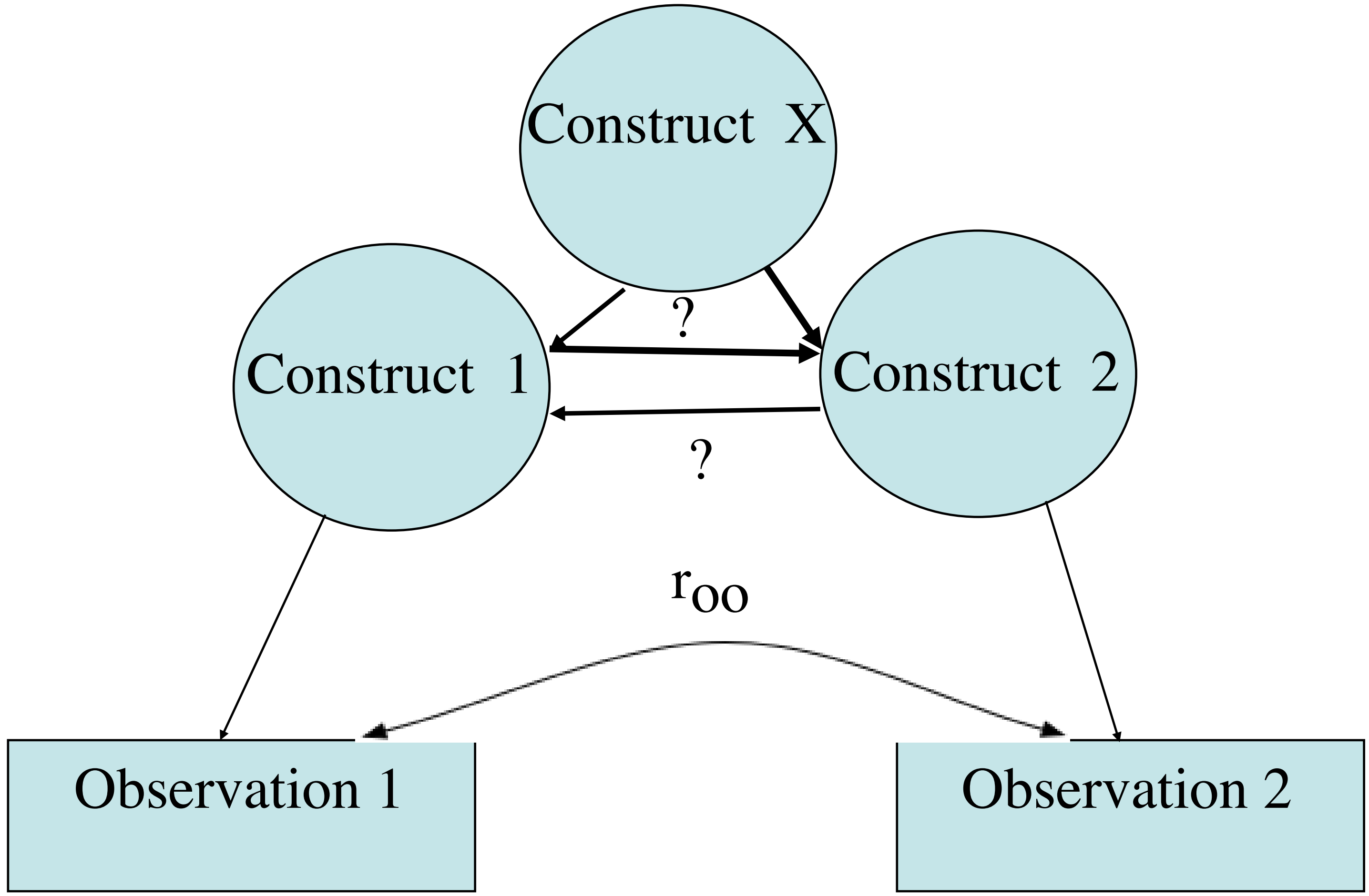
Theory and Theory Testing II: Experimental manipulation



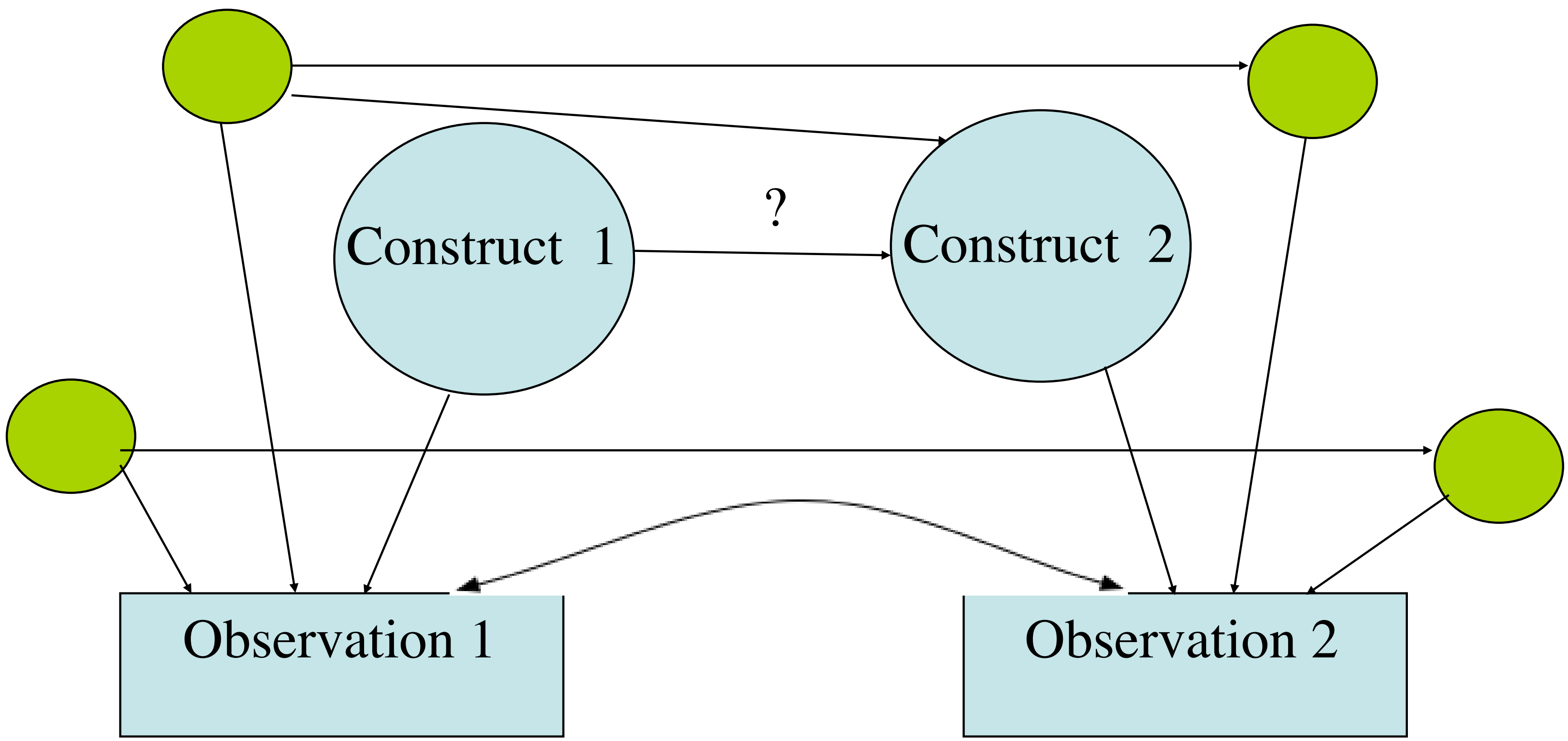
Theory and Theory Testing III: Correlational inference



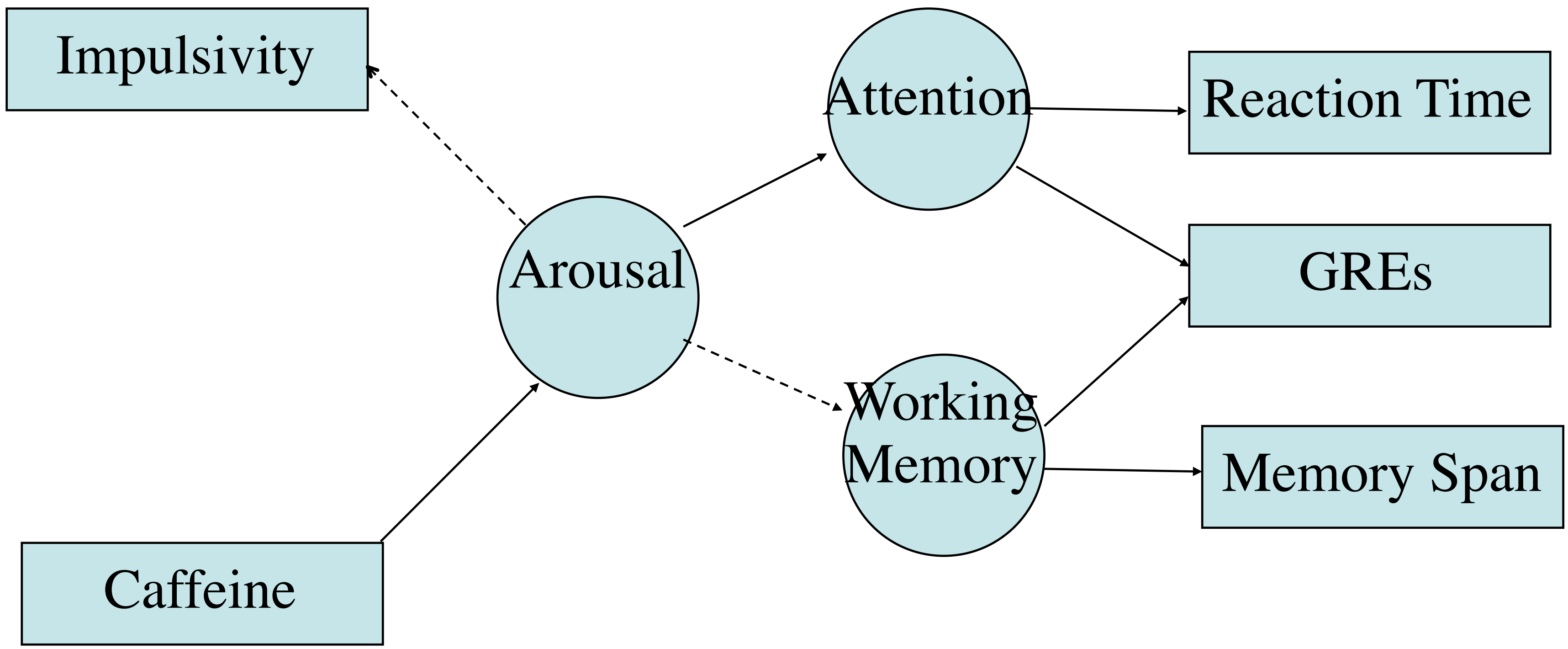
Theory and Theory Testing IV: Correlational inference



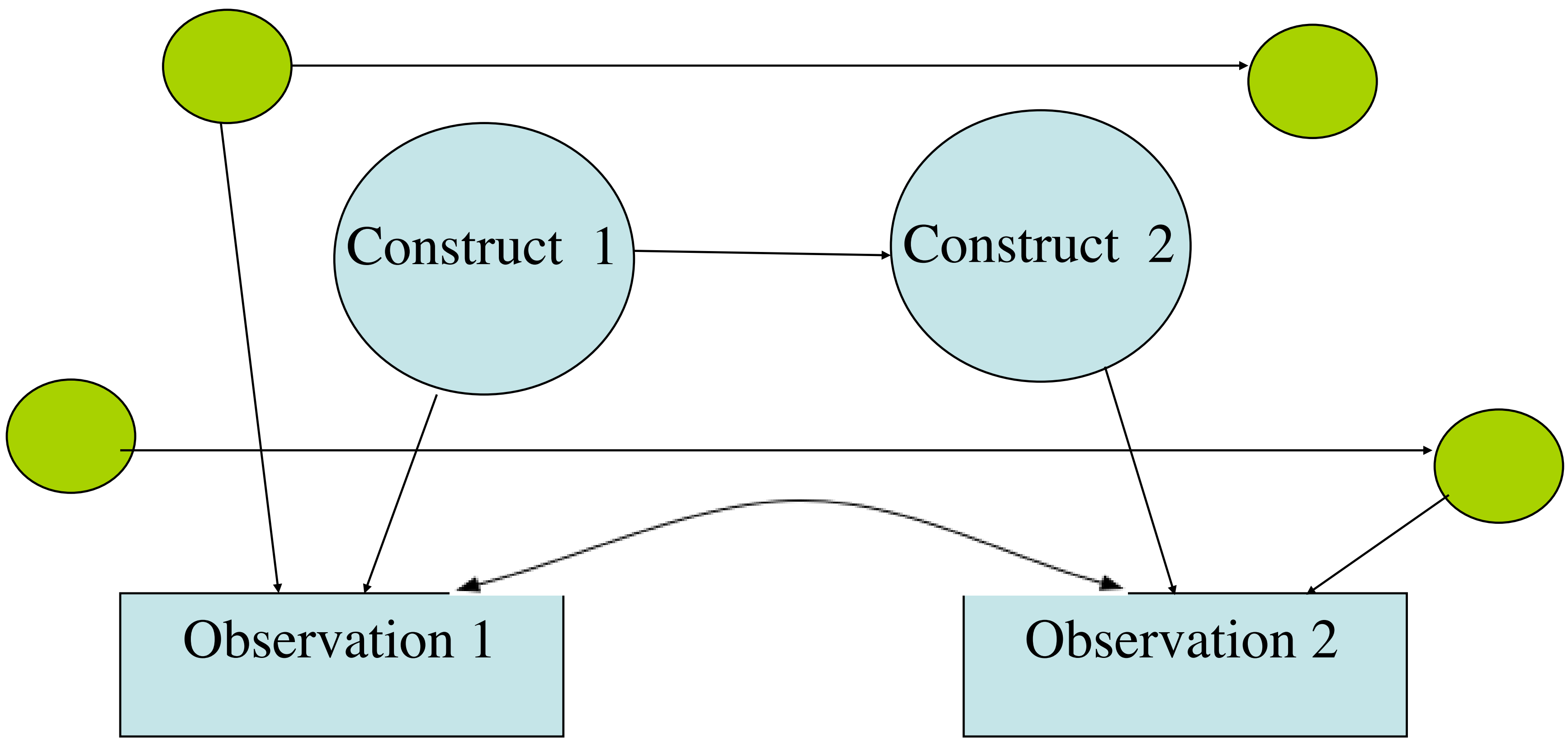
Theory and Theory Testing V: Alternative Explanations



Individual differences and general laws



Theory and Theory Testing VI: Eliminate Alternative Explanations

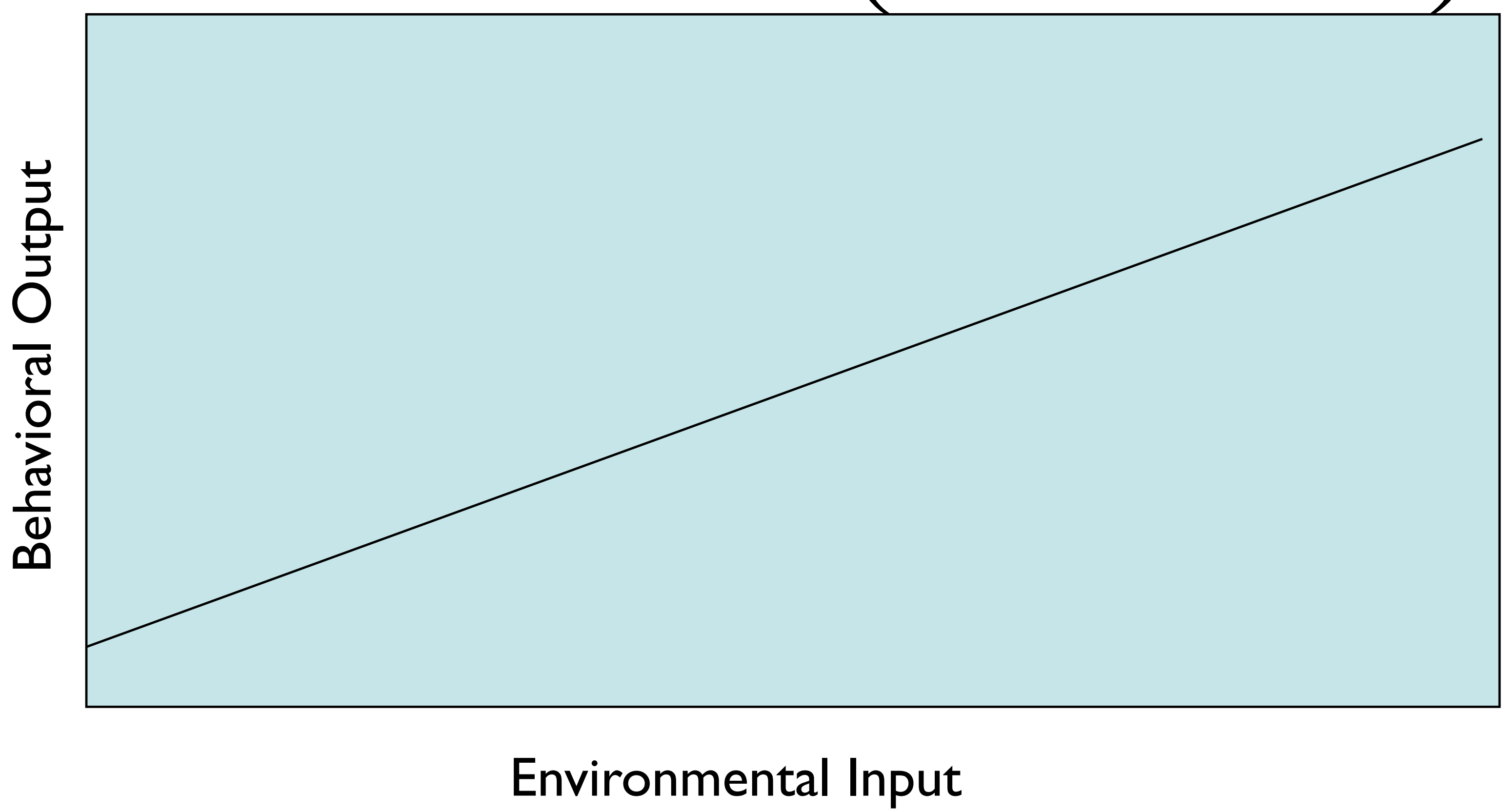


Types of Relationships

(Vale and Vale, 1969)

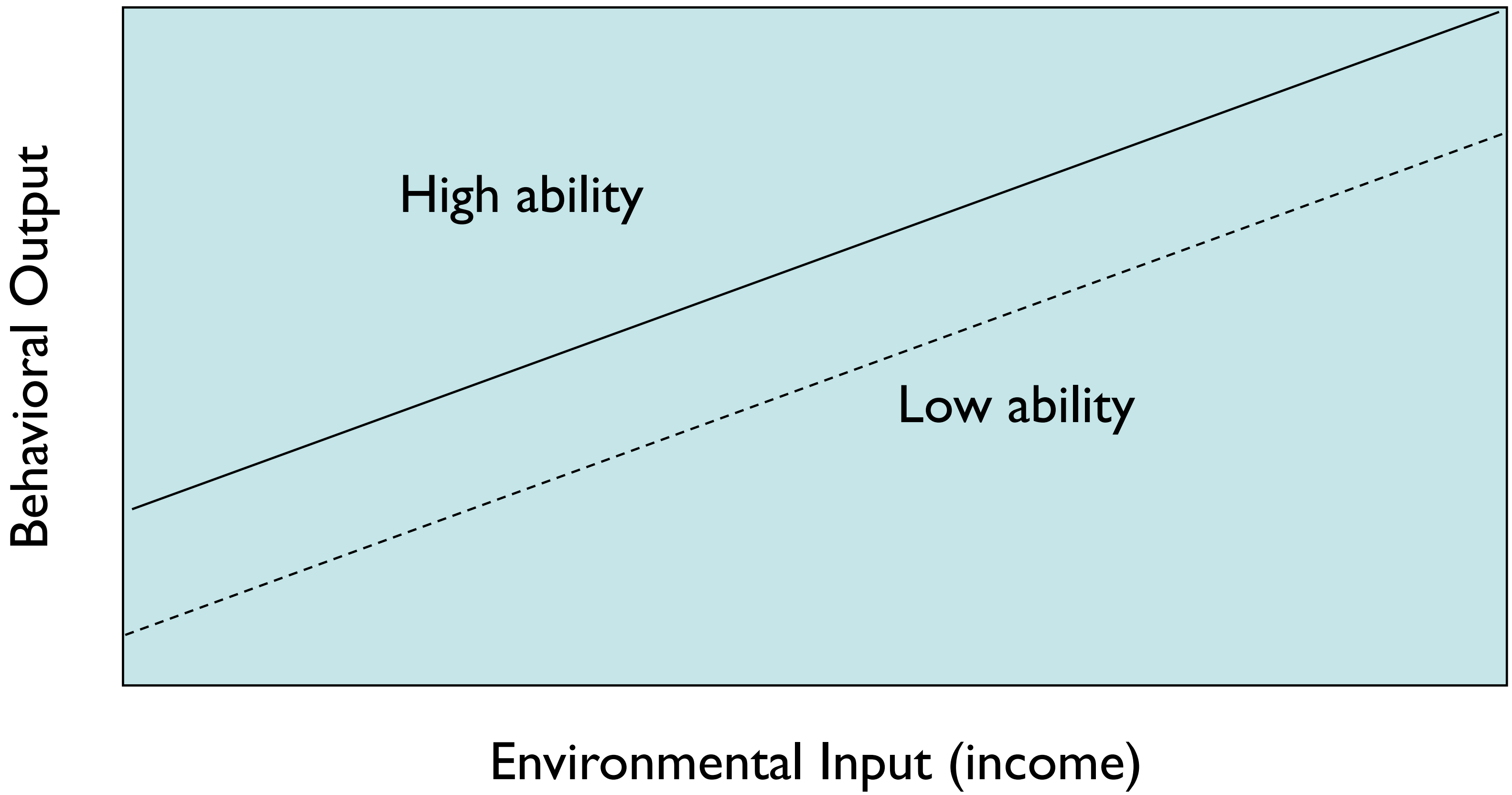
- Behavior = f(Situation)
- Behavior = f_1(Situation)+ f_2(Personality)
- Behavior = f_1(Situation)+f_2(Personality)+ f_3(Situation*Personality)
- Behavior = f_1(Situation * Personality)
- Behavior = idiosyncratic

Types of Relationships: Behavior = f(Situation)



Neuronal excitation = f(light intensity)

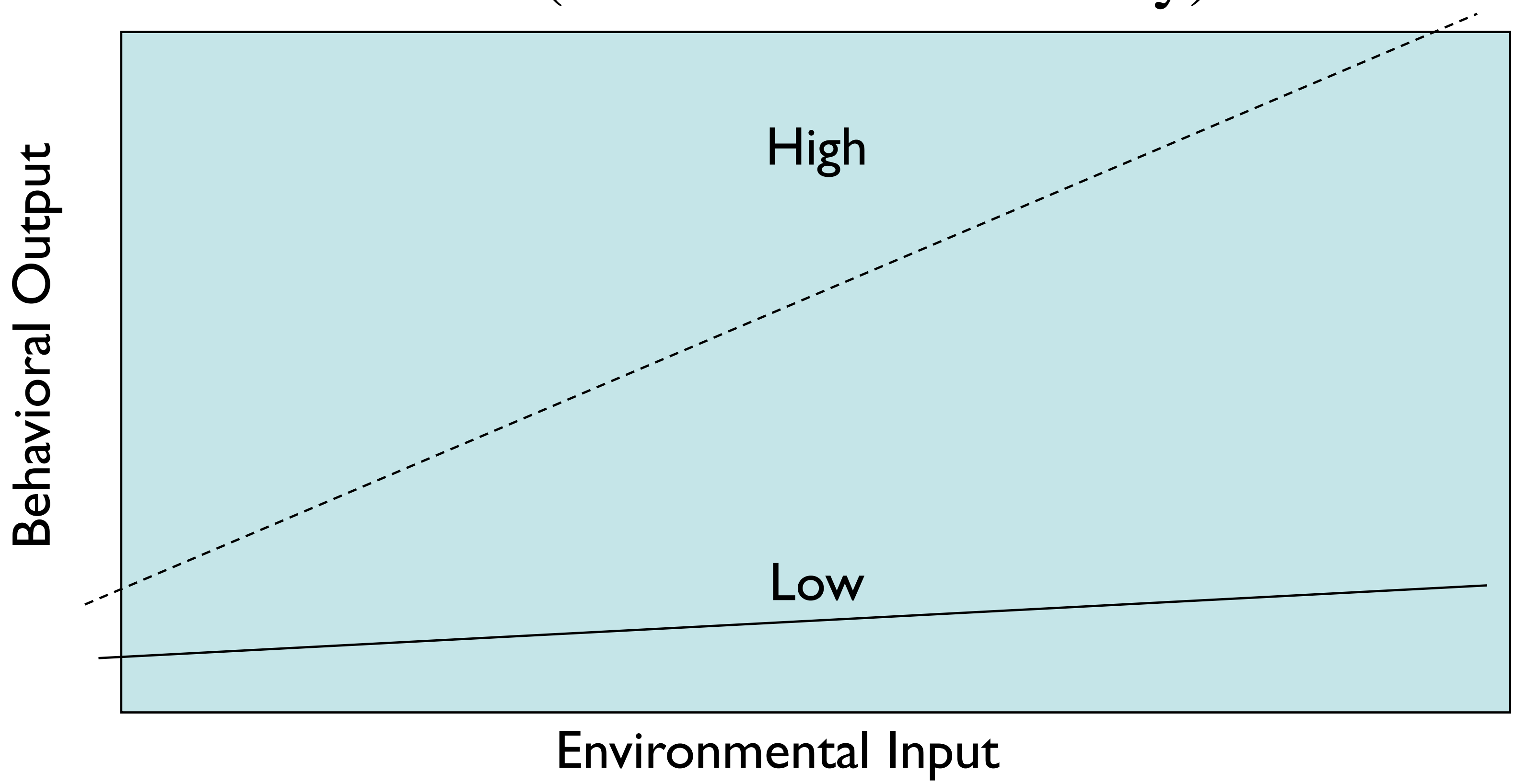
Types of Relationships: Behavior = f_1(Situation)+f_2(Person)



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

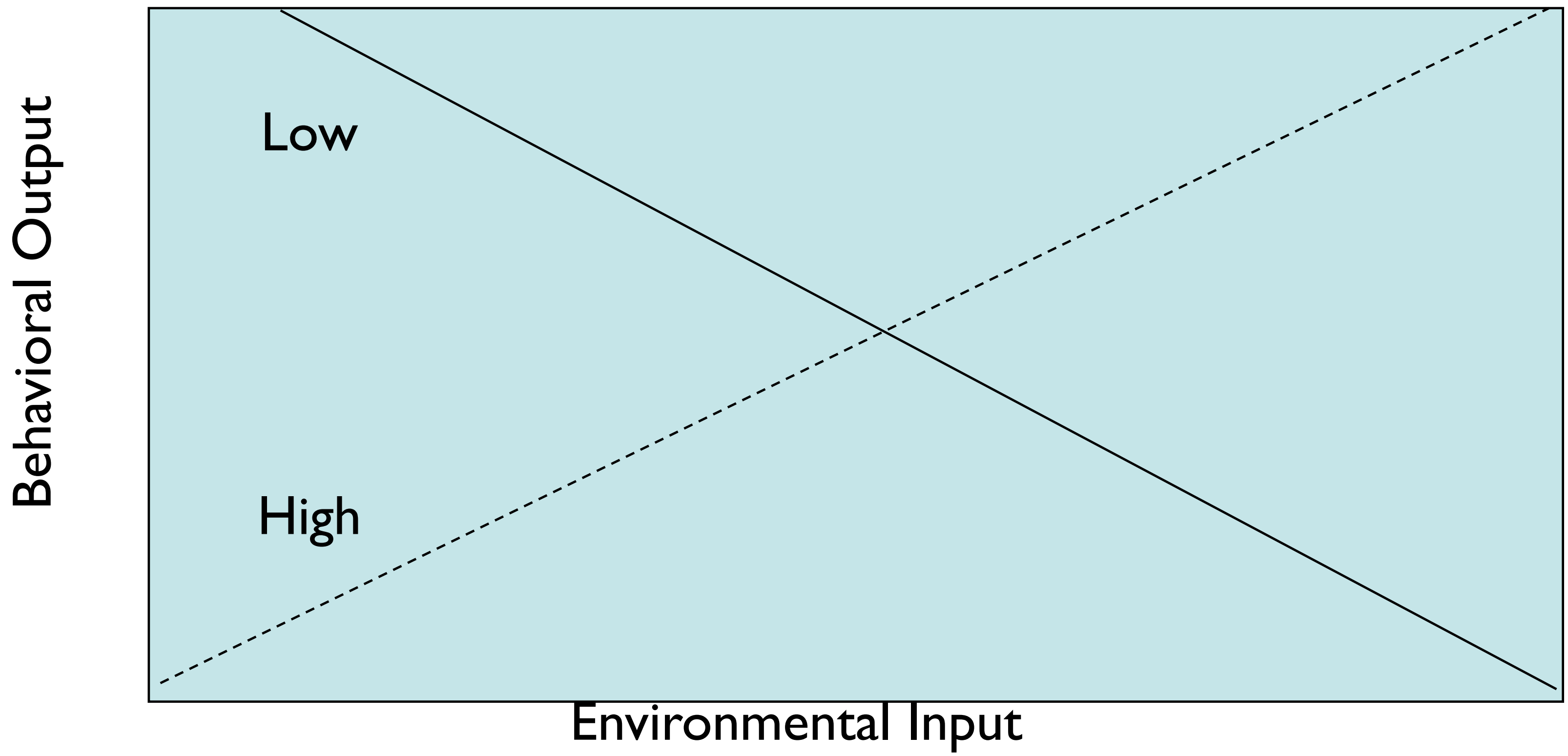
Types of Relationships:

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)

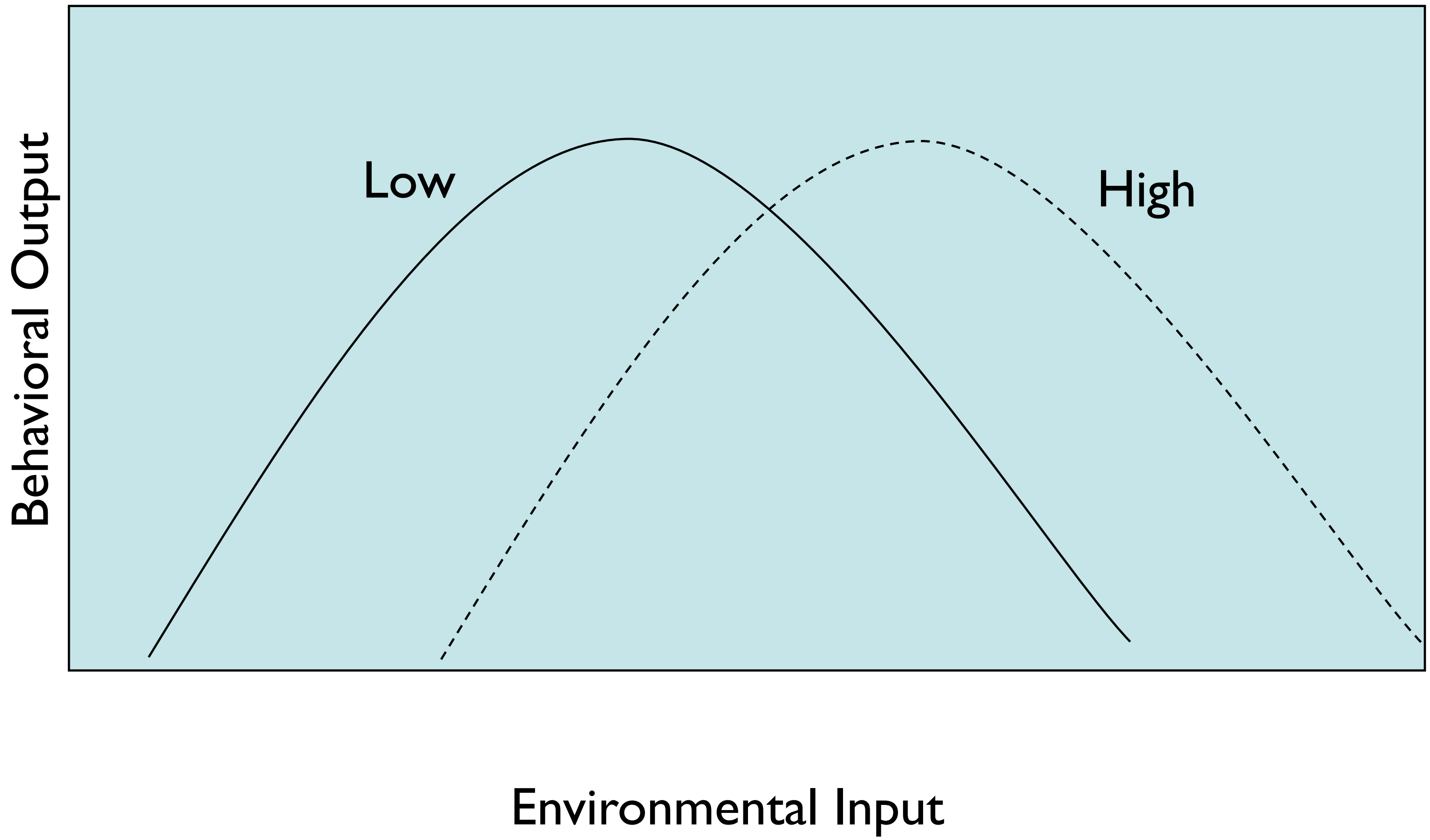
Types of Relationships: Behavior = f(Situation*Person)



Eating = f(preload * restraint)

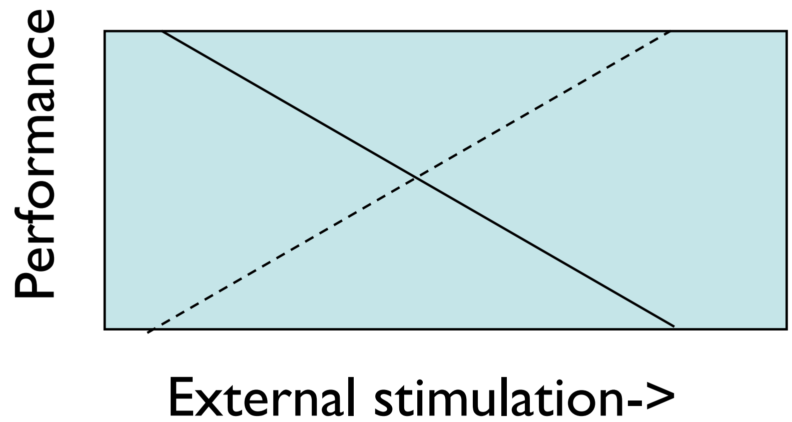
GRE = f(caffeine * impulsivity)

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



GRE = f(caffeine * impulsivity)

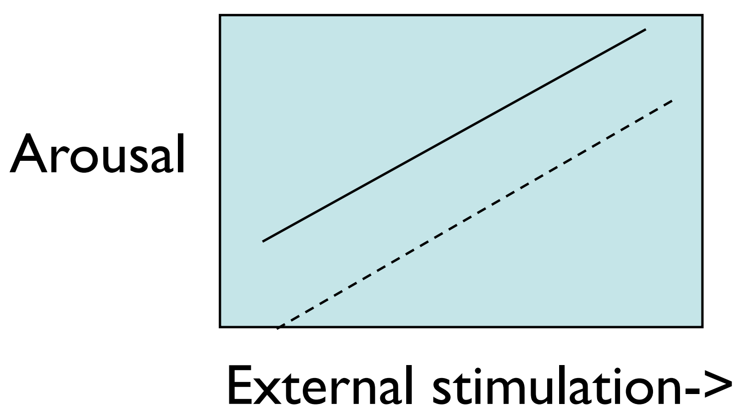
Persons, Situations, and Theory



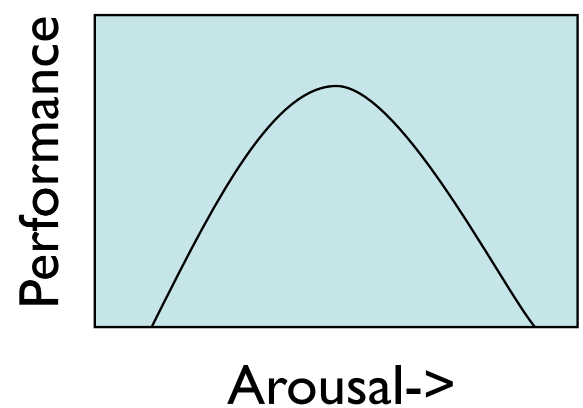
Observed relationship

Theoretical model

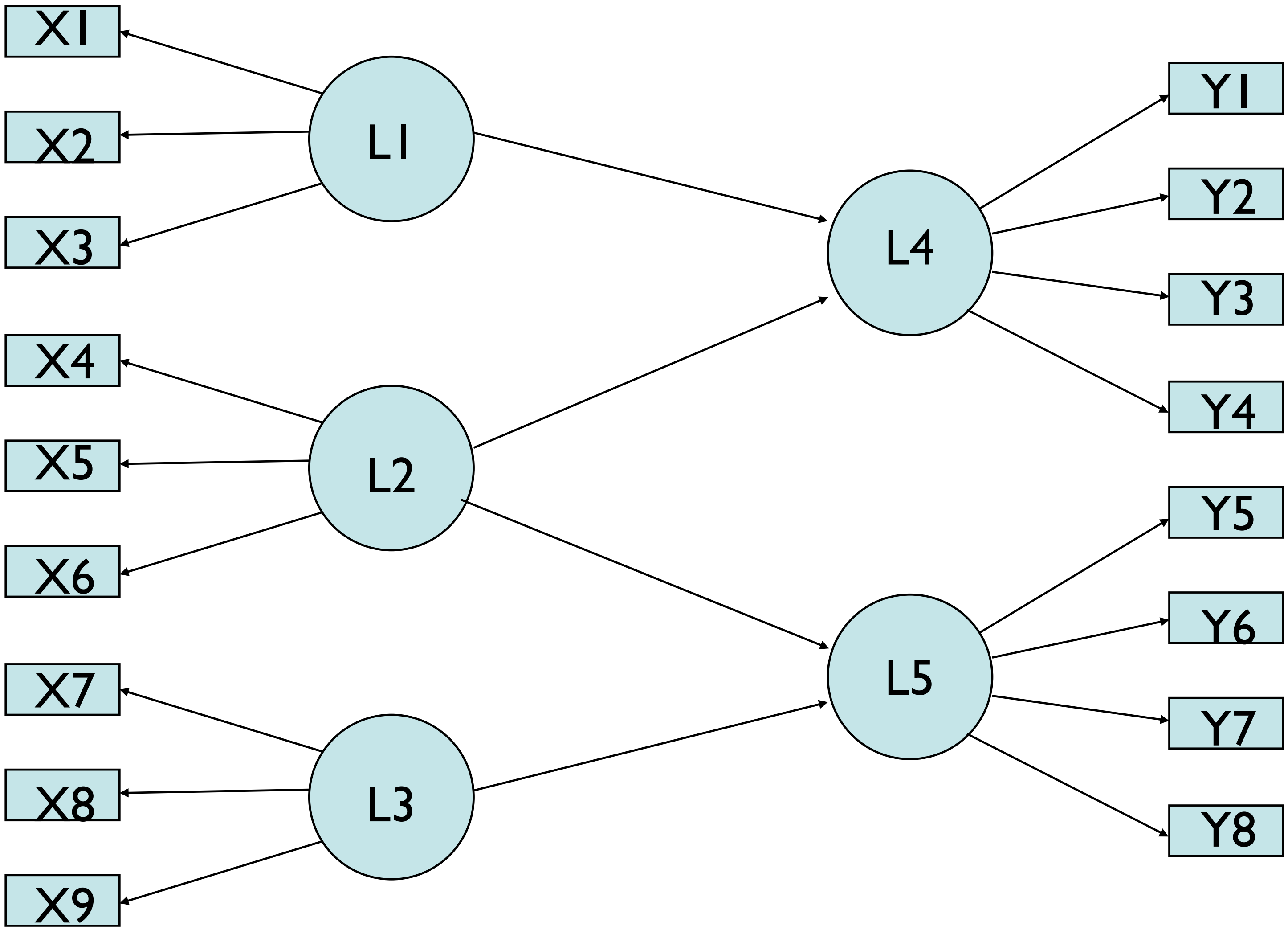
Individual Difference



General Law



Psychometric Theory: A conceptual Syllabus

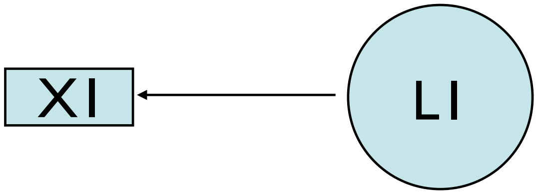


Data = Model + Residual

In all of psychometrics and statistics, five questions to ask are:

1. What is the model?
2. How well does it fit?
3. What are the plausible alternative models?
4. How well do they fit?
5. Is this better or worse than the current fit?

A Theory of Data: What can be measured



What is measured?

- Objects
- Individuals

What kind of measures are taken?

- Proximity (- distance)
- Order

What kind of comparisons are made?

- Single Dyads
- Pairs of Dyads

Assigning numbers to observations

2.718281828459050	3,412.1416
3.141592653589790	86,400
24	31,557,600
37	299,792,458
98.7	$6.022141 * 10^{23}$
365.25	42
365.25636305	X

Assigning numbers to observations: order vs. proximity

- Suppose we have observations X, Y, Z
- We assume each observation is a point on an attribute dimension (see Michell for a critique of the assumption of quantity) .
- Assign a number to each point.
- Two questions to ask:
  - What is the order of the points?
  - How far apart are the points?

Scaling of objects

- Consider $O = \{ o_1, o_2, \ldots o_n\}$ and
- $O \times O = \{(o_1, o_1), (o_1, o_2), \ldots (o_1, o_n),\ldots, (o_2, o_n), \ldots , (o_n, o_n)\}$
- Can we assign scale values to objects that satisfy an order relationship "≤"
  - $o_i \leq o_j$ and $o_i \geq o_j$ <=> $o_i = o_j$
  - $o_i \leq o_j$ and $o_j \leq o_k$ <=> $o_i \leq o_k$ (transitive)

Moh’s index of hardness

Mohs Hardness	Mineral	Scratch hardness
1	Talc	.59
2	Gypsum	.61
3	Calcite	3.44
4	Fluorite	3.05
5	Apaptite	5.2
6	Orthoclase Feldspar	37.2
7	Quartz	100
8	Topaz	121
9	Corundum	949
10	Diamond	85,300

Note the strong non-linearity of the top end of the scale

The Beaufort Scale

Force	Wind (Knots)	WMO Classification	Appearance of Wind Effects
0	Less than 1	Calm	Sea surface smooth and mirror-like
1	1-3	Light Air	Scaly ripples, no foam crests
2	4-6	Light Breeze	Small wavelets, crests glassy, no breaking
3	7-10	Gentle Breeze	Large wavelets, crests begin to break, scattered whitecaps
4	11-16	Moderate Breeze	Small waves 1-4 ft. becoming longer, numerous whitecaps
5	17-21	Fresh Breeze	Moderate waves 4-8 ft taking longer form, many whitecaps, some spray
6	22-27	Strong Breeze	Larger waves 8-13 ft, whitecaps common more spray
7	28-33	Near Gale	Sea heaps up, waves 13-20 ft, white foam streaks off breakers
8	34-40	Gale Moderately	high (13-20 ft) waves of greater length, edges of crests begin to break into spindrift, foam blown in streaks
9	41-47	Strong Gale	High waves (20 ft), sea begins to roll, dense streaks of foam, spray may reduce visibility
10	48-55	Storm	Very high waves (20-30 ft) with overhanging crests, sea white with densely blown foam, heavy rolling, lowered visibility
11	56-63	Violent Storm	Exceptionally high (30-45 ft) waves, foam patches cover sea, visibility more reduced
12	64+	Hurricane	Air filled with foam, waves over 45 ft, sea completely white with driving spray, visibility greatly reduced

Roughly linear with windspeed, but force of wind is quadratric effect of wind speed

Scaling of objects subjects as replicates

- Typical object scaling is concerned with order or location of objects
- Subjects are assumed to be random replicates of each other, differing only as a source of noise

Absolute scaling techniques

- “On a scale from 1 to 10” this ... is a ___?
- If A is 1 and B is 10, then what is C?
- College rankings based upon selectivity
- College rankings based upon “yield”
- Zagat ratings of restaurants

Absolute scaling difficulties

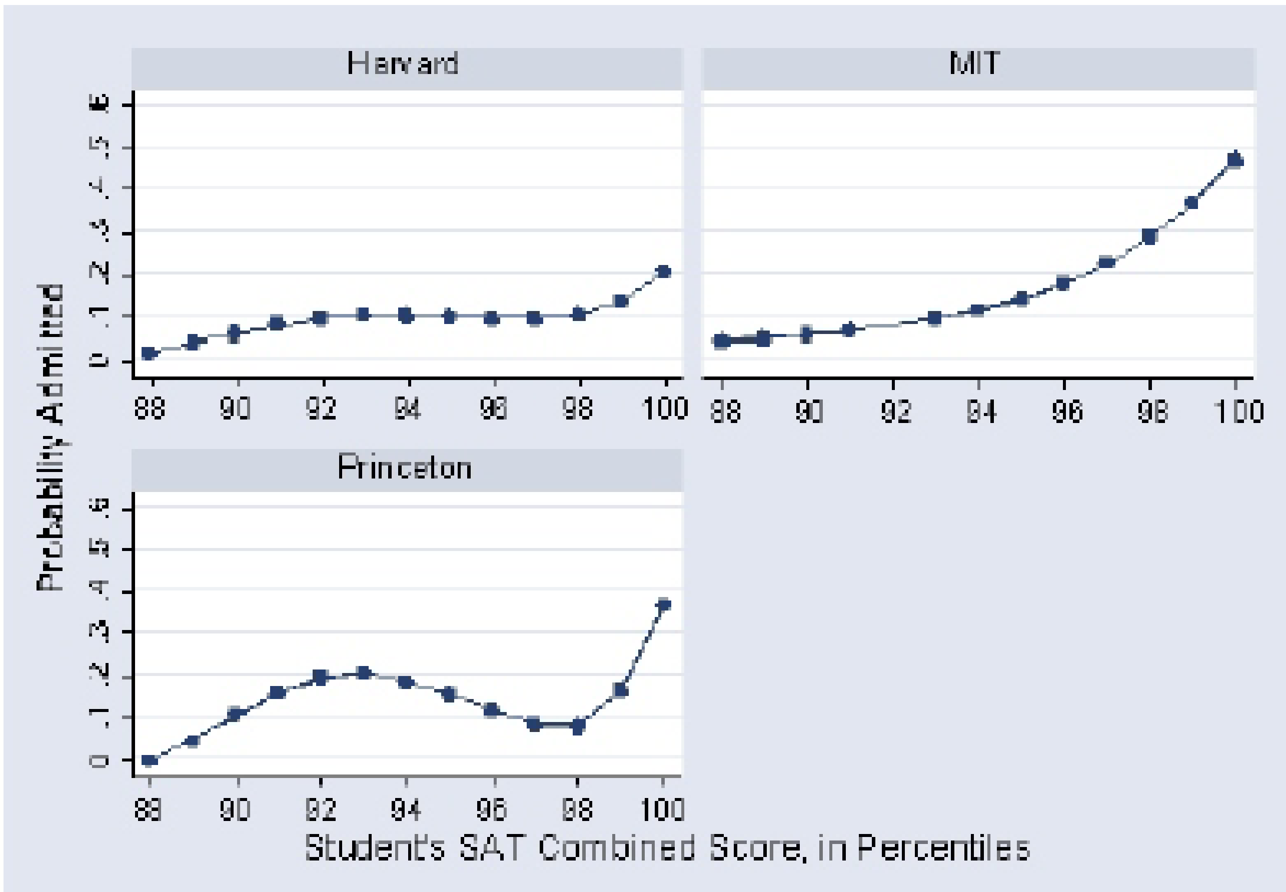
- "On a scale from 1 to 10" this ... is a ___?
  - sensitive to context effects
    - what if a new object appears?
    - Need unbounded scale
- If A is 1 and B is 10, then what is C?
  - results will depend upon A, B

Absolute scaling: artifacts

- College rankings based upon selectivity
  - accept/applied
    - encourage less able to apply
- College rankings based upon “yield”
  - matriculate/accepted
    - early admissions guarantee matriculation
    - don’t accept students who will not attend

College admission tricks

Increase the yield by rejecting students likely to go elsewhere.



Avery, C., Glickman, M, Hoxby, C., & Metrick, A. (2004) A revealed preference ranking of U.S. colleges and universities. http://www.nber.org/papers/w10803
(see also Avery, C., Glickman, M, Hoxby, C., & Metrick, A, 2013)

Models of scaling objects

- Assume each object (a, b,...z) has a scale value (A, B, ... Z) with some noise for each measurement.
- Probability of A > B increases with difference of between a and b
  - P(A>B) = f(a - b)
  - Can we find a function, f, such that equal differences in the latent variable (a, b, c) lead to equal differences in the observed variable?

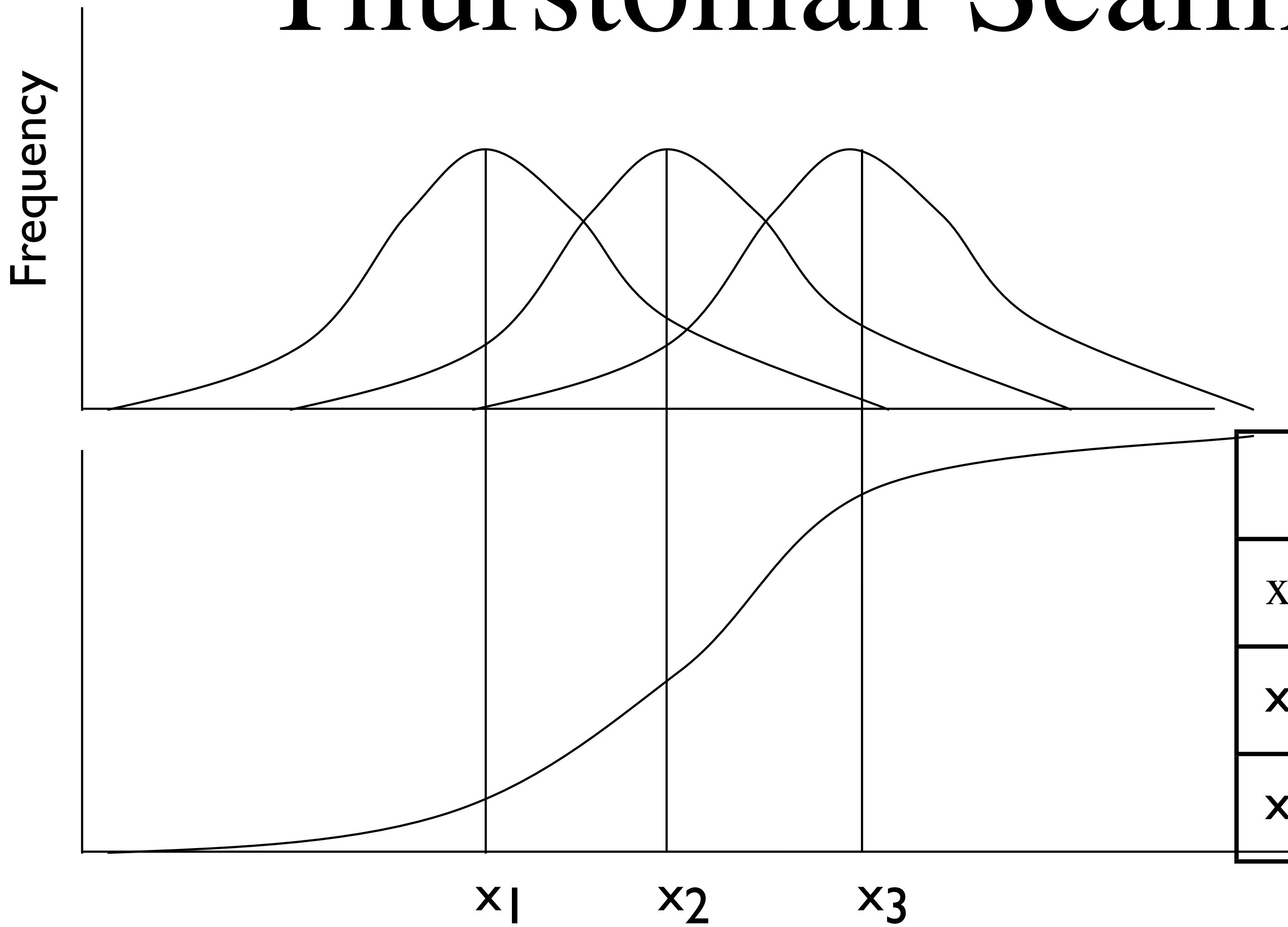
Models of scaling

- Given latent scores (a, b, ... z) find observed scores A=f(a), B = f(b), ... Z = f(z) such that iff a > b then A > B (an ordinal scale)
- Given latent scores (a, b, ... z) find observed scores A=f(a), B = f(b), ... Z = f(z) such that iff a-b > c-d then A-B > C - D (an interval scale)
- Given latent scores (a, b, ... z) find observed scores A=f(a), B = f(b), ... Z = f(z) such that iff a/b > c/d then A/B > C/ D (a ratio scale)

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 (Thurstone 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

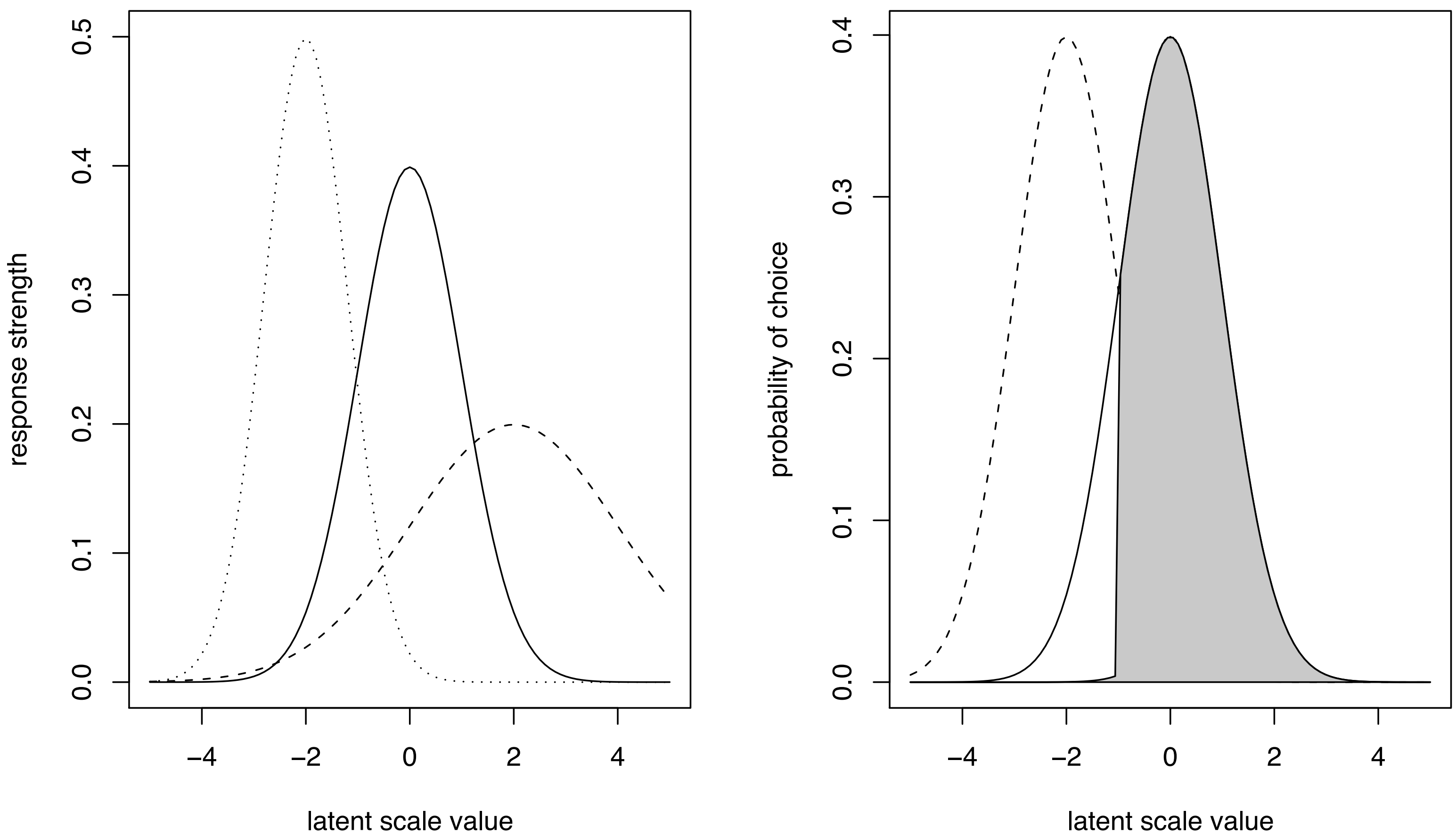
Thurstonian Scaling



	x_1	x_2	x_3
x_1	0.5	0.7	0.9
x_2	0.3	0.5	0.7
x_3	0.1	0.3	0.5

p(col > row)

Thurstone comparative judgment



Thurstone scaling:step 1: choice

```
> data(vegetables)
> round(veg,2)
```

	Turn	Cab	Beet	Asp	Car	Spin	S.Beans	Peas	Corn
Turn	0.50	0.82	0.77	0.81	0.88	0.89	0.90	0.89	0.93
Cab	0.18	0.50	0.60	0.72	0.74	0.74	0.81	0.84	0.86
Beet	0.23	0.40	0.50	0.56	0.74	0.68	0.84	0.80	0.82
Asp	0.19	0.28	0.44	0.50	0.56	0.59	0.68	0.60	0.73
Car	0.12	0.26	0.26	0.44	0.50	0.49	0.57	0.71	0.76
Spin	0.11	0.26	0.32	0.41	0.51	0.50	0.63	0.68	0.63
S.Beans	0.10	0.19	0.16	0.32	0.43	0.37	0.50	0.53	0.64
Peas	0.11	0.16	0.20	0.40	0.29	0.32	0.47	0.50	0.63
Corn	0.07	0.14	0.18	0.27	0.24	0.37	0.36	0.37	0.50

Thurstone scaling: step 2: normal transformation

```
> normal.values <- qnorm(as.matrix(veg))
> round(normal.values,2)
```

	Turn	Cab	Beet	Asp	Car	Spin	S.Beans	Peas	Corn
Turn	0.00	0.91	0.74	0.88	1.17	1.24	1.28	1.24	1.45
Cab	-0.91	0.00	0.26	0.59	0.65	0.63	0.88	1.02	1.07
Beet	-0.74	-0.26	0.00	0.15	0.63	0.46	1.02	0.83	0.91
Asp	-0.88	-0.59	-0.15	0.00	0.15	0.22	0.46	0.26	0.61
Car	-1.17	-0.65	-0.63	-0.15	0.00	-0.02	0.19	0.55	0.72
Spin	-1.24	-0.63	-0.46	-0.22	0.02	0.00	0.33	0.47	0.33
S.Beans	-1.28	-0.88	-1.02	-0.46	-0.19	-0.33	0.00	0.07	0.36
Peas	-1.24	-1.02	-0.83	-0.26	-0.55	-0.47	-0.07	0.00	0.33
Corn	-1.45	-1.07	-0.91	-0.61	-0.72	-0.33	-0.36	-0.33	0.00

Thurstone Step 3: average z score and rescale

Sums of z scores
Average z score
Rescale to set minimum to 0

	sums	means	values
Turn	-8.89	-0.99	0.00
Cab	-4.19	-0.47	0.52
Beet	-3.00	-0.33	0.65
Asp	-0.07	-0.01	0.98
Car	1.16	0.13	1.12
Spin	1.40	0.16	1.14
S.Beans	3.71	0.41	1.40
Peas	4.10	0.46	1.44
Corn	5.77	0.64	1.63

```
> sums <- colSums(normal.values)
> means <- colMeans(normal.values)
> values <- means - min(values)
> thurstone.df <- data.frame(sums,means,values)
> round(thurstone.df,2)
```

Goodness of fit test: $1 - \text{residual}^2/\text{original}^2 = .99$

Create a function to do this

```
> thurstone
function (x, ranks = FALSE, digits = 2)
{
    cl <- match.call()
    if (ranks) {
        choice <- choice.mat(x)
    }
    else {
        if (is.matrix(x))
            choice <- x
        choice <- as.matrix(x)
    }
    scale.values <- colMeans(qnorm(choice)) -
min(colMeans(qnorm(choice)))
    model <- pnorm(-scale.values %+% t(scale.values))
    error <- model - choice
    fit <- 1 - (sum(error * error)/sum(choice * choice))
    result <- list(scale = round(scale.values, digits), GF = fit,
        residual = error, Call = cl)
    class(result) <- c("psych", "thurstone")
    return(result)
}
```

Thurstone Model

```
> veg.scale <- thurstone(veg)  #Apply our new function
> veg.scale
Thurstonian scale (case 5) scale values
Call: thurstone(x = veg)
   Turn      Cab     Beet      Asp      Car     Spin S.Beans     Peas     Corn
   0.00     0.52     0.65     0.98     1.12     1.14     1.40     1.44     1.63

 Goodness of fit of model   0.99

> values <- veg.scale$scale
> model <-   -values %+% t(values)
> colnames(model) <- rownames(model) <- names(values)
> model
              Turn   Cab  Beet   Asp   Car  Spin S.Beans  Peas Corn
      Turn     0.00  0.52  0.65  0.98  1.12  1.14    1.40  1.44 1.63
      Cab     -0.52  0.00  0.13  0.46  0.60  0.62    0.88  0.92 1.11
      Beet    -0.65 -0.13  0.00  0.33  0.47  0.49    0.75  0.79 0.98
      Asp     -0.98 -0.46 -0.33  0.00  0.14  0.16    0.42  0.46 0.65
      Car     -1.12 -0.60 -0.47 -0.14  0.00  0.02    0.28  0.32 0.51
      Spin    -1.14 -0.62 -0.49 -0.16 -0.02  0.00    0.26  0.30 0.49
      S.Beans -1.40 -0.88 -0.75 -0.42 -0.28 -0.26    0.00  0.04 0.23
      Peas    -1.44 -0.92 -0.79 -0.46 -0.32 -0.30   -0.04  0.00 0.19
      Corn    -1.63 -1.11 -0.98 -0.65 -0.51 -0.49   -0.23 -0.19 0.00
```

Thurstone: model

```
> model <- pnorm(model) #convert Z scores to probability
> round(model,2)
```

	Turn	Cab	Beet	Asp	Car	Spin	S.Beans	Peas	Corn
Turn	0.50	0.70	0.74	0.84	0.87	0.87	0.92	0.93	0.95
Cab	0.30	0.50	0.55	0.68	0.73	0.73	0.81	0.82	0.87
Beet	0.26	0.45	0.50	0.63	0.68	0.69	0.77	0.79	0.84
Asp	0.16	0.32	0.37	0.50	0.56	0.56	0.66	0.68	0.74
Car	0.13	0.27	0.32	0.44	0.50	0.51	0.61	0.63	0.69
Spin	0.13	0.27	0.31	0.44	0.49	0.50	0.60	0.62	0.69
S.Beans	0.08	0.19	0.23	0.34	0.39	0.40	0.50	0.52	0.59
Peas	0.07	0.18	0.21	0.32	0.37	0.38	0.48	0.50	0.58
Corn	0.05	0.13	0.16	0.26	0.31	0.31	0.41	0.42	0.50

Thurstone: Data-Model

```
> error <- veg - model
> round(error,2)
```

	Turn	Cab	Beet	Asp	Car	Spin	S.Beans	Peas	Corn
Turn	0.00	0.12	0.03	-0.03	0.01	0.02	-0.02	-0.03	-0.02
Cab	-0.12	0.00	0.05	0.05	0.02	0.00	0.00	0.02	-0.01
Beet	-0.03	-0.05	0.00	-0.07	0.06	-0.01	0.07	0.01	-0.02
Asp	0.03	-0.05	0.07	0.00	0.01	0.02	0.01	-0.08	-0.01
Car	-0.01	-0.02	-0.06	-0.01	0.00	-0.02	-0.04	0.08	0.07
Spin	-0.02	0.00	0.01	-0.02	0.02	0.00	0.03	0.06	-0.06
S.Beans	0.02	0.00	-0.07	-0.01	0.04	-0.03	0.00	0.01	0.05
Peas	0.03	-0.02	-0.01	0.08	-0.08	-0.06	-0.01	0.00	0.05
Corn	0.02	0.01	0.02	0.01	-0.07	0.06	-0.05	-0.05	0.00

Summarize Residuals

```
describe(error,skew=FALSE)
```

	var	n	mean	sd	median	trimmed	mad	min	max	range	se
Turn	1	9	-0.01	0.05	0.00	-0.01	0.03	-0.12	0.03	0.15	0.02
Cab	2	9	0.00	0.05	0.00	0.00	0.02	-0.05	0.12	0.17	0.02
Beet	3	9	0.00	0.05	0.01	0.00	0.04	-0.07	0.07	0.14	0.02
Asp	4	9	0.00	0.04	-0.01	0.00	0.03	-0.07	0.08	0.14	0.01
Car	5	9	0.00	0.05	0.01	0.00	0.01	-0.08	0.06	0.14	0.02
Spin	6	9	0.00	0.03	0.00	0.00	0.03	-0.06	0.06	0.12	0.01
S.Beans	7	9	0.00	0.04	0.00	0.00	0.03	-0.05	0.07	0.12	0.01
Peas	8	9	0.00	0.05	0.01	0.00	0.06	-0.08	0.08	0.16	0.02
Corn	9	9	0.01	0.04	-0.01	0.01	0.02	-0.06	0.07	0.13	0.01

Find fit

- Many indices of fit
- Typical is 1-error^2/data^2
- Fit of Thurstone =
  - > 1-sum(error^2)/sum(veg^2)
  - [1] 0.99

Alternative scaling models

- Thurstone assumes normal deviations
- Logistic model produces similar results
  - used in scaling chess players, sports teams
    - win/loss record
  - scaling of colleges by where students choose to go (choice of A vs. B)
    - more difficult to fake

Compare to other scaling methods

- Thurstone assumes normal error of preference
- logistic model is alternative model
- so are other rank difference models
- all about the same in terms of fit

At least two ways to collect choice data

- Paired comparisons:
  - Is X > Y
  - Is Y > Z
  - ... n*(n-1)/2 pairs
- Rank orders (X>Y>Z>W) => a set of pairs
  - X>Y, X>Z, Y>Z, X>W, Y> W, Z>W

Thurstonian scaling in R

- code for Thurstonian case V is in psych
  - data(vegetables)
  - thu <- thurstone(veg)
  - thu   #shows values and fits
  - thu$residual #shows residuals
- brief discussion of Thurstonian and alternative scaling models with links at
  - http://personality-project.org/r/thurstone.html

What is this thing called R?

- A quick introduction to R: gettingstarted
- personality-project.org/r/psych

Assigning numbers: do they form a metric space?

- Suppose we have observations X, Y, Z
- We assume each observation is a point on (possibly many) attribute dimension(s)
- Assign a number to each point.
- Do these numbers form a metric space?
- Requires finding a distance between points

Metric spaces and the axioms of a distance measure

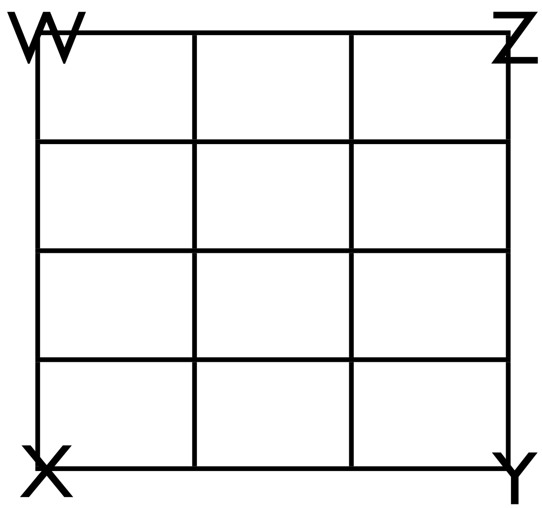
- A metric space is a set of points with a distance function, D, which meets the following properties
- Distance is symmetric, positive definite, and satisfies the triangle inequality:
  - $D(X, Y) = D(Y, X)$ (symmetric)
  - $D(X,Y) \geq 0$ (non negativity)
  - $D(X,Y) = 0$ iff $X=Y$ ($D(X,X)=0$ reflexive)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

Two unidimensional metric spaces

	X	Y	Z
X	0	1	2
Y	1	0	1
Z	2	1	0
att	1	2	3

	X	Y	Z
X	0	3	8
Y	3	0	5
Z	8	5	0
att	1	4	9

Multidimensional spaces using alternative metrics



Euclidian

	X	Y	Z	W
X	0	3	5	4
Y	3	0	4	5
Z	5	4	0	3
W	4	5	3	0

City block

	X	Y	Z	W
X	0	3	7	4
Y	3	0	4	7
Z	7	4	0	3
W	4	7	3	0

A non metric space

	X	Y	Z
X	0	1	2
Y	1	0	2
Z	0	0	0
att	1	1	4

Multidimensional scaling

- Given a n * n distance matrix, is it possible to represent the data in a k dimensional space?
- How well does that model fit?
- How sensitive is the model to transformations of the original distances?
- Need to find distances
  - absolute distance between pairs
  - ranks of distances between pairs of pairs

Distances between US cities

	ATL	BOS	ORD	DCA	DEN	LAX	MIA	JFK	SEA	SFO	MSY
ATL	0	934	585	542	1209	1942	605	751	2181	2139	424
BOS	934	0	853	392	1769	2601	1252	183	2492	2700	1356
ORD	585	853	0	598	918	1748	1187	720	1736	1857	830
DCA	542	392	598	0	1493	2305	922	209	2328	2442	964
DEN	1209	1769	918	1493	0	836	1723	1636	1023	951	1079
LAX	1942	2601	1748	2305	836	0	2345	2461	957	341	1679
MIA	605	1252	1187	922	1723	2345	0	1092	2733	2594	669
JFK	751	183	720	209	1636	2461	1092	0	2412	2577	1173
SEA	2181	2492	1736	2328	1023	957	2733	2412	0	681	2101
SFO	2139	2700	1857	2442	951	341	2594	2577	681	0	1925
MSY	424	1356	830	964	1079	1679	669	1173	2101	1925	0

cities

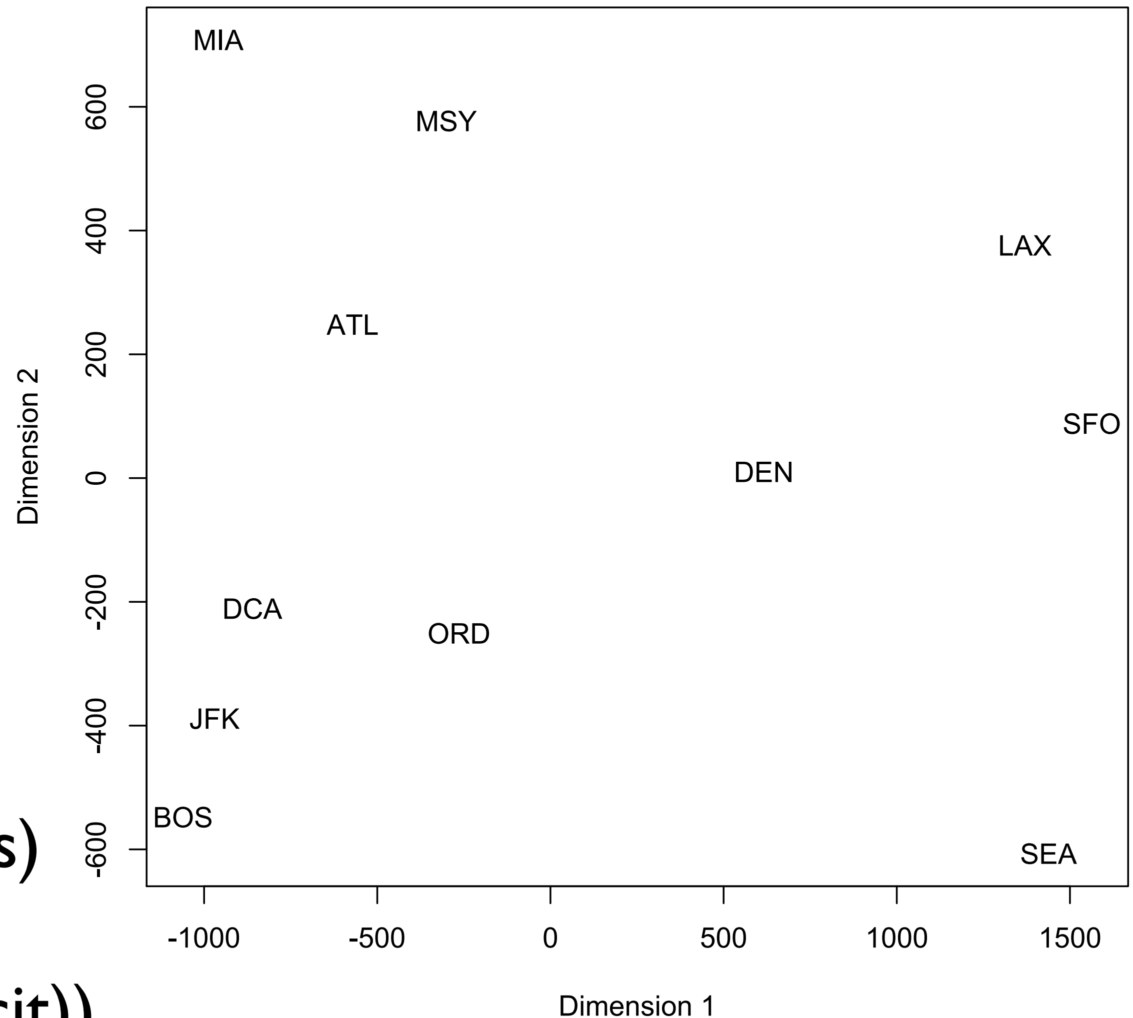
Multidimensional Scaling

	Dimension 1	Dimension 2
ATL	-571	248
BOS	-1061	-548
ORD	-264	-251
DCA	-861	-211
DEN	616	10
LAX	1370	376
MIA	-959	708
JFK	-970	-389
SEA	1438	-607
SFO	1563	88
MSY	-301	577

cmdscale(cities)

round(cmdscale(cities),0)

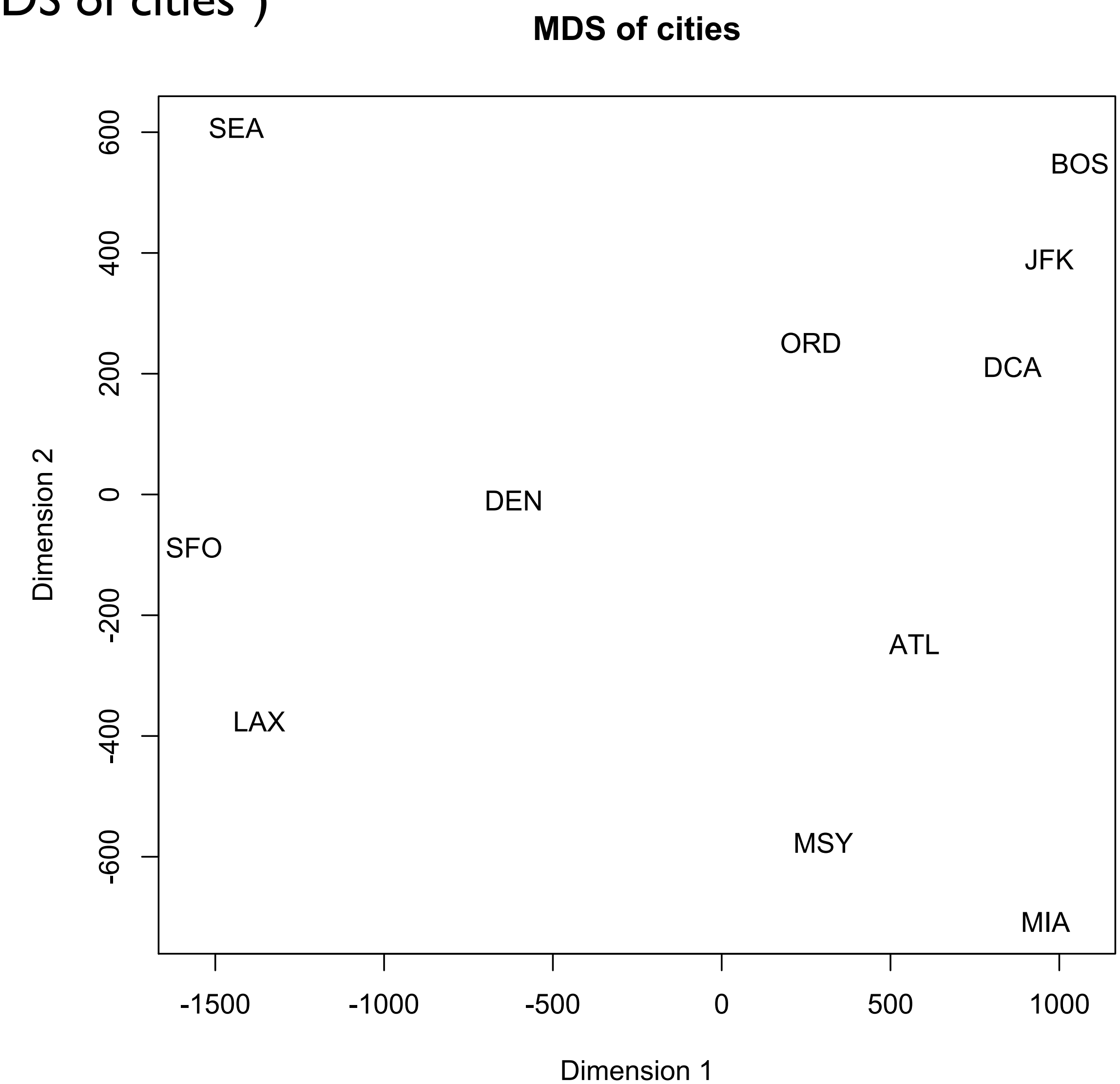
Spatial representation



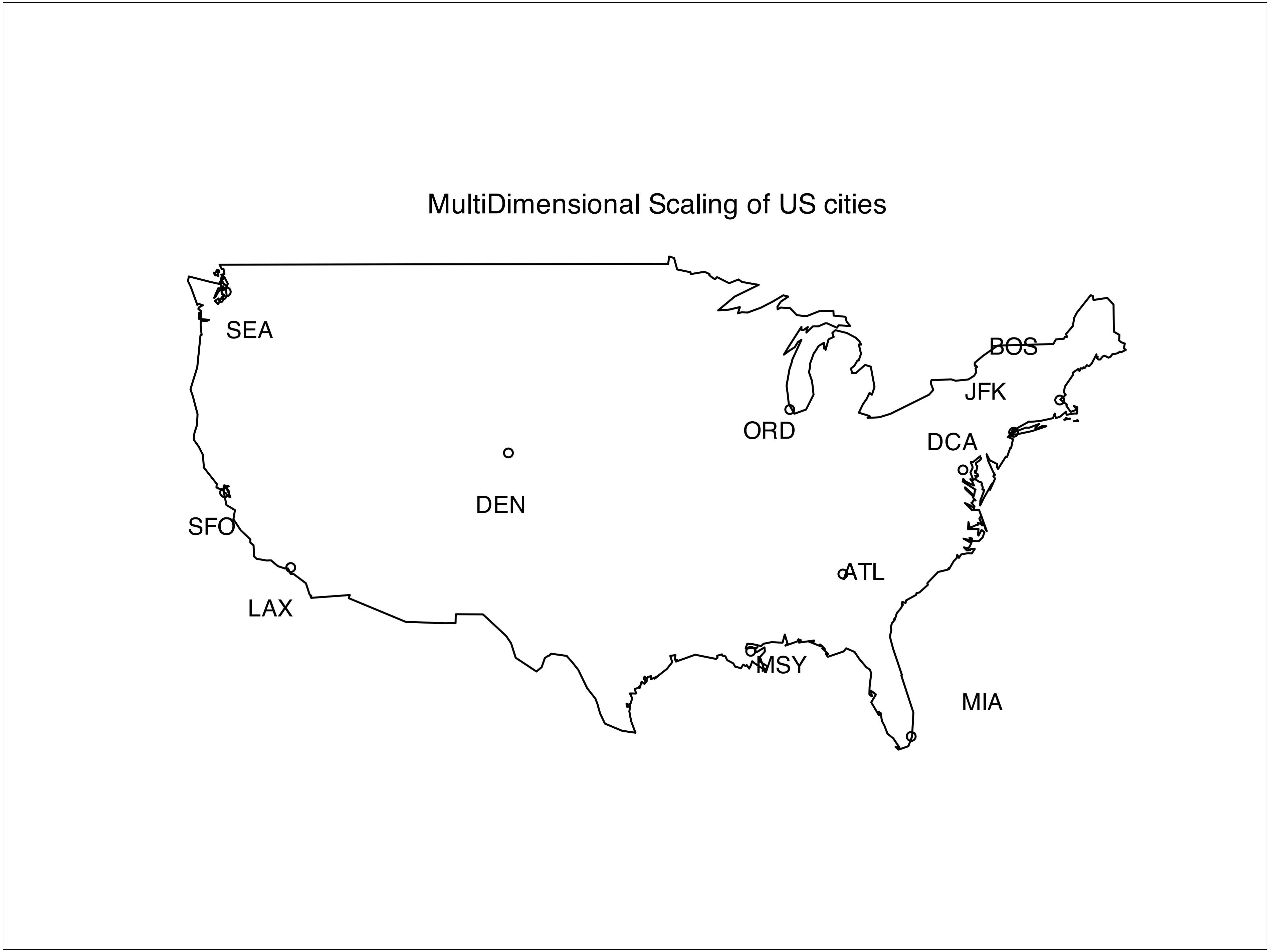

```
> cit <-cmdscale(cities)
> plot(cit,typ="n")
> text(cit,rownames(cit))
```

A more familiar map

```
> plot(-cit,typ="n",main="MDS of cities")
> text(-cit,rownames(cit))
```



Compare with classic solution

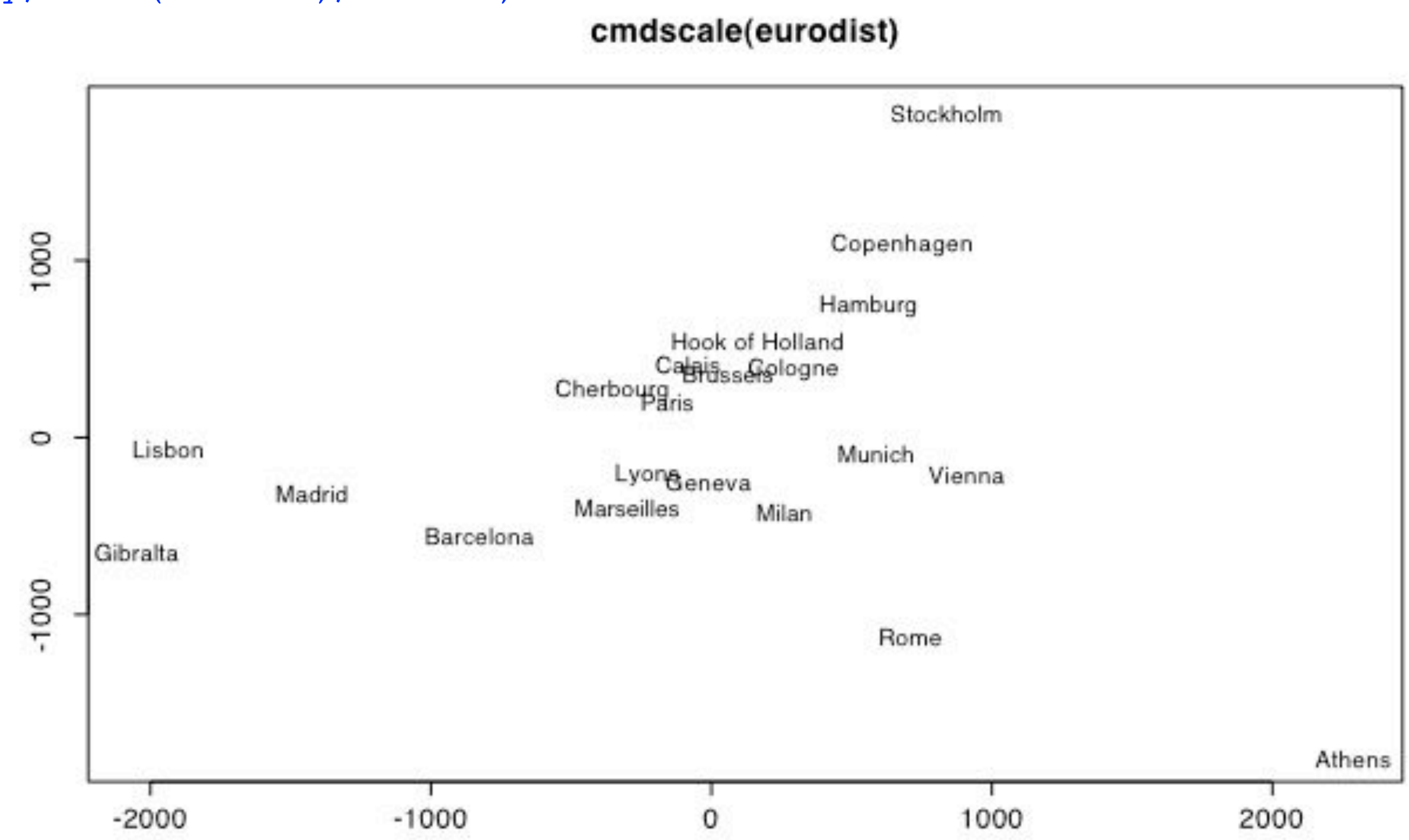

MultiDimensional Scaling of US cities
SEA
BOS
JFK
ORD
DCA
DEN
SFO
ATL
LAX
MSY
MIA

R code for MDS

- http://personality-project.org/r/mds.html
- compare cmdscale (metric mds) with isoMDS (non-metric scaling)
- ALSCAL and KYST are standard packages in SPSS
- Individual Differences models of MDS include INDSCAL and INDIFF
- ALSCAL is now available in R

Metric scaling of 28 European cities

```
loc <- cmdscale(eurodist)
x <- loc[,1]
y <- -loc[,2]
plot(x, y, type="n", xlab="", ylab="", main="cmdscale(eurodist)")
text(x, y, names(eurodist), cex=0.8)
```



Types of data collected vs. types of questions asked

- Ask Si about
  - O
  - O x O
- infer
  - O x O
  - (O x O) x (O x O)
  - S x O

	s_1	...	s_n	o_1	...	o_n
s_1						
...						
s_n						
o_1						
...						
o_n						

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

	Single Dyads	Pairs of Dyads
Single Stimuli	Measurement (S*O) $s_i < o_j$ (Abilities) $\|s_i - o_j\| < d$ (attitudes)	$\|s_i - o_j\| < \|s_k - o_l\|$ Unfolding (S*O)*(S*O) Preferential choice
Pairs of Stimuli	Scaling of stimuli (O * O) $o_i < o_j$	MDS (O*O)*(O*O) $\|o_i - o_j\| < \|o_k - o_l\|$

Individual differences in Multidimensional Scaling
S * (O*O) * (O*O)

Coombs' typology

- O x O   ($o_i < o_j$)   Scaling
- (OxO) x (OxO) $|o_i - o_j| < |o_k - o_l|$ MDS
- S x O   (two types of comparisons)
  - ($s_i < o_j$)   measurement of ability
  - $|s_i - o_j| < d$ measurement of attitude
- (SxO) x (SxO) preferential choice
  - $|s_i - o_j| < |s_k - o_l|$ or $|s_i - o_j| < |s_i - o_l|$

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? $|s_i - o_j| < |s_i - o_l|$

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

Preferential Choice: Individual (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

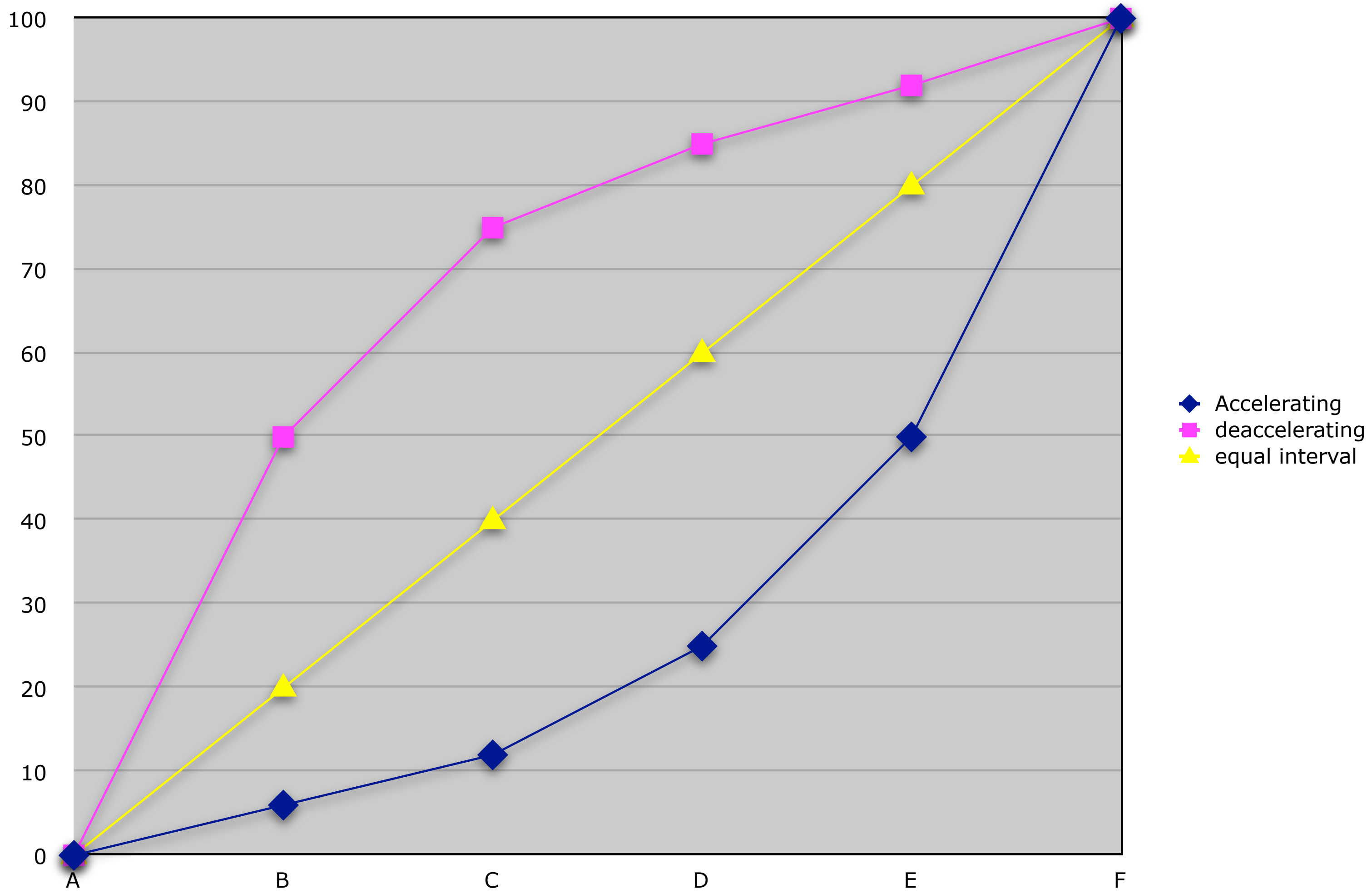
Preferential choice-underlying model

- 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.

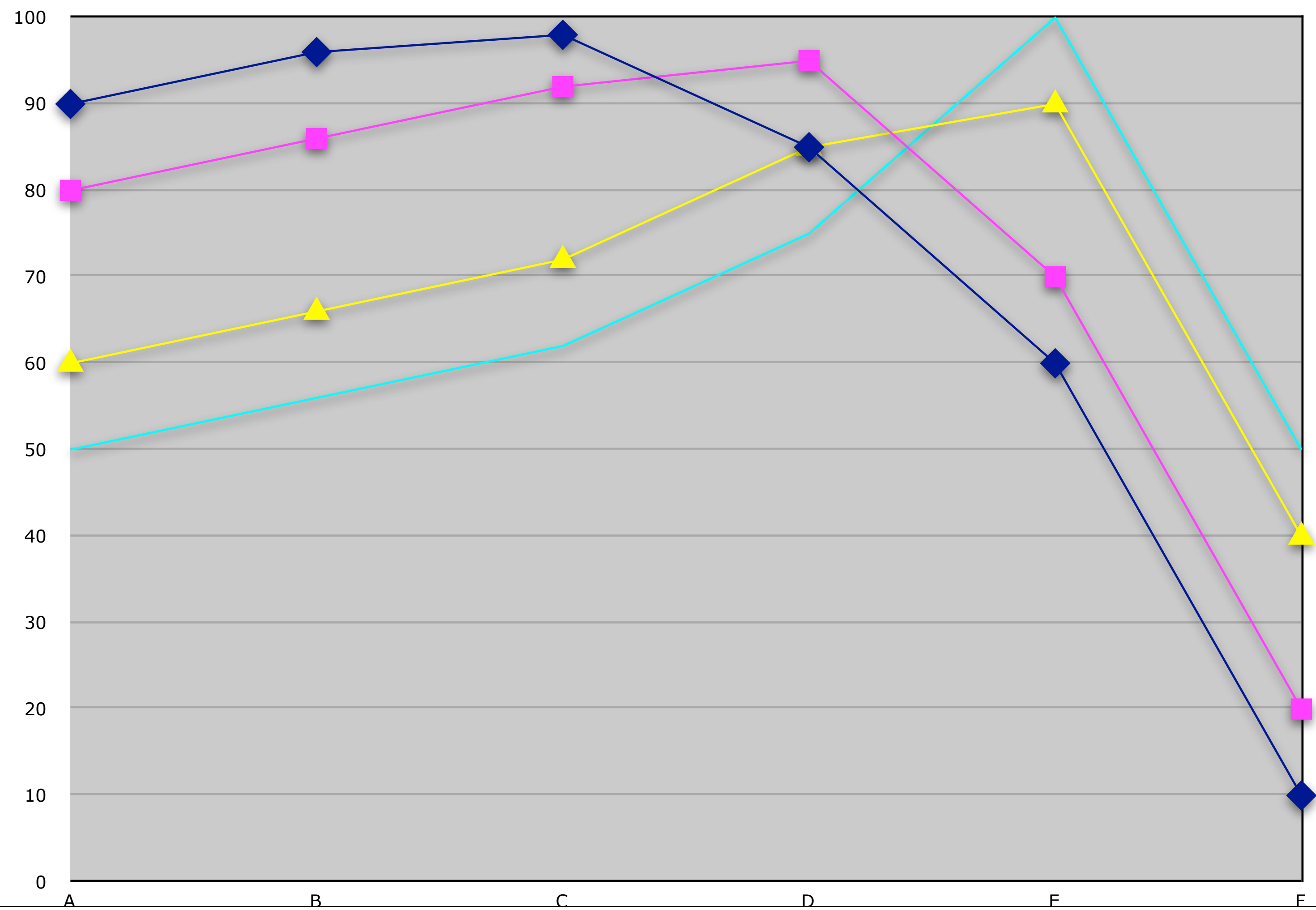
Questions about Scale Ratings

- Do subjects understand instructions?
- Can people give accurate representations of scale value to objects?
- Can we find subject locations in preferential space?

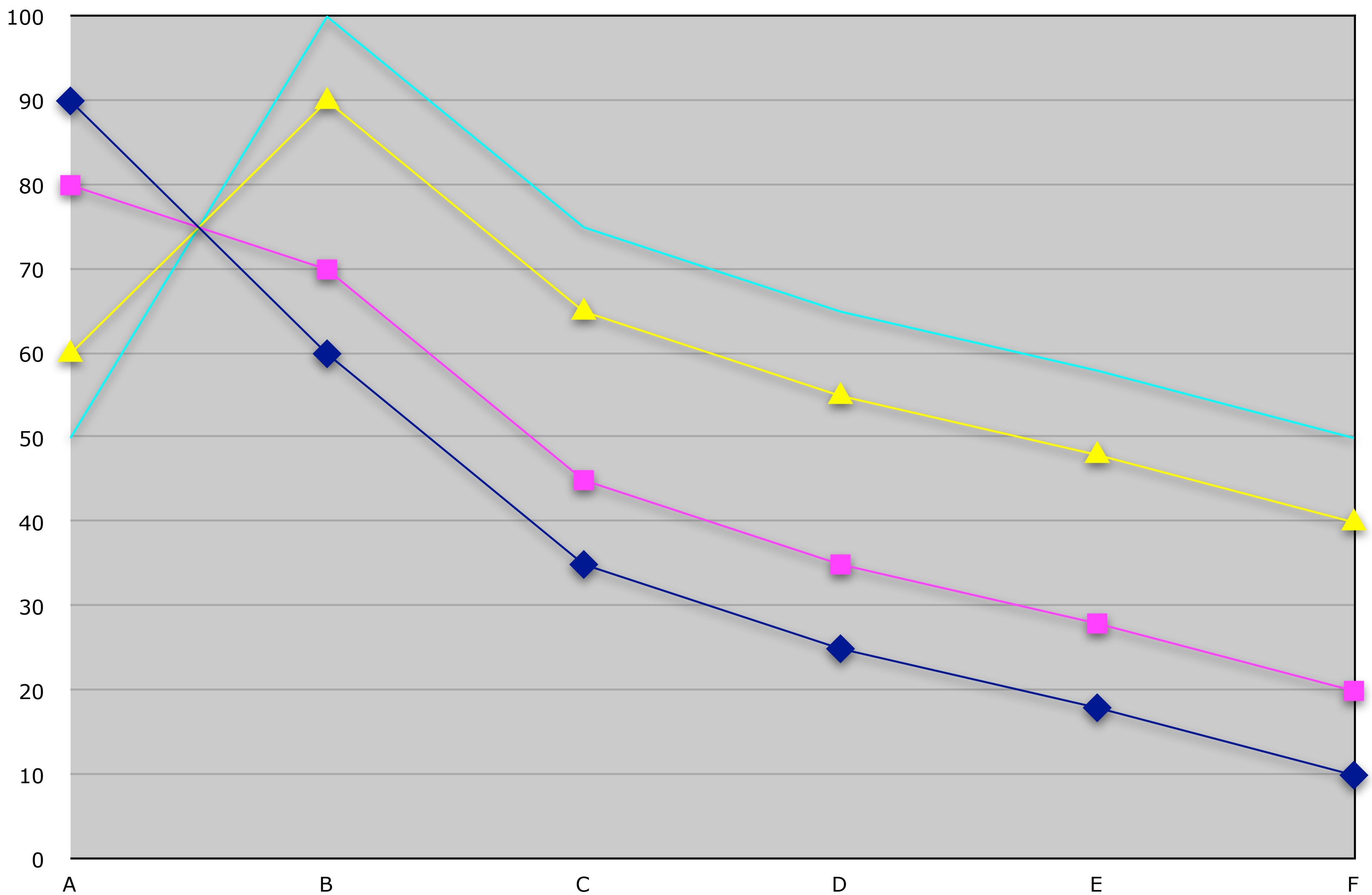
Alternative Joint scales



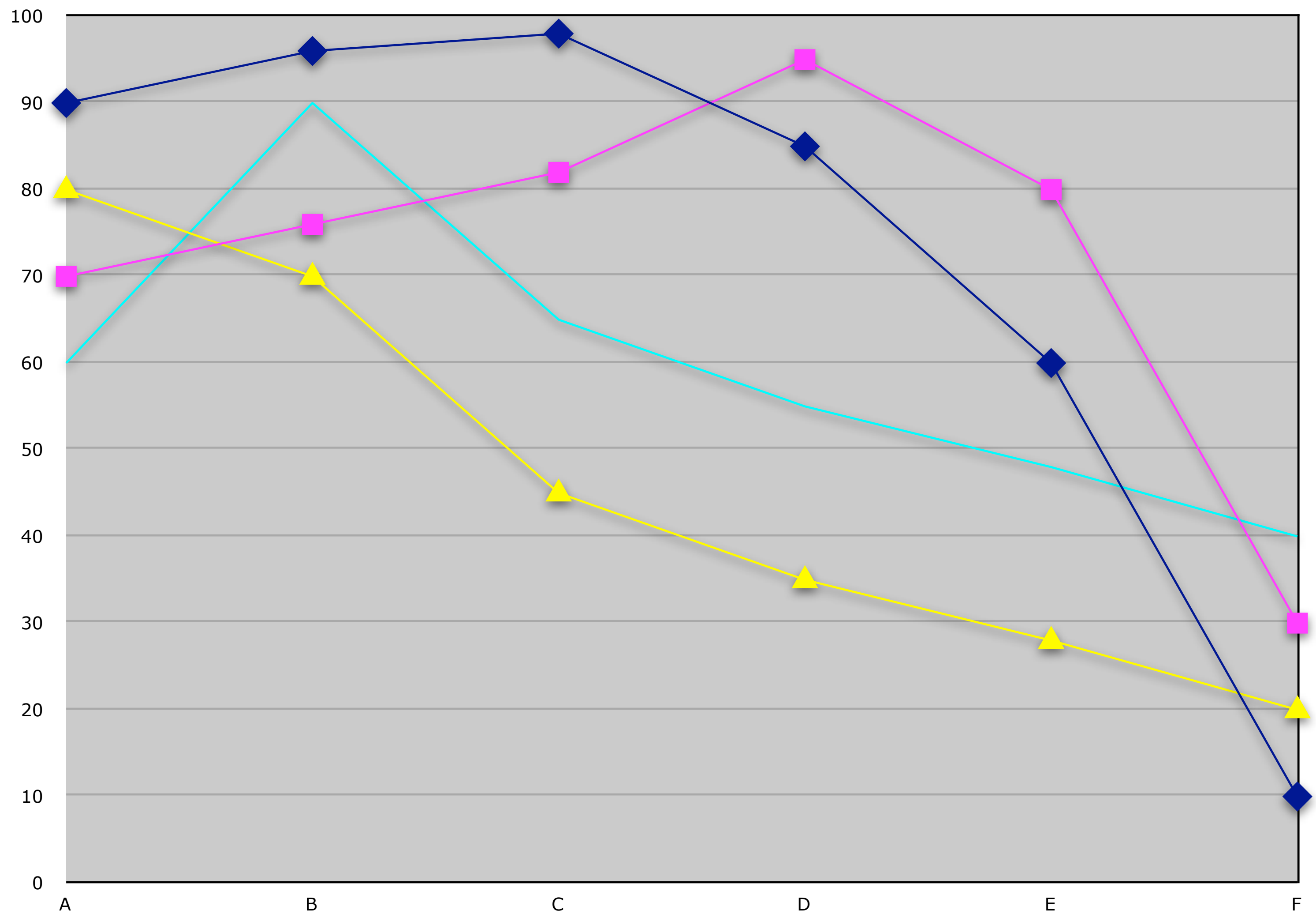
4 Individual scales from the accelerating Joint scale



Individual scales from the deaccelerating Joint scale



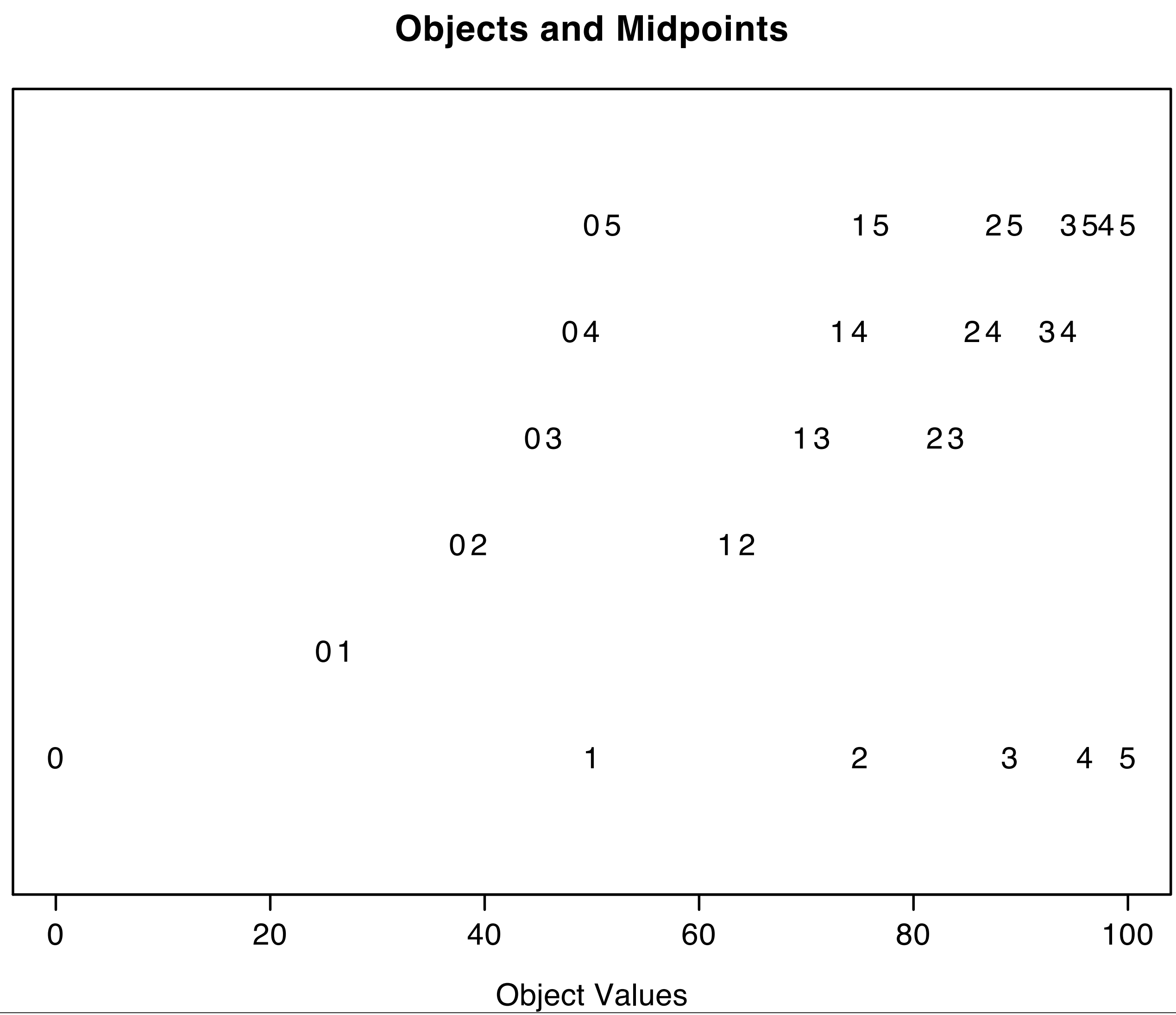
Individual scales from two Joint scales



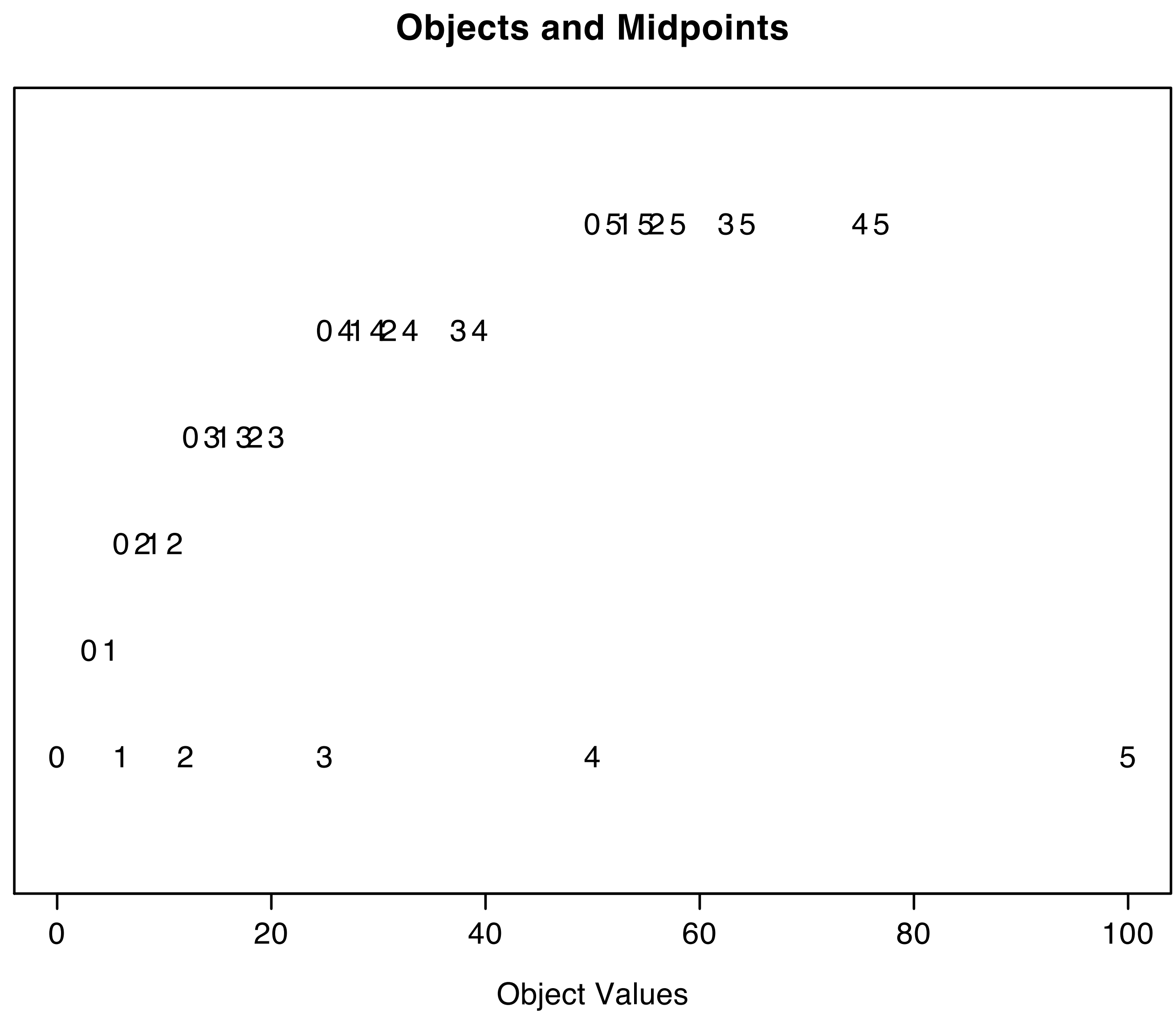
Unfolding of Preferences

- 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



Joint scales, Points and Midpoints
Accelerating scale



I scales and midpoints example 1

- Preference Orders: Midpoints crossed
- (Individual Scales)

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: Midpoints crossed
- (Individual Scales)

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

- Let x|y mean midpoint of x and y then the ordering of the midpoints provides information
- 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)

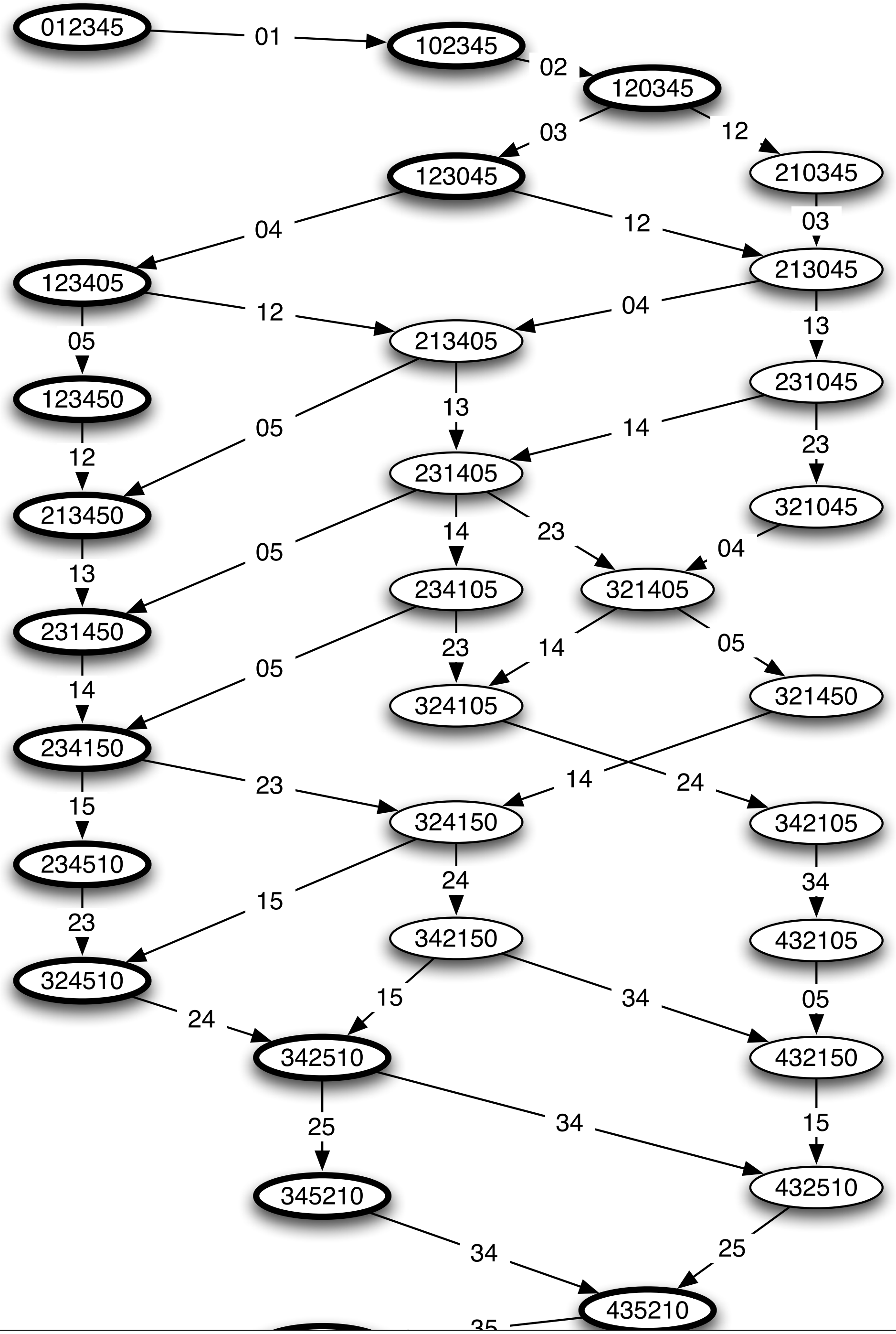
- Partial Orders of distances
- (04) > (03) > (02) > (12) > (34)
- (04) > (03) > (02) > (01) > (24) > (34)
- (04) > (03) > (02) > (01) > (24) > (23)

Family size Joint scale fitted from class data

- Alternative models:
  - Accelerating differences between children
  - De-accelerating differences
  - Equal spaced differences

405 data I scales and Midpoints

Count	I Scale	Midpoints "crossed"														
6	012345															
	102345	01														
3	210345	01	02	12												
2	213045	01	02	12	03											
3	213450	01	02	12	03			04		05						
	231045	01	02	12	03	13										
	231405	01	02	12	03	13		04								
	231450	01	02	12	03	13		04		05						
3	321045	01	02	12	03	13	23									
3	321450	01	02	12	03	13	23	04		05						
2	324150	01	02	12	03	13	23	04	14	05						
	324510	01	02	12	03	13	23	04	14	05		15				
	342150	01	02	12	03	13	23	04	14	05	24					
2	342510	01	02	12	03	13	23	04	14	05	24	15				
	342105	01	02	12	03	13	23	04	14		24					
	432105	01	02	12	03	13	23	04	14		24		34			
3	432150	01	02	12	03	13	23	04	14	05	24		34			
	453210	01	02	12	03	13	23	04	14	05	24	15	34	25	35	
2	543210	01	02	12	03	13	23	04	14	05	24	15	34	25	35	45



012345
01
102345
02
120345
03
12
123045
210345
03
04
12
123405
213045
12
04
05
213405
13
123450
231045
05
13
14
12
231405
23
213450
321045
05
14
23
04
13
234105
321405
231450
05
23
14
05
14
324105
321450
234150
14
24
23
15
324150
342105
234510
15
24
34
23
342150
432105
324510
15
34
05
24
342510
432150
25
34
15
345210
432510
25
34
435210
35

Partial metric information: Midpoint orders implies order of distance

Midpoint order			distance
12	03	<=>	(01)<(23)
23	04	<=>	(02)<(34)
24	15	<=>	(12)<(45)
34	25	<=>	(23)<(45)
05	24	<=>	(45)<(02)
24	15	<=>	(12)<(45)

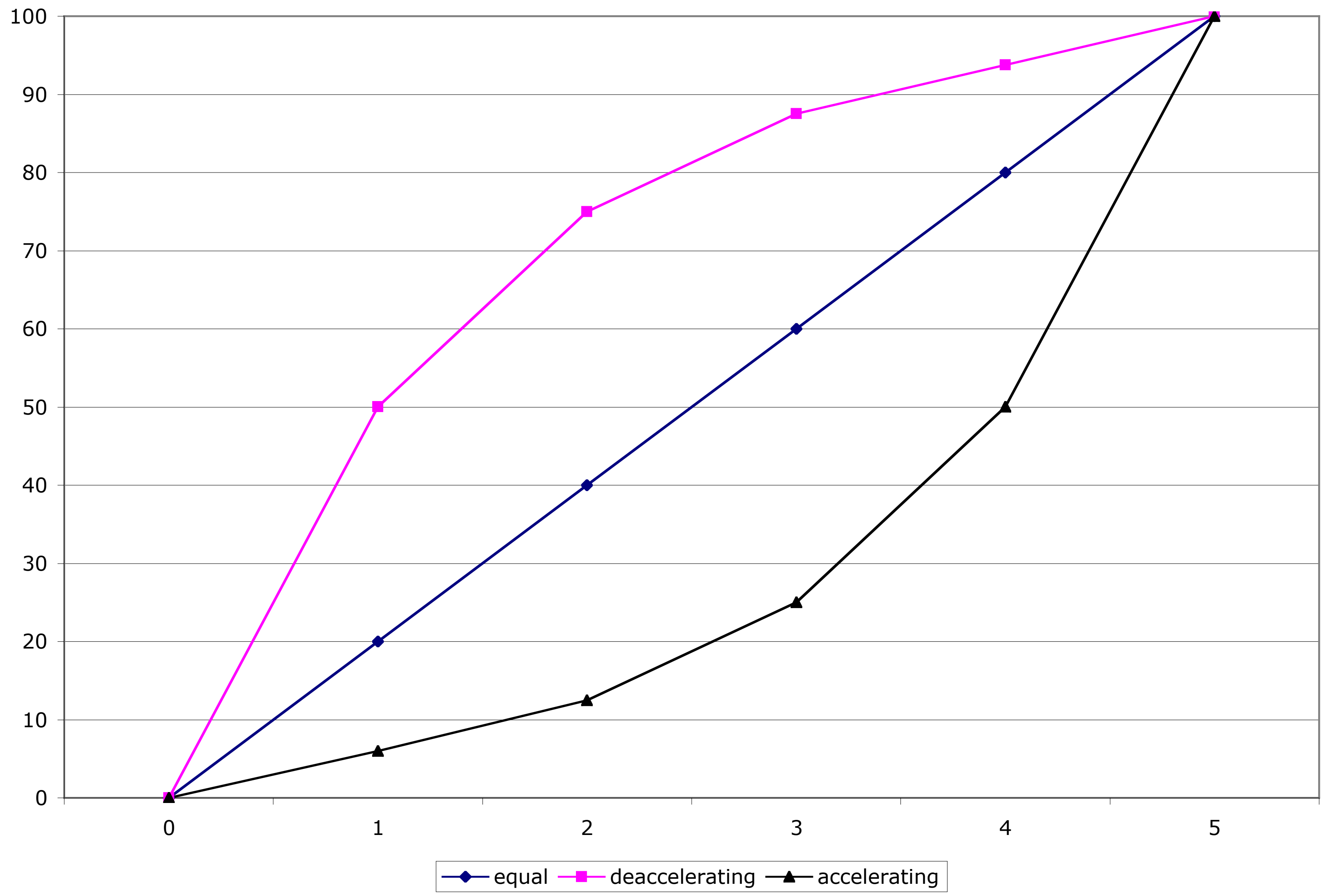
Partial orders

(01)<(23)	(23)<(45)	(45)<(02)	(02)<(34)
(01)<(02)	(02)<(34)		
(01)<(23)<(45)<(02)<(34)			
(12)<(45)<(02)<(34)			

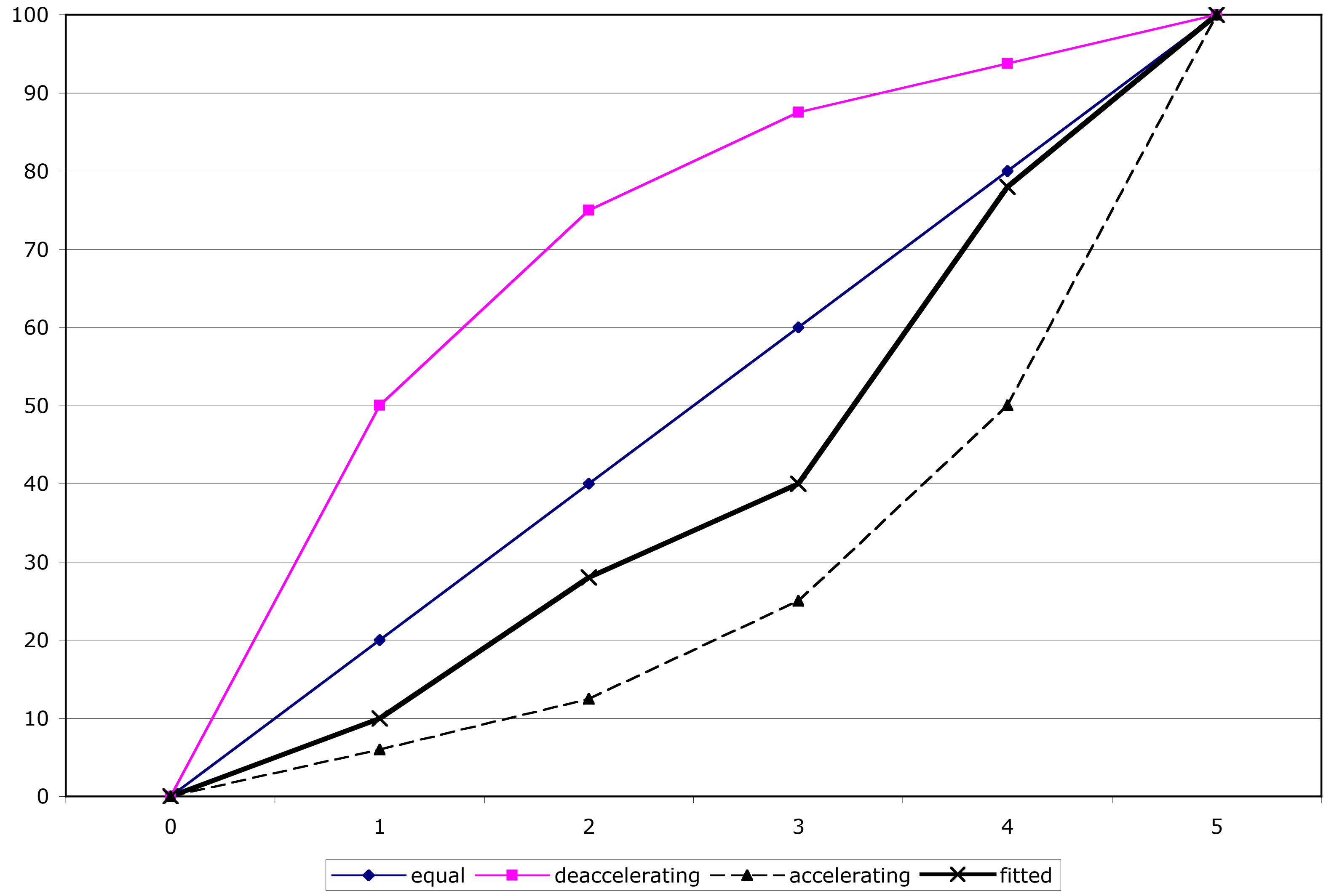
Distances as deltas

0	1	2	3	4	5
	a+b	b+c+d	a+b+c	a+2b+c+d+e	a+b+c+d

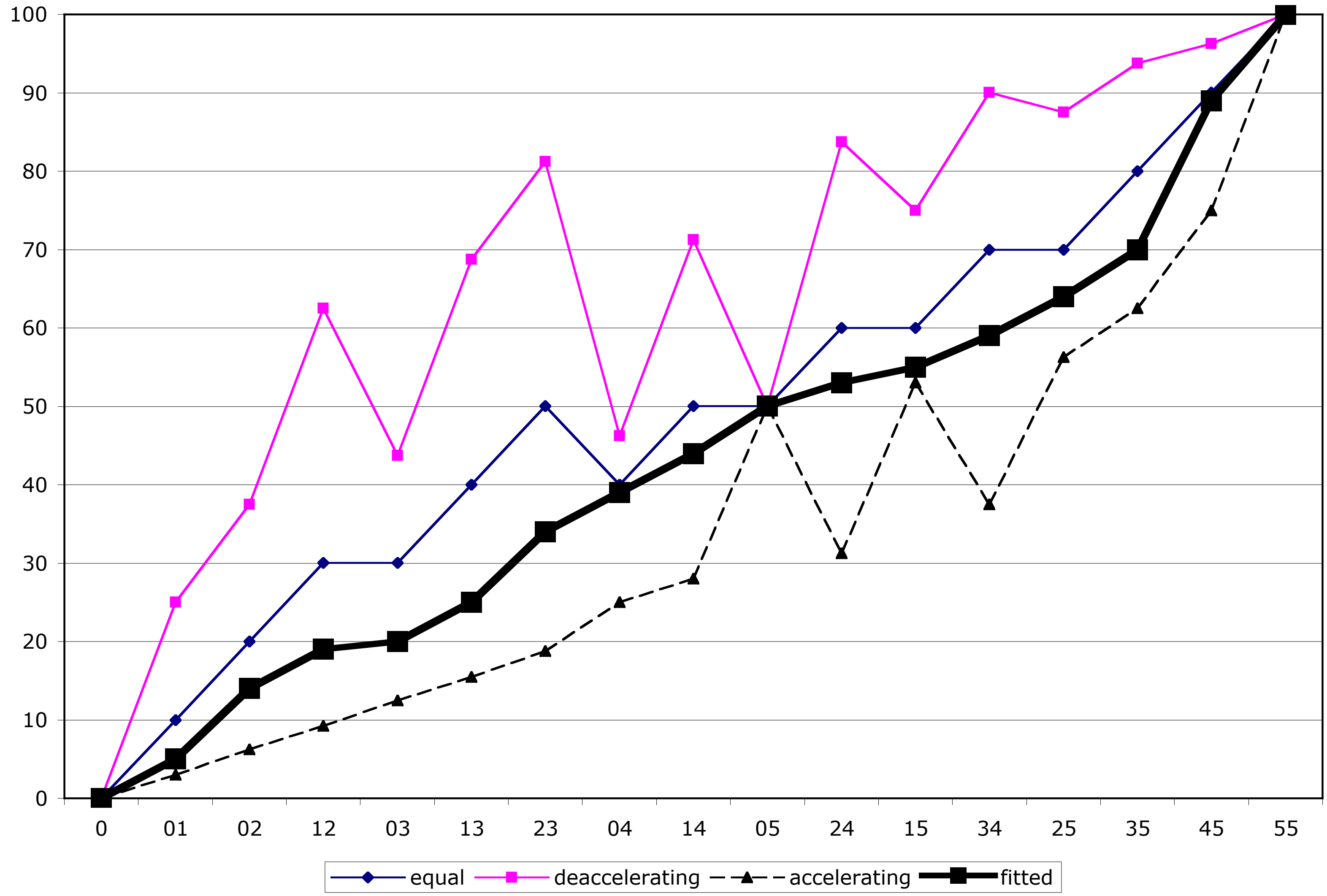
3 a priori models



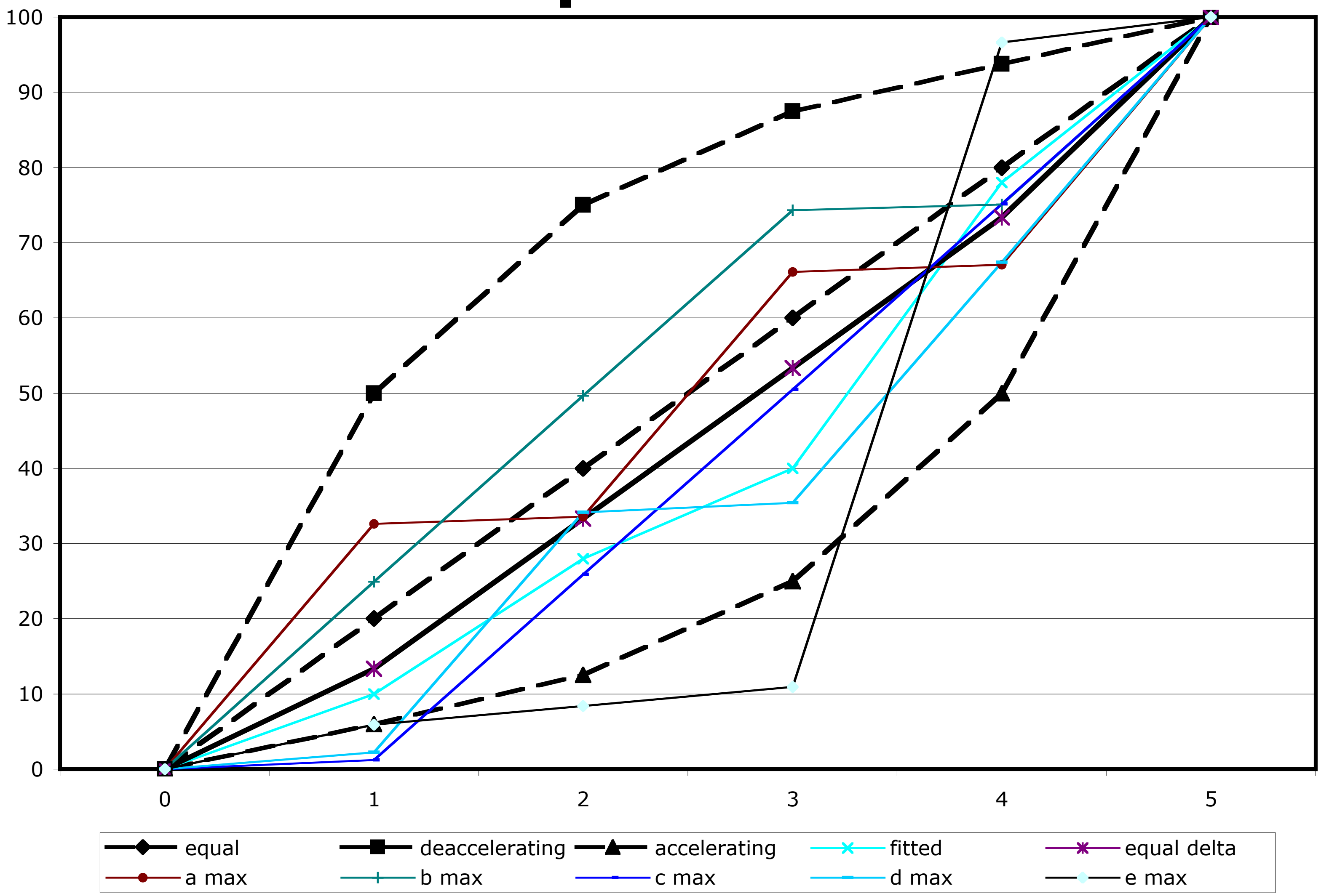
3 a priori + 1 fitted model



3 prior + 1 fitted models and implied midpoint orders



The range of acceptable fits versus a priori models



Measurement (S * O)

- Ordering of abilities: $s_i < o_j$

  Is a subject less than an object (i.e. does the subject miss the item).

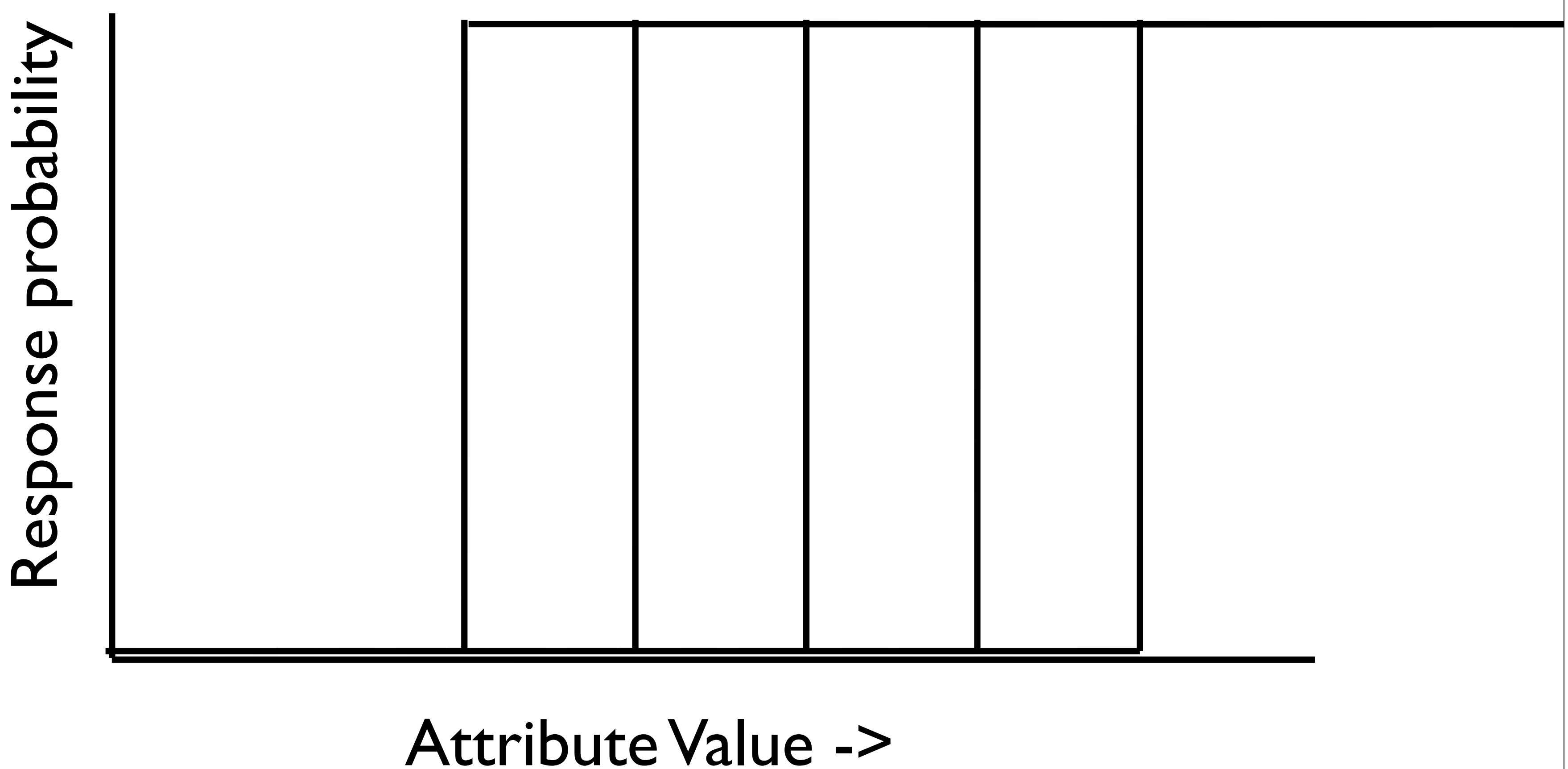
  Order the items in terms of difficulty, and subjects in terms of ability.

  example: high jump or cognitive ability test

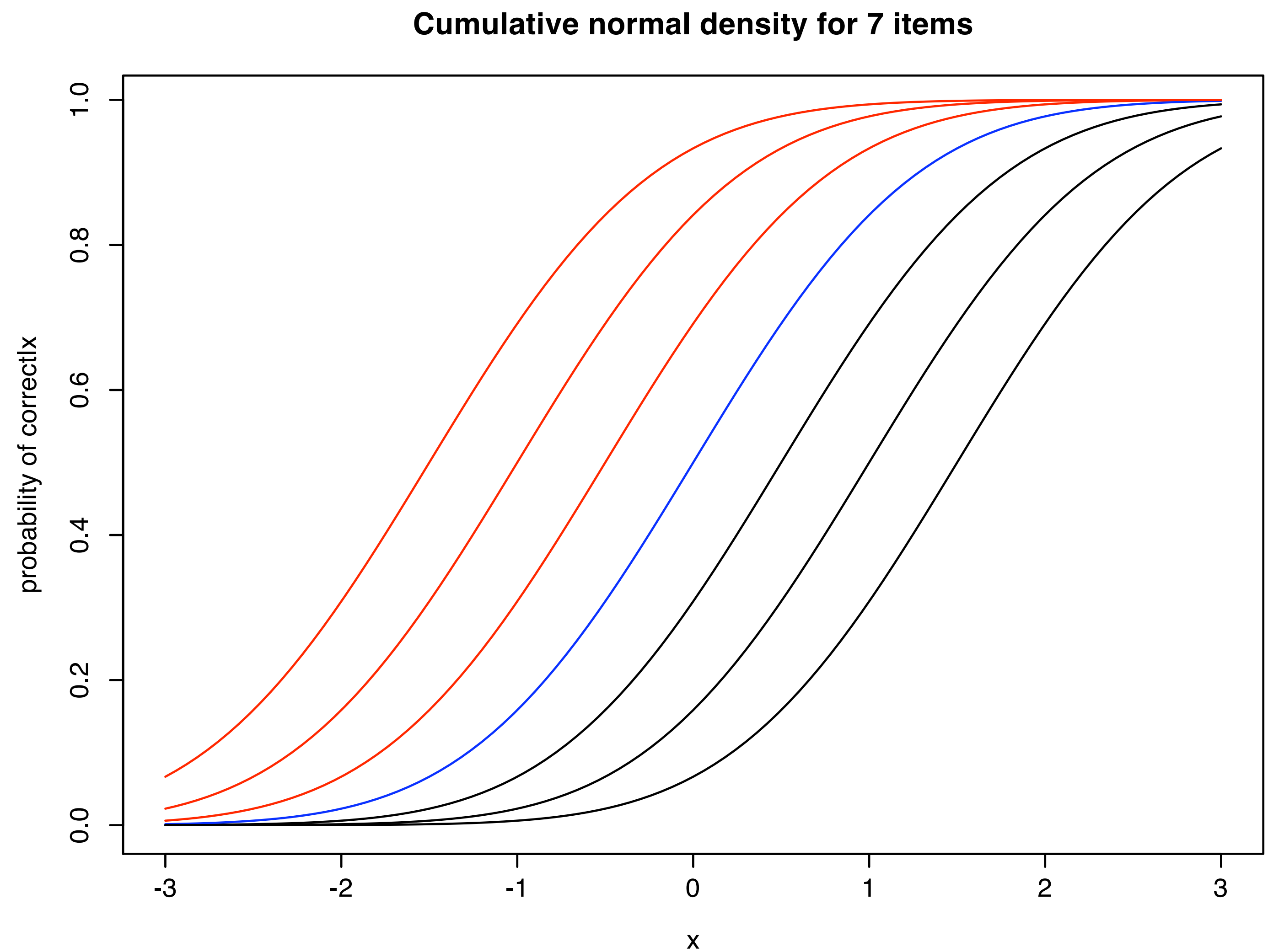
- Proximity of attitudes $|s_i - o_j| < d$

  Subject agrees (endorses) an item if d < some threshold
  Subject rejects the item if d > threshold

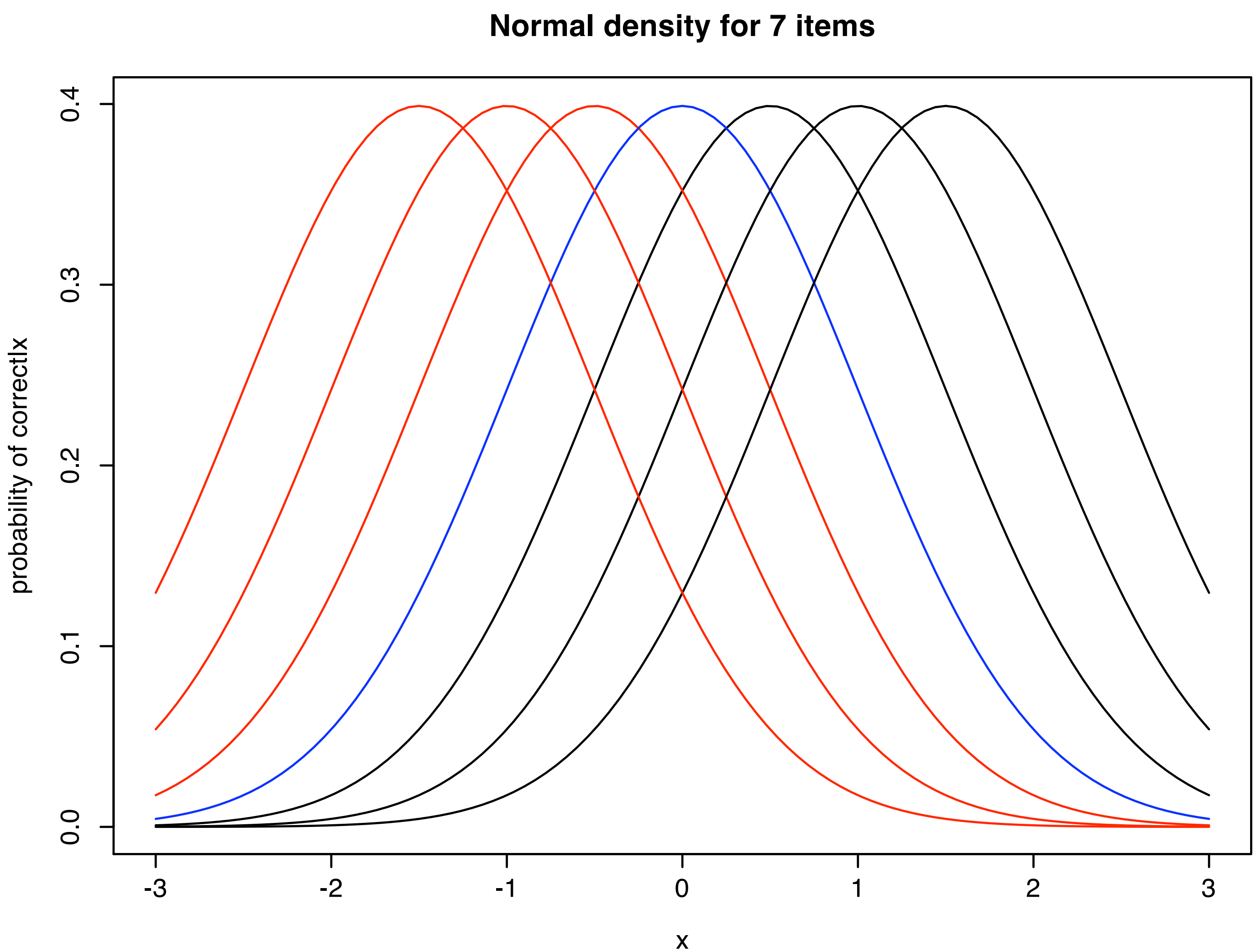
Error free models of ability: The Guttman scale



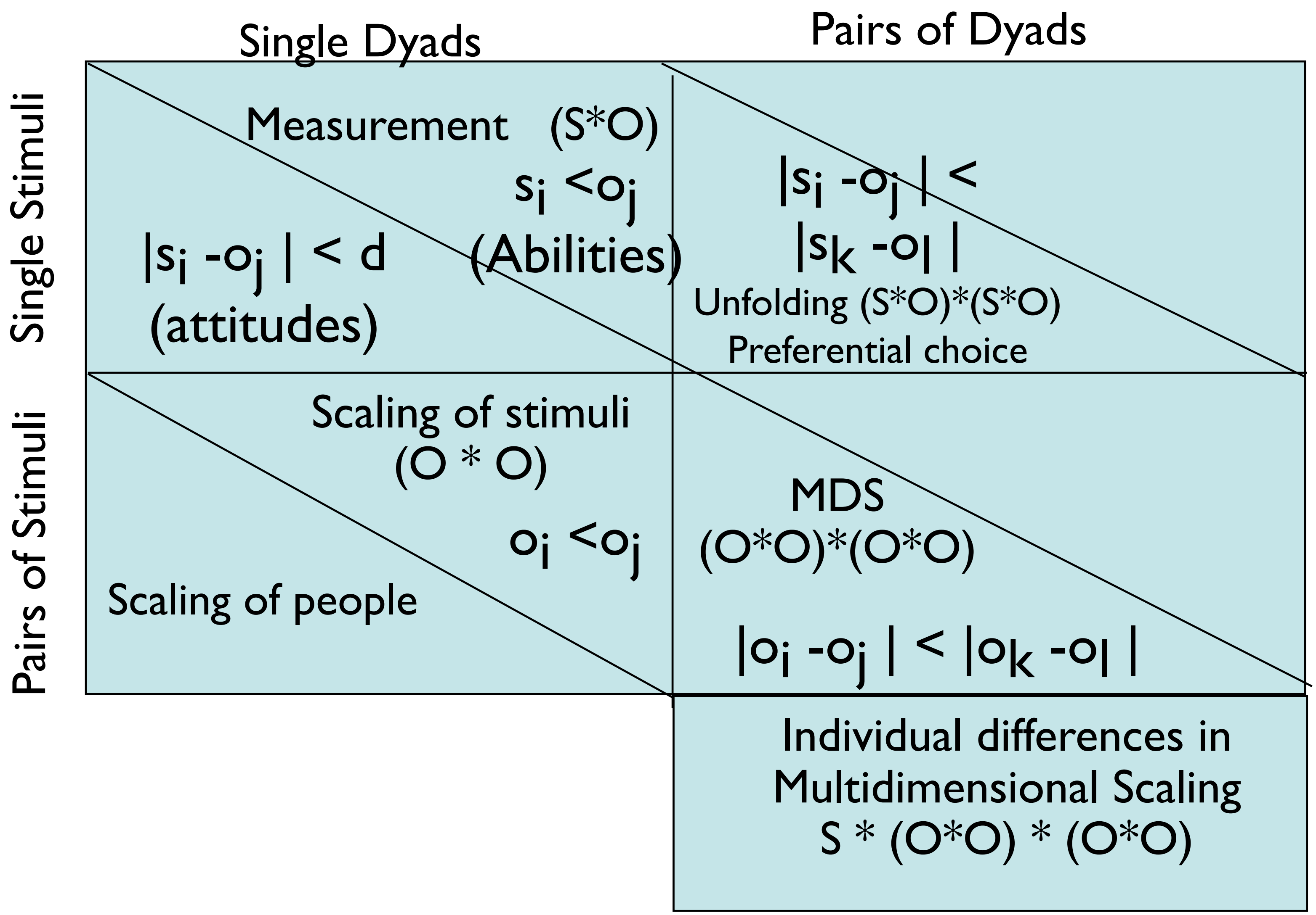
Error models of ability: Normal Ogive/Logistic of order



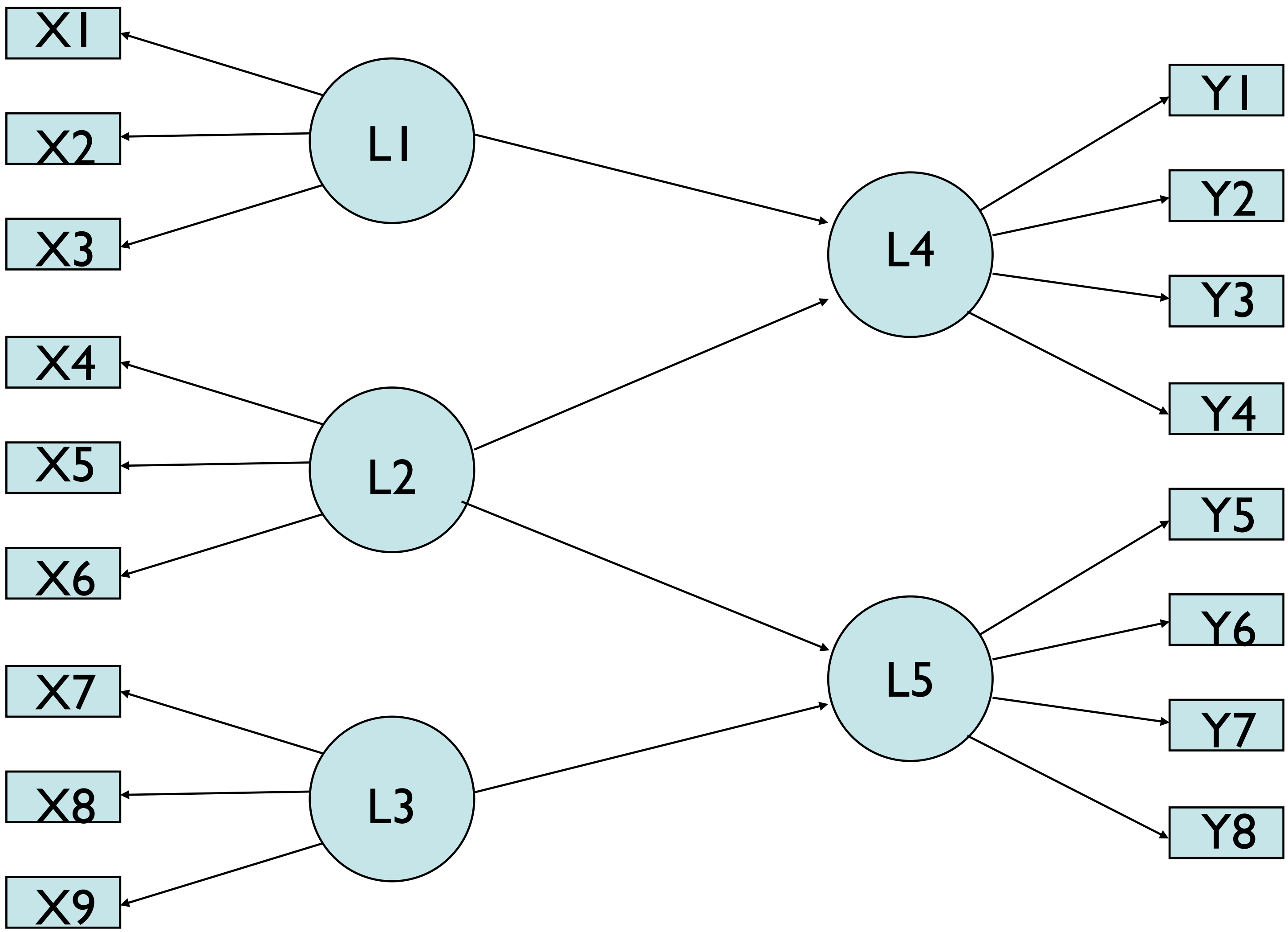
Measuring Attitudes: distance from ideal point => unfolding



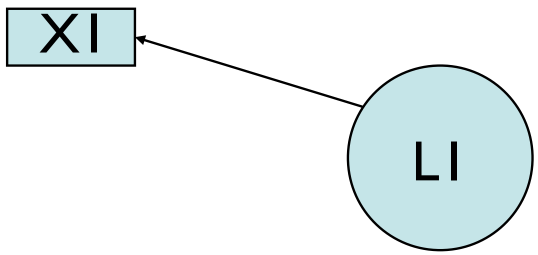
Coombs typology of data



Psychometric Theory: A conceptual Syllabus



Measurement and scaling



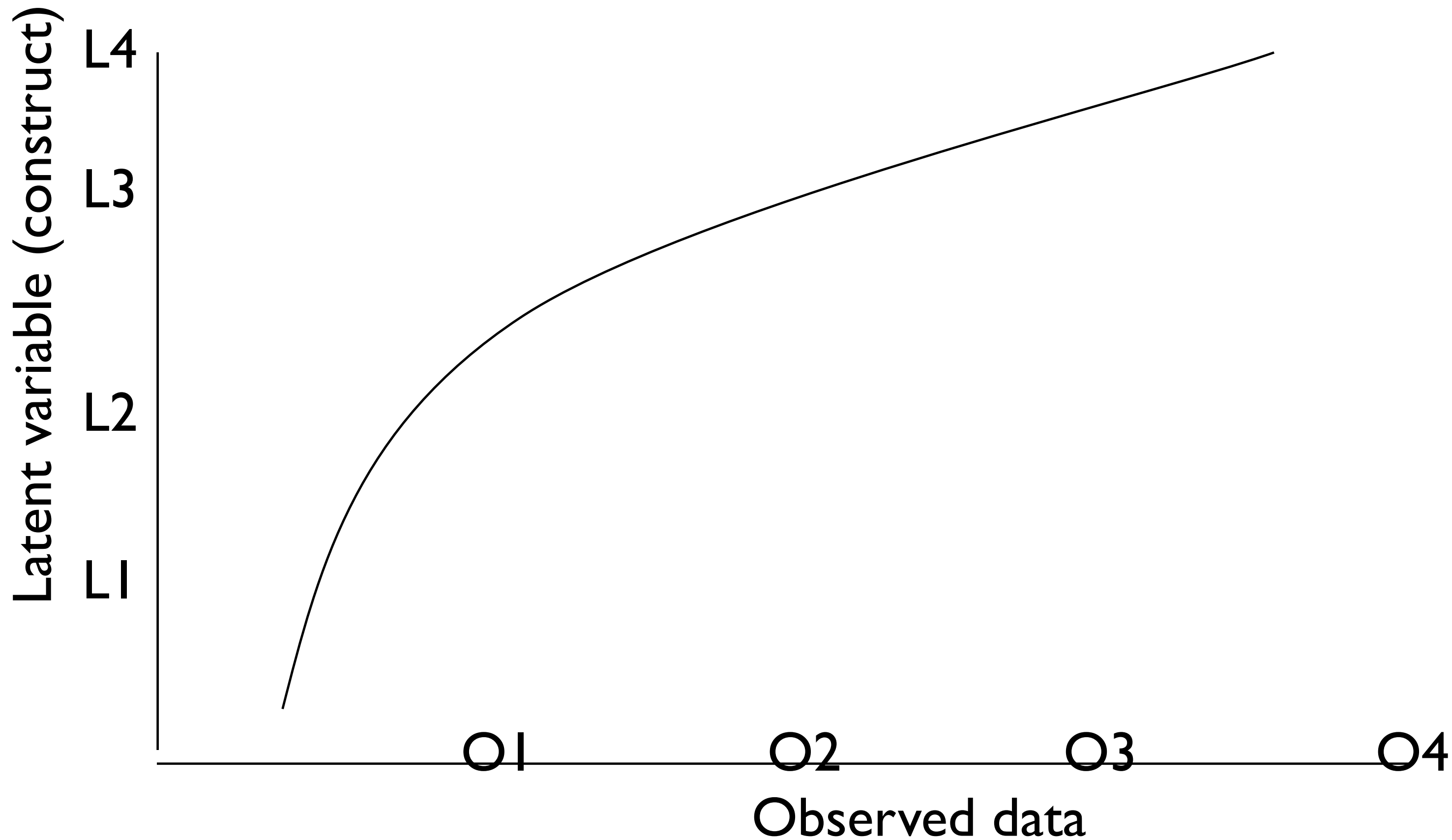
Inferring latent values from observed values

Types of Scales: Inferences from observed variables to Latent variables

- Nominal
- Ordinal
- Interval
- Ratio

- Categories
- Ranks (x > y)
- Differences
  - X-Y > W-V
- Equal intervals with a zero point =>
  - X/Y > W/V

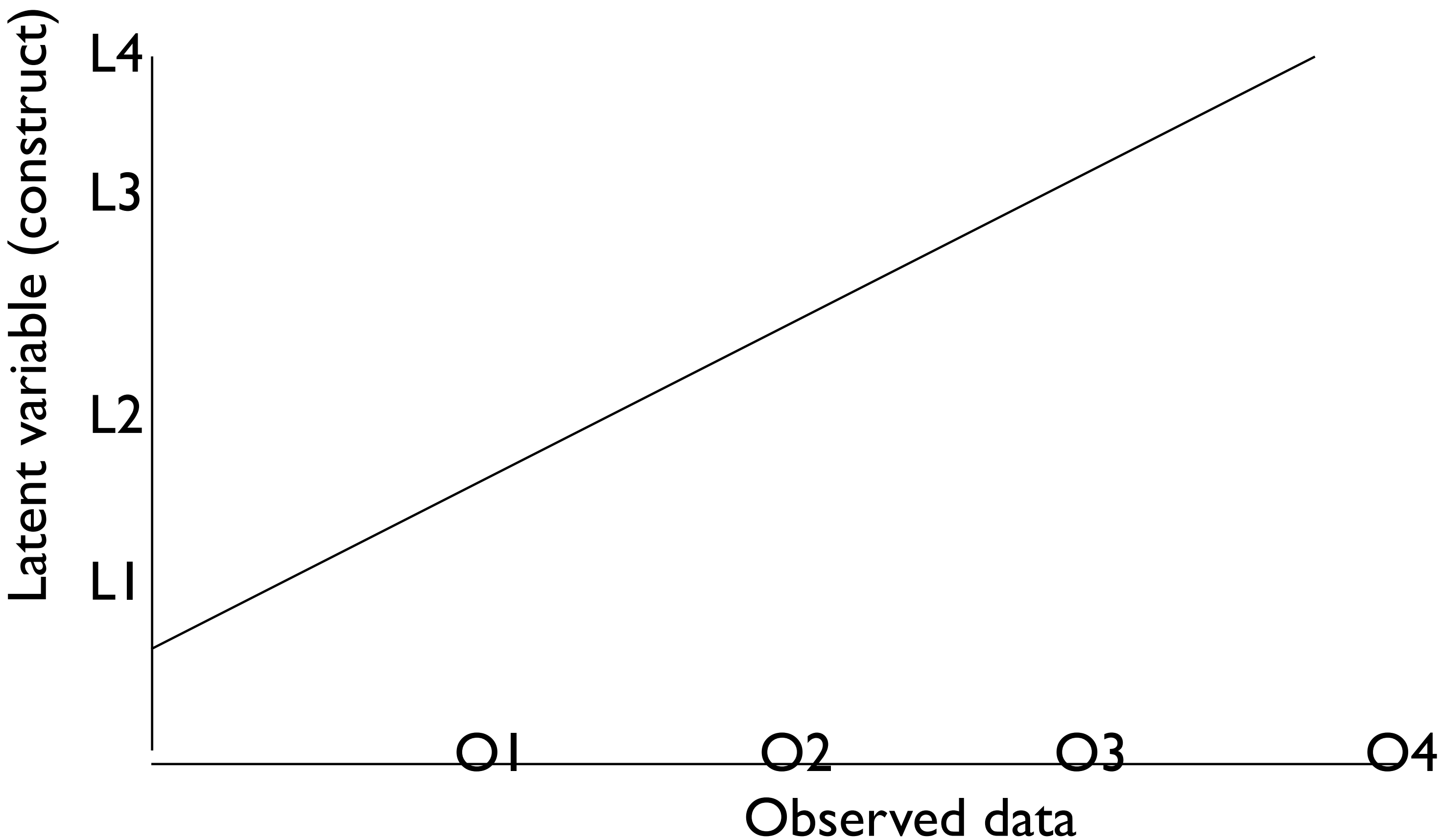
Mappings and inferences



Ordinal Scales

- Any monotonic transformation will preserve order
- Inferences from observed to latent variable are restricted to rank orders
- Statistics: Medians, Quartiles, Percentiles

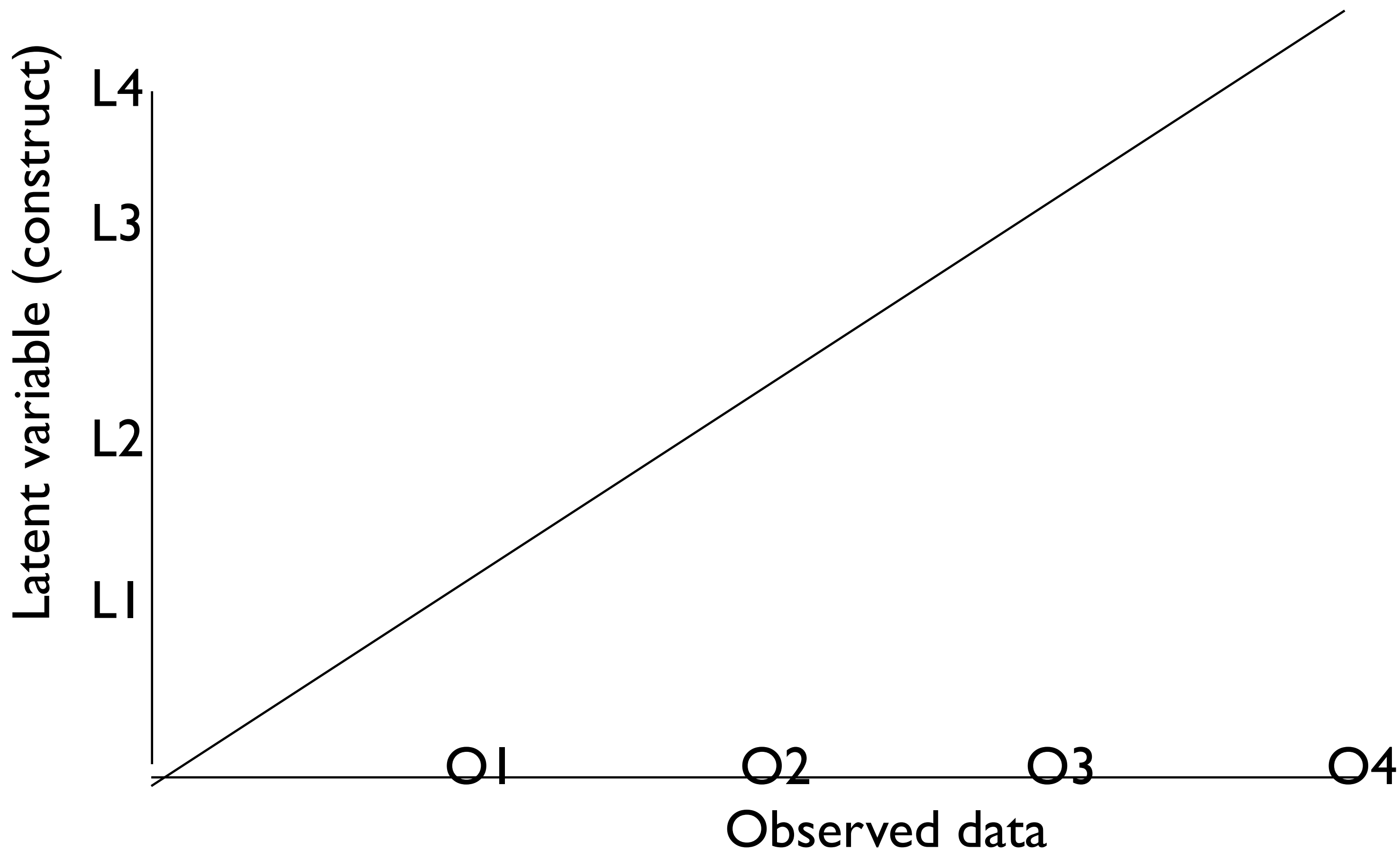
Mappings and inferences



Interval Scales

- Possible to infer the magnitude of differences between points on the latent variable given differences on the observed variable
  X is as much greater than Y as Z is from W
- Linear transformations preserve interval information
- Allowable statistics: Means, Variances

Mappings and inferences



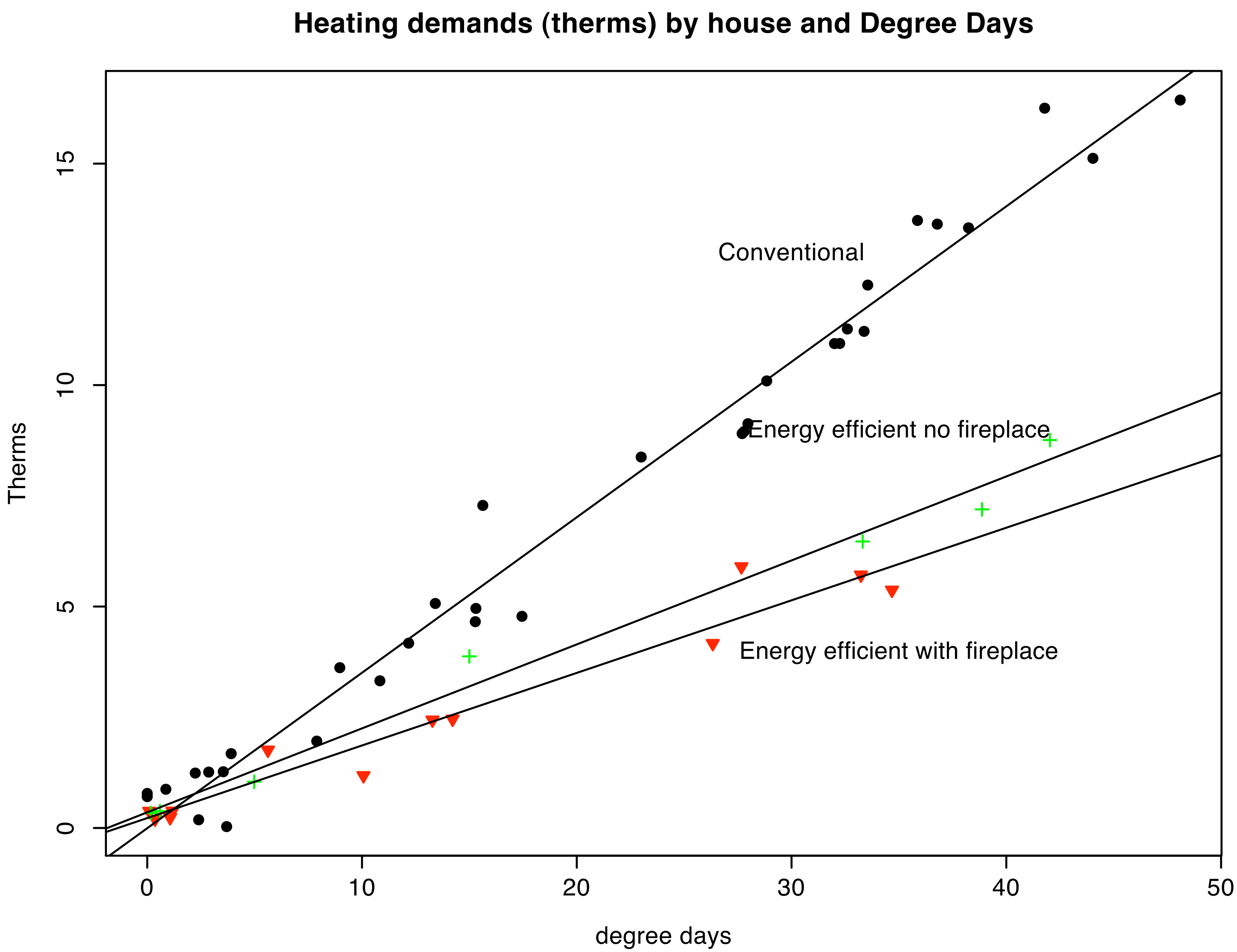
Ratio Scales

- Interval scales with a zero point
- Possible to compare ratios of magnitudes (X is twice as long as Y)

The search for appropriate scale

- Is today colder than yesterday? (ranks)
- Is the amount that today is colder than yesterday more than the amount that yesterday was colder than the day before? (intervals)
  - 50 F - 39 F < 68 F - 50 F
  - 10 C - 4 C < 20 C - 10 C
  - 283K - 277K < 293K -283K
- How much colder is today than yesterday?
  - (Degree days as measure of energy use)
  - K as measure of molecular energy

Gas consumption by degree days (65-T)



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.

Effect of teaching upon performance

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.

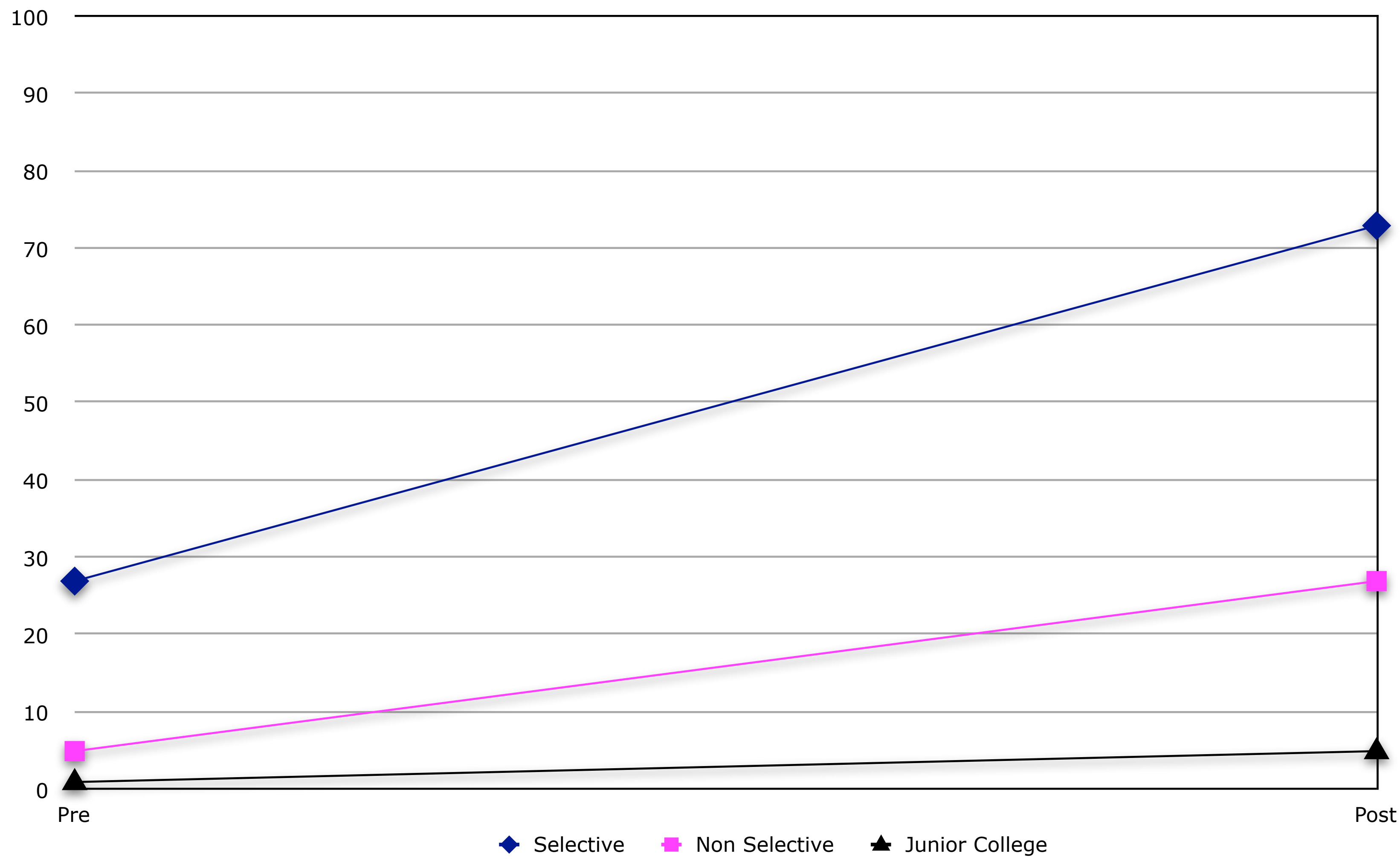
A writing test was given to the entering students at three institutions in the Boston area. After one year, a similar writing test was given again. Although there was some attrition from each sample, the researchers report data only for those who finished one year. The pre and post test scores as well as the change scores were as shown below:

Effect of teaching upon performance

	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.

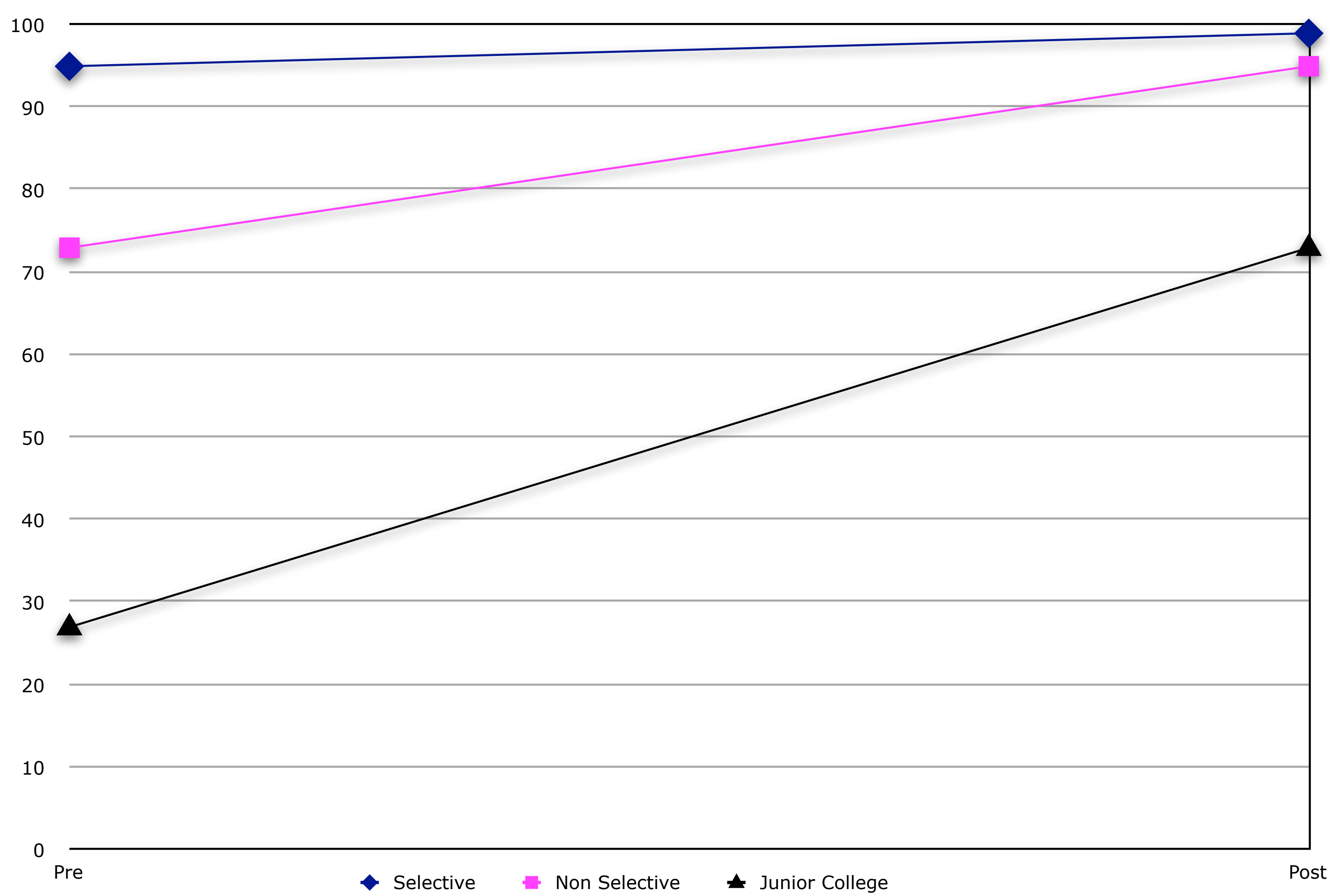
Effect of Teaching upon Performance?



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:

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

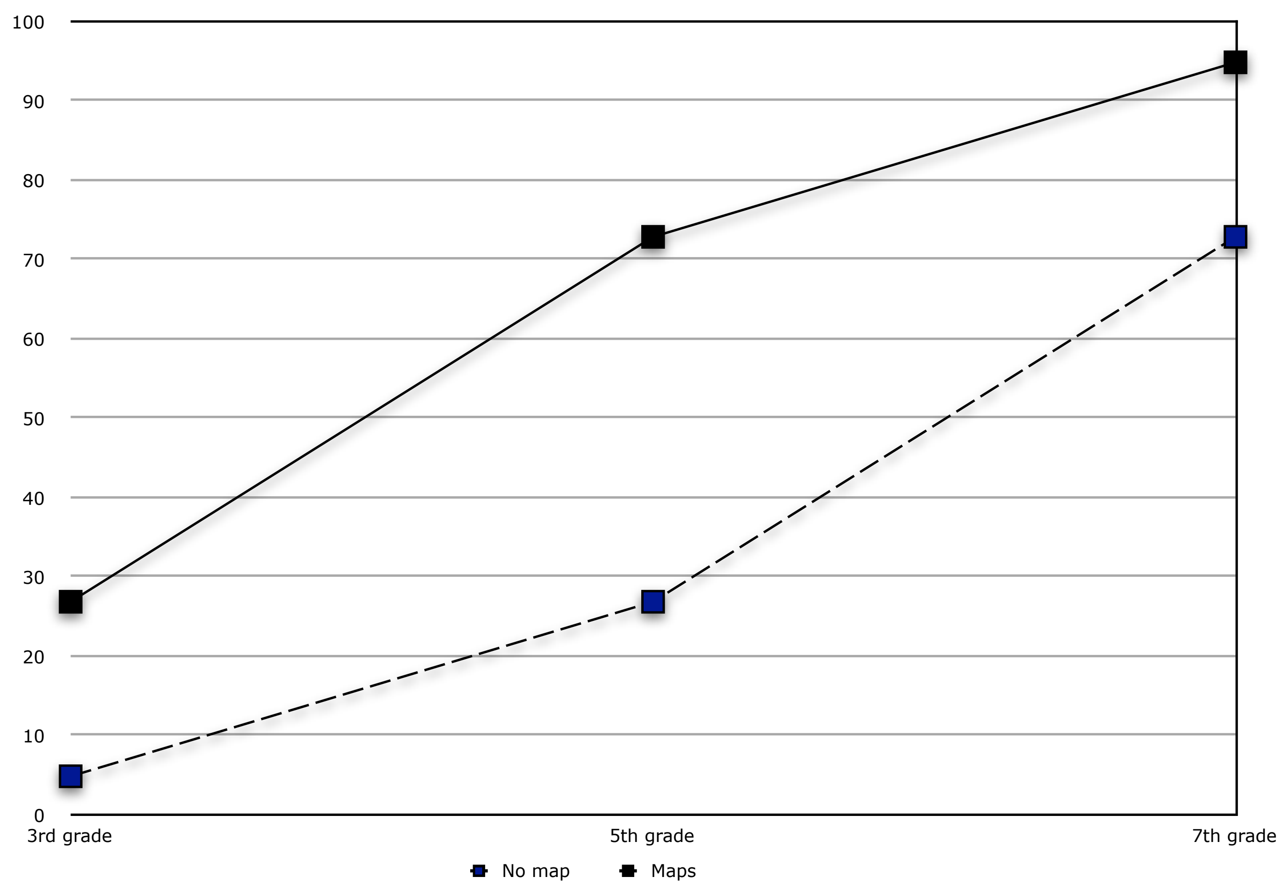
Effect of Teaching upon



A leading cognitive developmentalist believed that there is a critical stage for learning spatial representations using maps. Children younger than this stage are not helped by maps, nor are children older than this stage. He randomly assigned 3rd, 5th, and 7th grade students into two conditions (nested within grade), control and map use. Performance was measured on a task of spatial recall (children were shown toys at particular locations in a set of rooms and then asked to find them again later. Half the children were shown a map of the rooms before doing the task.

	No map	Maps
3rd grade	5	27
5th grade	27	73
7th grade	73	95

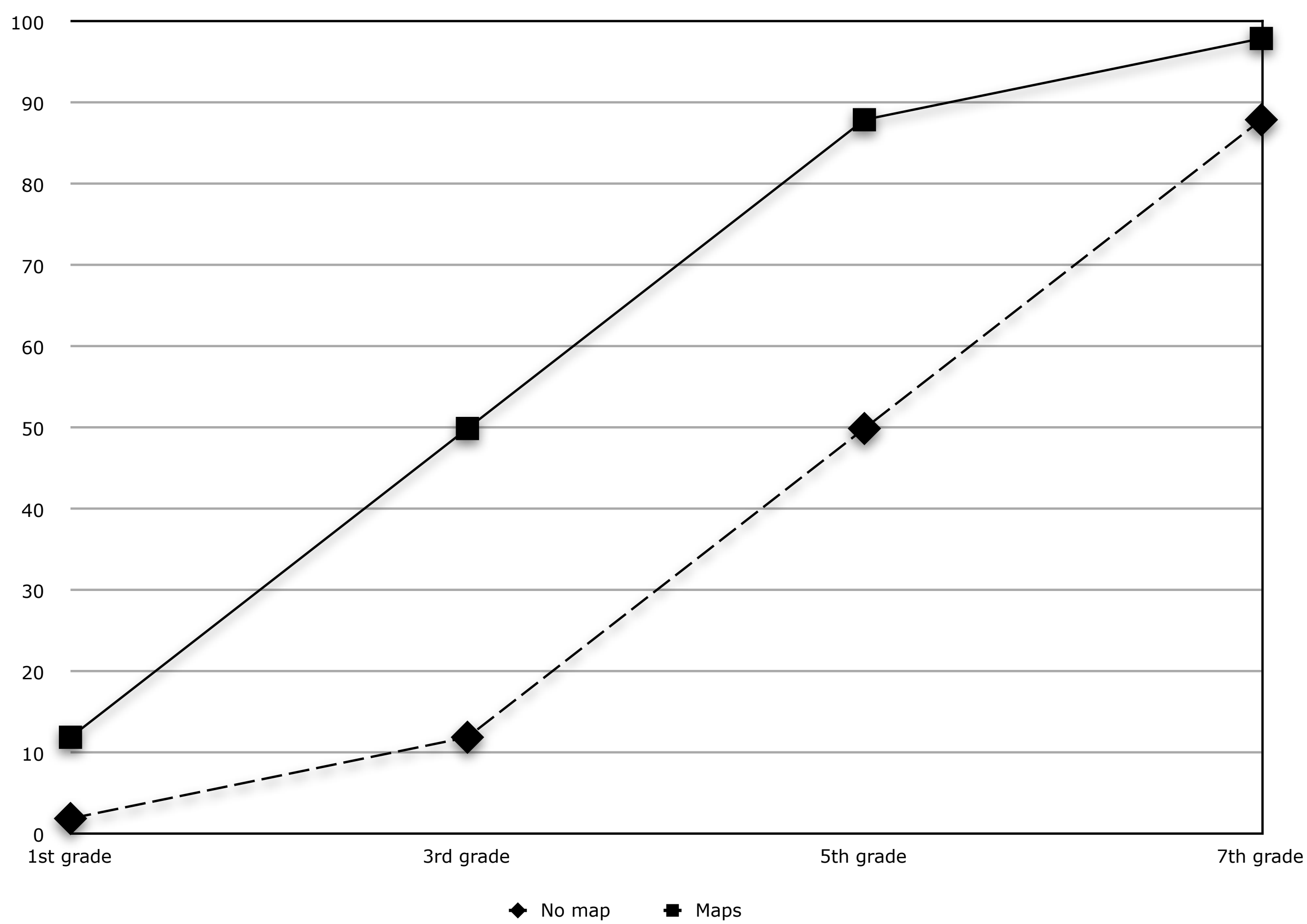
Spatial reasoning facilitated by maps at a critical age



Another cognitive developmentalist believed that there is a critical stage but that it appears earlier than previously thought. Children younger than this stage are not helped by maps, nor are children older than this stage. He randomly assigned 1st, 3rd, 5th, and 7th grade students into two conditions (nested within grade), control and map use. Performance was measured on a task of spatial recall (children were shown toys at particular locations in a set of rooms and then asked to find them again later. Half the children were shown a map of the rooms before doing the task.

	No map	Maps
1st grade	2	12
3rd grade	12	50
5th grade	50	88
7th grade	88	98

Spatial Reasoning is facilitated by map use at a

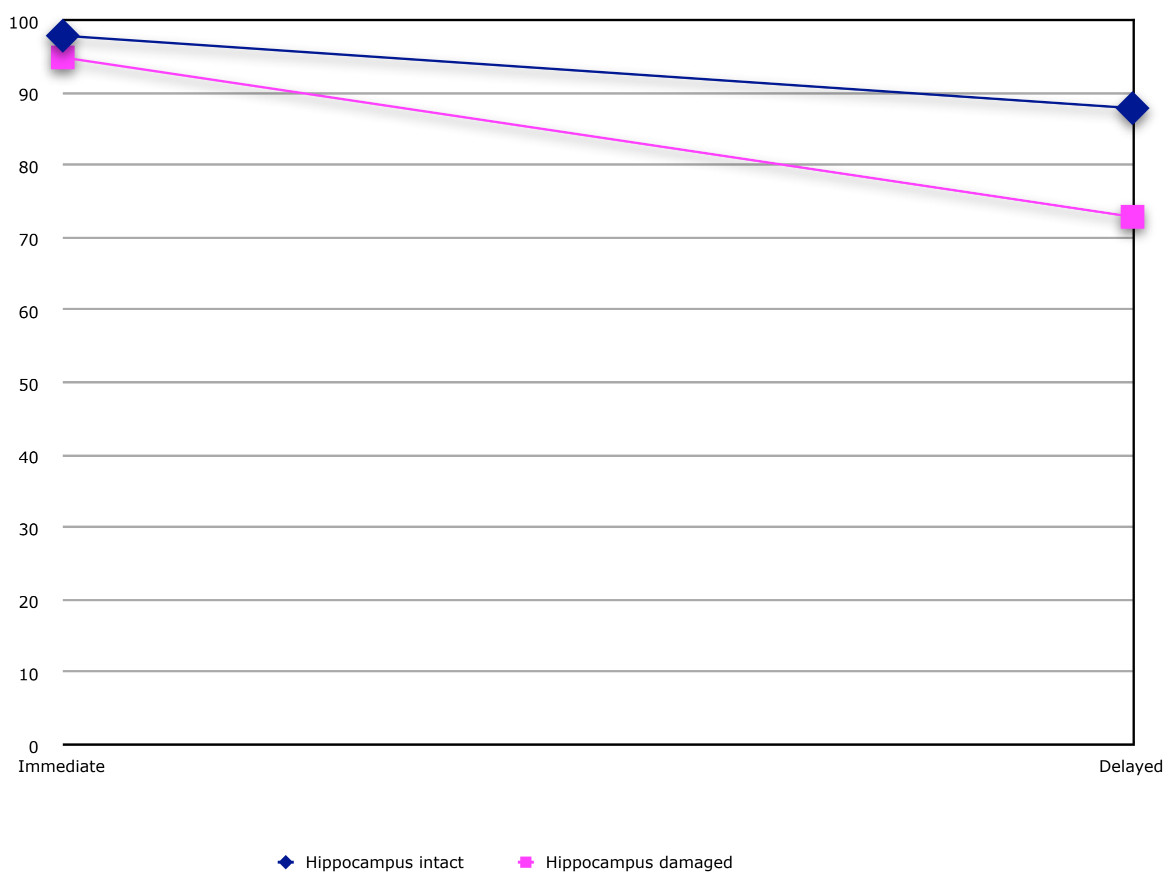


Cognitive-neuro psychologists believe that damage to the hippocampus affects long term but not immediate memory. As a test of this hypothesis, an experiment is done in which subjects with and without hippocampal damage are given an immediate and a delayed memory task. The results are impressive:

	Immediate	Delayed
Hippocampus intact	98	88
Hippocampus damaged	95	73

From these results the investigator concludes that there are much larger deficits for the hippocampal damaged subjects on the delayed rather than the immediate task. The investigator believes these results confirm his hypothesis. Comment on the appropriateness of this conclusion.

Memory = f(hippocampal damange * temporal delay)

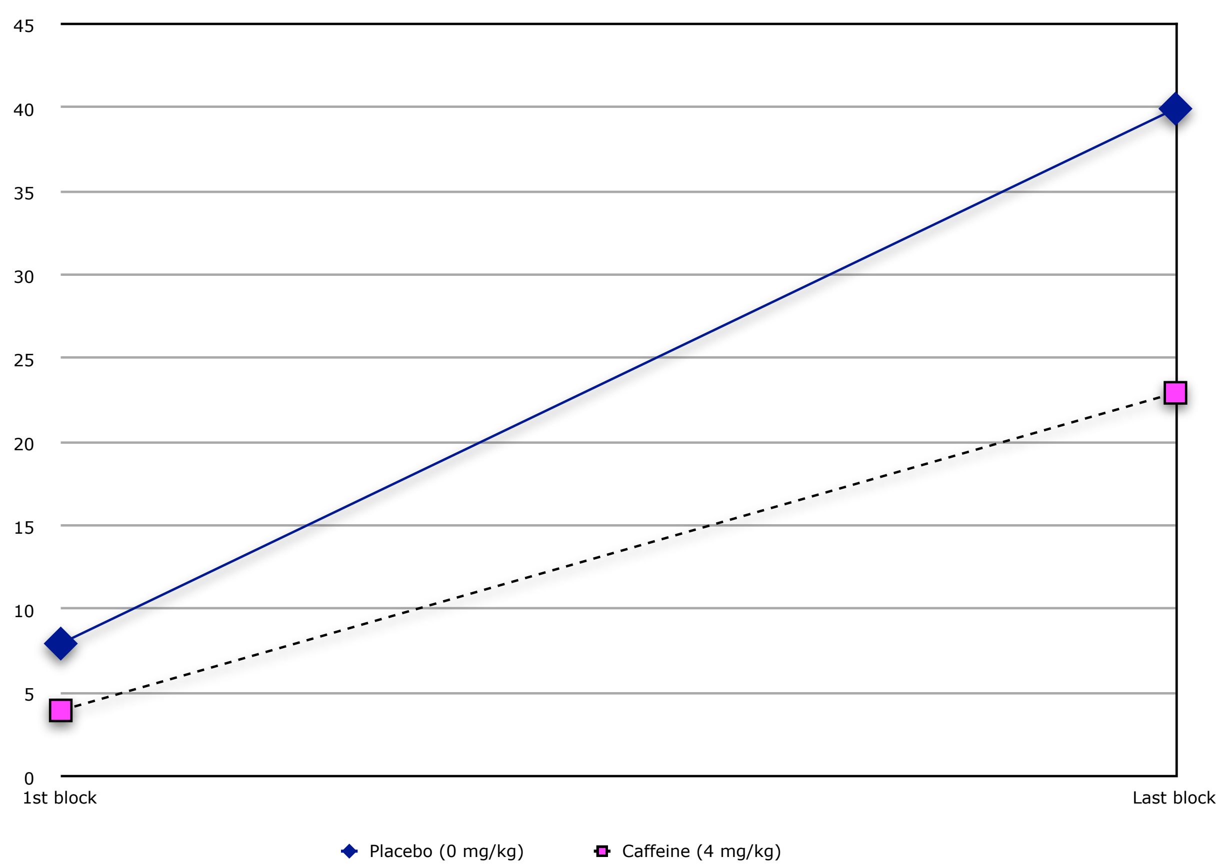


Errors=f(caffeine * time on task)

An investigator believes that caffeine facilitates attentional tasks such that require vigilance. Subjects are randomly assigned to conditions and receive either 0 or 4mg/kg caffeine and then do a vigilance task. Errors are recorded during the first 5 minutes and the last 5 minutes of the 60 minute task. The number of errors increases as the task progresses but this difference is not significant for the caffeine condition and is for the placebo condition.

	1^{st} block	Last block
Placebo (0 mg/kg)	8	40
Caffeine (4 mg/kg)	4	23

Errors=f(caffeine * time on task)



Measuring Arousal

Arousal is a fundamental concept in many psychological theories. It is thought to reflect basic levels of alertness and preparedness. Typical indices of arousal are measures of the amount of palmer sweating. This may be indexed by the amount of electricity that is conducted by the fingertips. Alternatively, it may be indexed (negatively) by the amount of skin resistance of the finger tips. The Galvanic Skin Response (GSR) reflects moment to moment changes, SC and SR reflect longer term, basal levels.

High skin conductance (low skin resistance) is thought to reflect high arousal.

Measuring Arousal

Anxiety is thought to be related to arousal. The following data were collected by two different experimenters. One collected Resistance data, one conductance data.

	Resistance	Conductance
Anxious	2, 2	.5, .5
Low anx	1, 5	1, .2

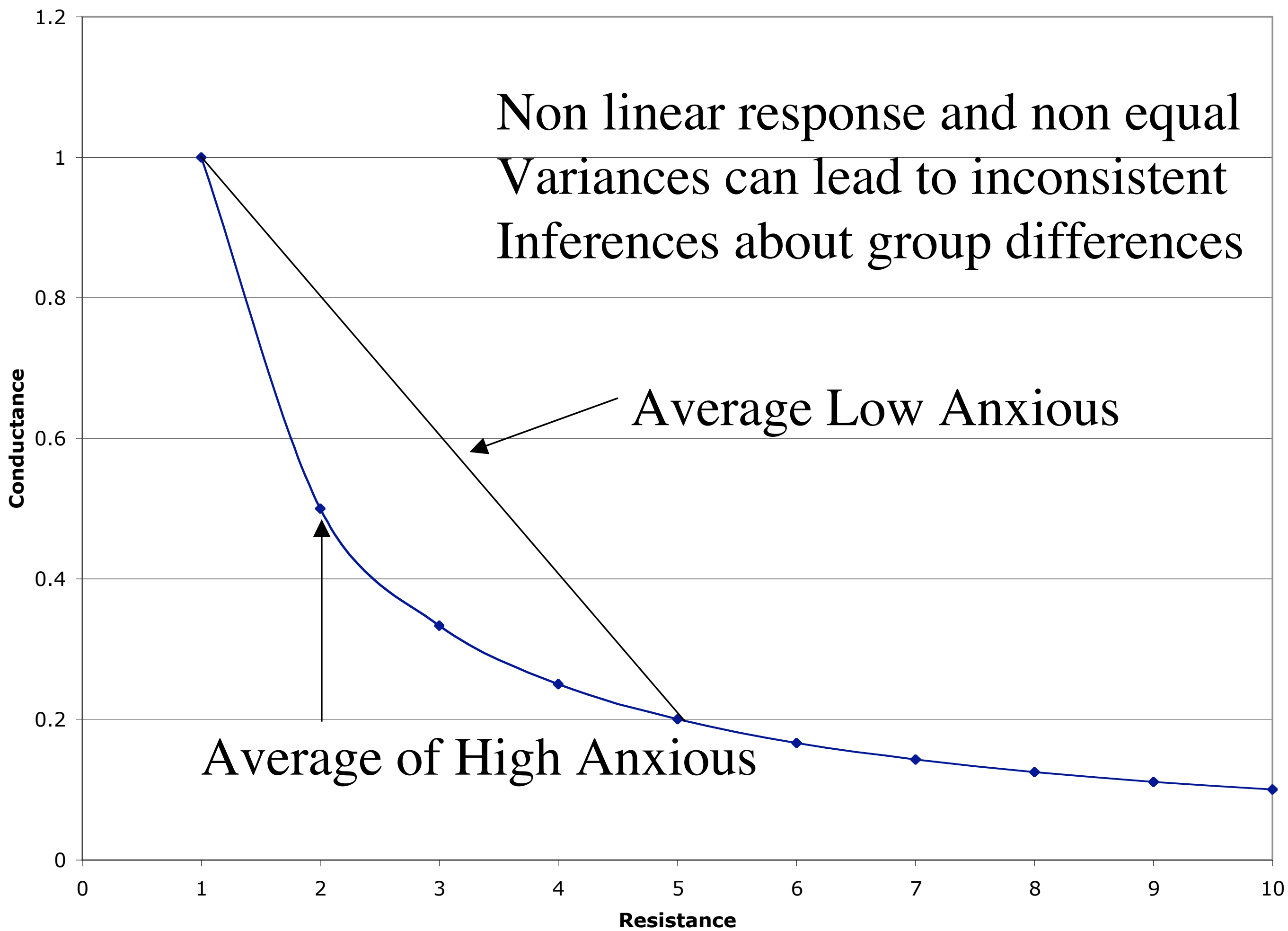
The means were

	Resistance	Conductance
Anxious	2	.5
Low anx	3	.6

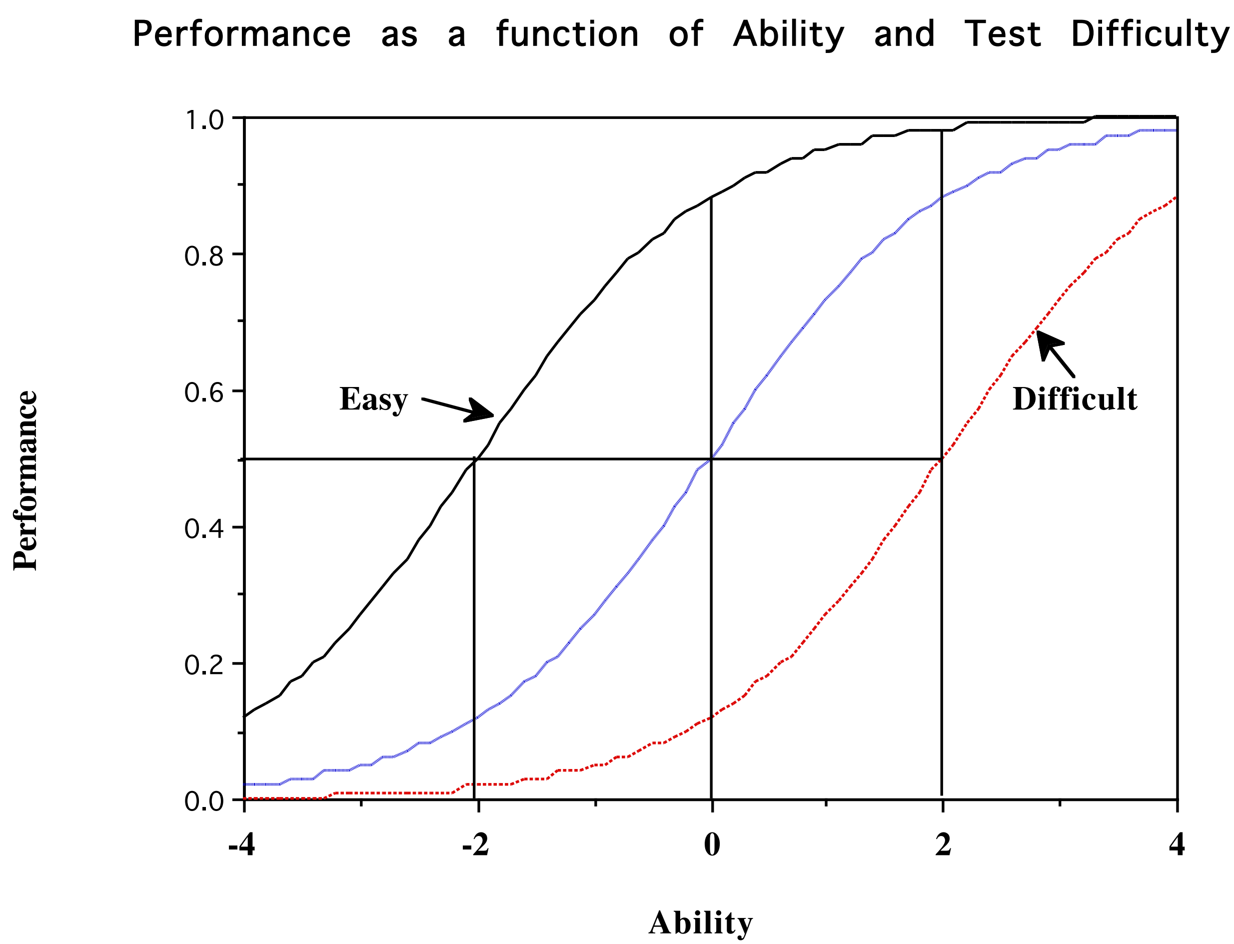
Experimenter 1 concluded that the low anxious had higher resistances, and thus were less aroused. But experimenter 2 noted that the low anxious had higher levels of skin conductance, and were thus more aroused.

How can this be?

Conductance = 1/Resistance



Performance and task difficulty



Performance, ability, and task difficulty

	Difficulty				
	-2	-1	0	1	2
Latent Ability					
-4.00	0.12	0.05	0.02	0.01	0.00
-2.00	0.50	0.27	0.12	0.05	0.02
0.00	0.88	0.73	0.50	0.27	0.12
2.00	0.98	0.95	0.88	0.73	0.50
4.00	1.00	0.99	0.98	0.95	0.88
Change from					
-4 to -2	0.38	0.22	0.10	0.04	0.02
-2 to -0	0.38	0.46	0.38	0.22	0.10
0 to 2	0.10	0.22	0.38	0.46	0.38
2 to 4	0.02	0.04	0.10	0.22	0.38

Performance and Task Difficulty

Note that equal differences along the latent ability dimension result in unequal differences along the observed performance dimension. Compare particularly performance changes resulting from ability changes from -2 to 0 to 2 units.

This is taken from the standard logistic transformation used in Item Response Theory that maps latent ability and latent difficulty into observed scores. IRT attempts to estimate difficulty and ability from the observed patterns of performance.

$$\text{Performance} = 1/(1+\exp^{(\text{difficulty}-\text{ability})})$$

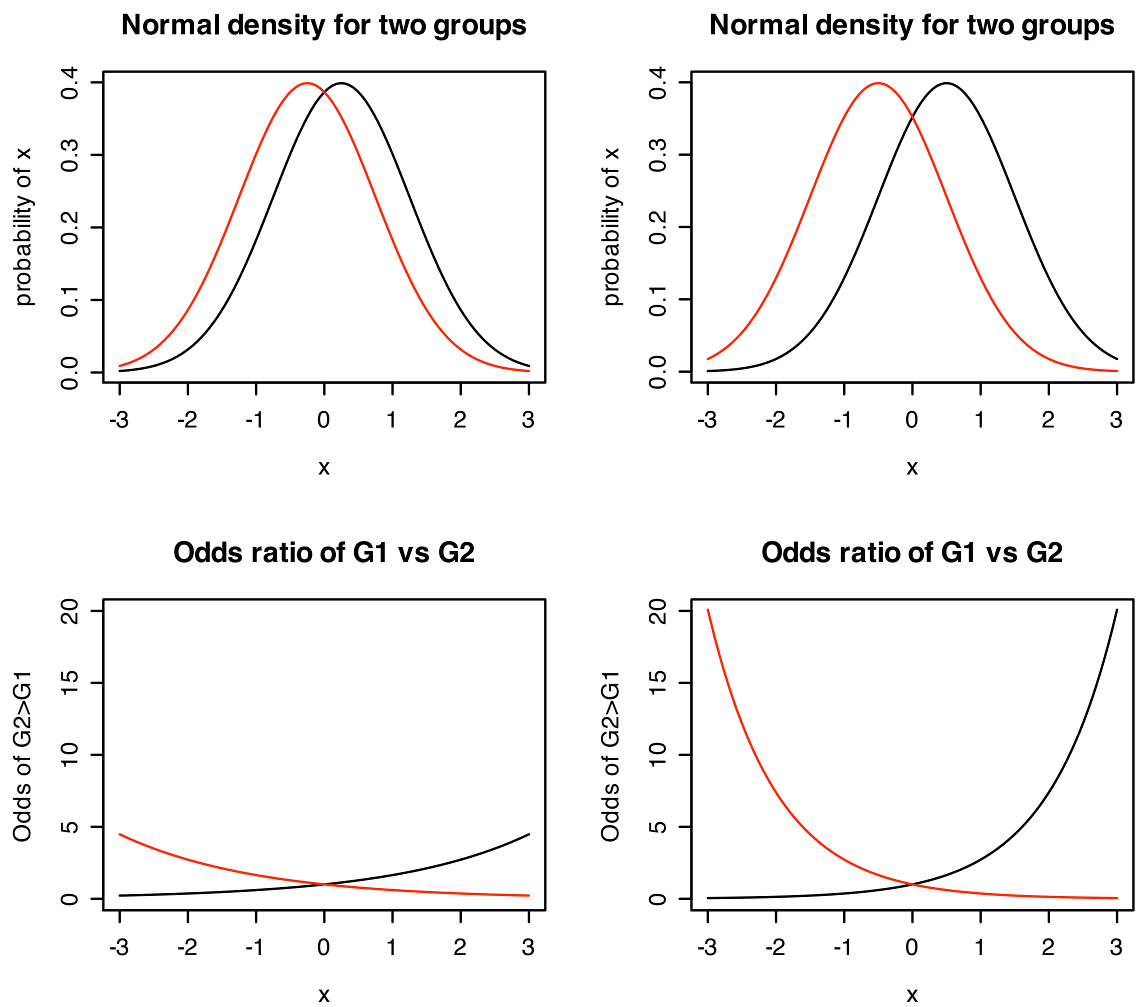
Decision making and the benefit of extreme selection ratios

- Typical traits are approximated by a normal distribution.
- Small differences in means or variances can lead to large differences in relative odds at the tails
- Accuracy of decision/prediction is higher for extreme values.
- Do we infer trait mean differences from observing differences of extreme values?

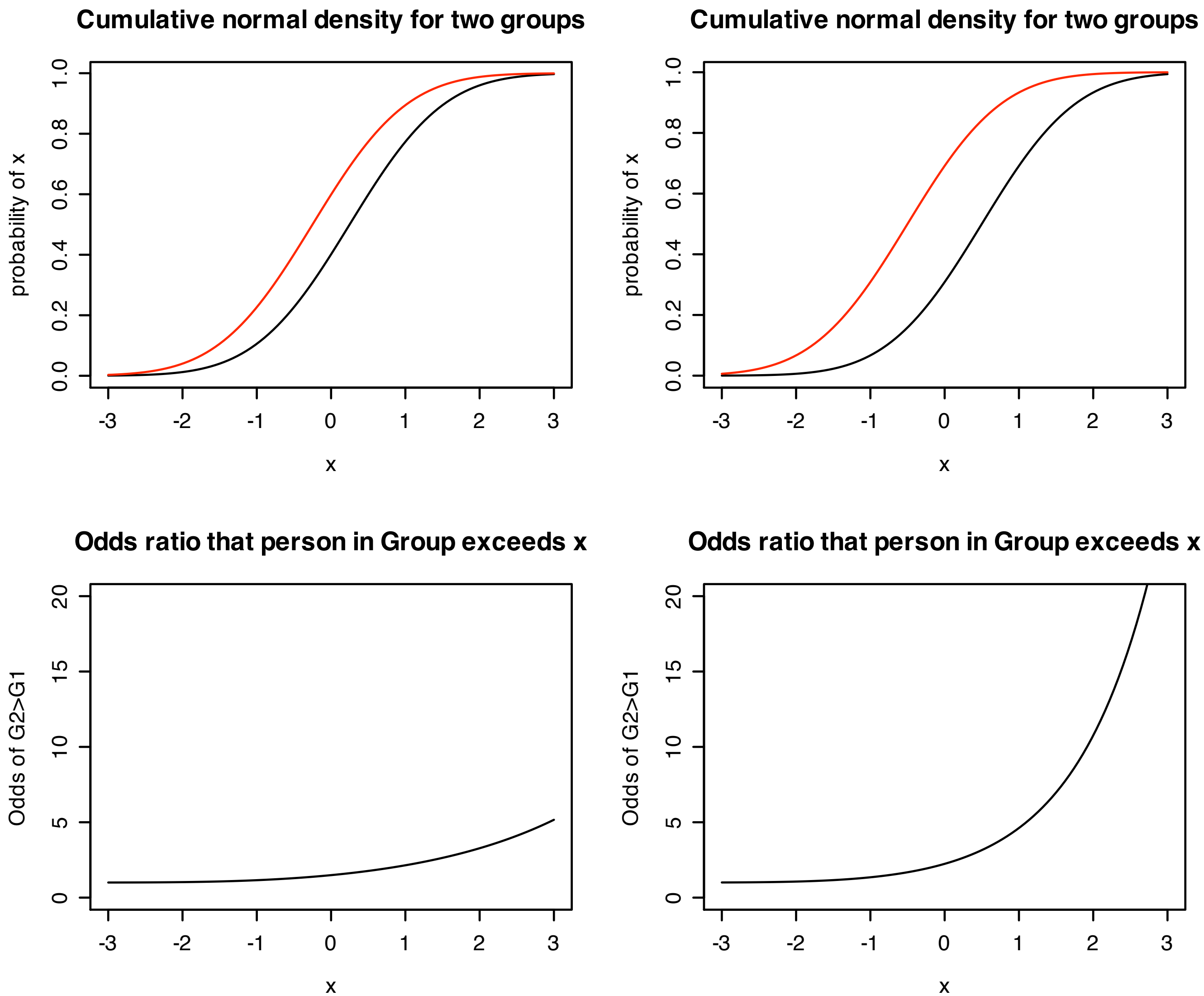
- (code for these graphs at
  - http://personality-project.org/r/extremescores.r)

Odds ratios as f(mean difference, extremity)

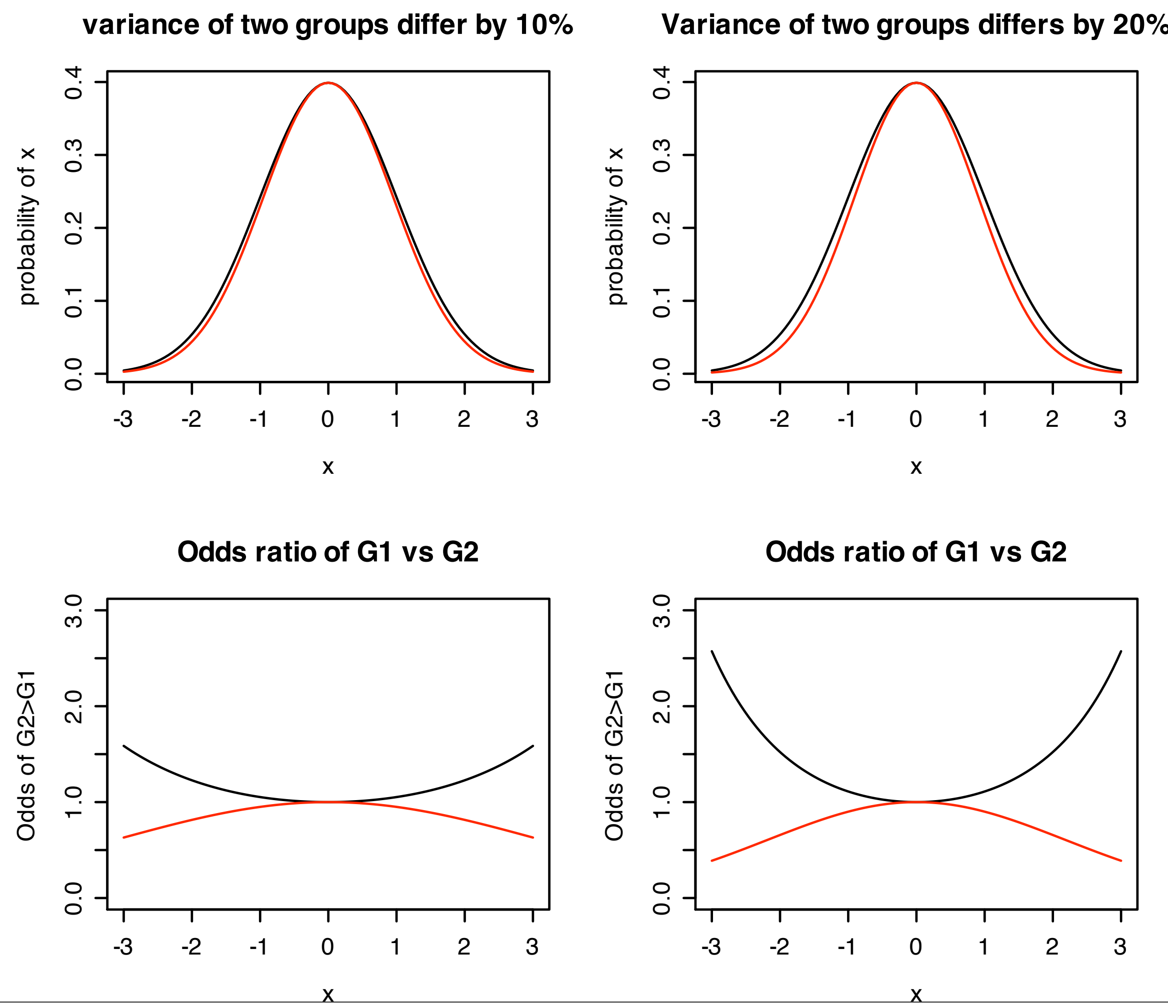
Difference =.5 sigma Difference =1.0 sigma



The effect of group differences on likelihood of extreme scores



The effect of differences of variance on odds ratios at the tails



Percentiles are not a linear metric and percentile odds are even worse!

- When comparing changes due to interventions or environmental trends, it is tempting to see how many people achieve a certain level (eg., of educational accomplishment, or of obesity), but the magnitude of such changes are sensitive to starting points, particularly when using percentiles or even worse, odds of percentiles.
- Consider the case of obesity:

Obesity gets worse over time

- "Over the last 15 years, obesity in the US has doubled, going from one in 10 to one in five. But the prevalence of *morbid* obesity has quadrupled, meaning that **the number of people 100 pounds overweight has gone from one in 200 to one in 50**. And the number of people roughly 150 pounds overweight has increased by a factor of 5, spiraling from one in 2000 to one in 400."
- "… The fact that super obesity is increasing faster than other categories of overweight suggests a strong environmental component (such as larger portions). If this were a strictly genetic predisposition, the numbers would rise only in proportion to the increase in other weight categories." (Tufts Health Newsletter, Dec. 2003, p 2)

Is obesity getter worse for the super obese? - Seemingly

Label	Definition	Odds	Change in Odds
Obese	BMI = 30 40 lb for 5’5”	1/10 to 1/5	2
Morbid Obese	BMI = 40 100 lb	1/200 to 1/50	4
Super Obese	BMI = 50 150 lb	1/2000 to 1/400	5

Is obesity getter worse for the super obese? -- No

Label	Definition	Odds	Change in Odds	z score	Change in z
Obese	BMI = 30	1/10 to 1/5	2	-1.28 -.84	0.44
Morbid Obese	BMI = 40	1/200 to 1/50	4	-2.58 -2.05	0.53
Super Obese	BMI = 50	1/2000 to 1/400	5	-3.29 -2.81	0.48

Psychometric Theory: A conceptual Syllabus