Psychology 454: Psychological Measurement using latent variables

William Revelle
Northwestern University

Winter, 2007

http://personality-project.org/revelle/syllabi/454/454.syllabus.pdf

Overview

I. Goals

A. To understand the fundamental concepts in latent variable modeling

B. To understand how to evaluate the quality of models when applied to data by understanding various threats to validity

C. To learn how to apply these concepts to real data sets using a variety of standard statistical packages

Texts and readings

I. Loehlin, J. C. Latent Variable Models (4th ed). Lawrence Erlbaum Associates, Mahwah, N.J. 2004

II. multiple web based readings

   A. e.g., Shrout, Widaman, etc.

III. syllabus and handouts available at http:personality-project.org/revelle/syllabi/454/454.syllabus.pdf

Overview

I. Observed Variables, Latent Variables, and theory
   A. Path models with observed variables (regression)
   B. Path models with observed and latent variables (sem)
II. Review of correlation, regression, reliability and matrix algebra
III. Estimation of models (OLS, WLS, ML)
   A. Evaluation of models
IV. Complex models
   A. multiple time points
   B. multiple groups

Overview - continued

I. Theory and estimation

A. Application of particular computer programs

B. Open Source (my preference)

1. R (download or use through SSCC)

2. Mx (download or use through SSCC)

C. Proprietary (if you want)

1. LISREL, Prelis (available through SSCC)

2. EQS, AMOS

Requirements/Evaluation

I. Basic knowledge of psychometrics
II. Familiarity with basic matrix algebra (or, at least, a willingness to learn)
III. Willingness to use computer programs - comparing alternative solutions, playing with data
IV. Willingness to ask questions
V. Weekly problem sets/final brief paper

Overview: Outline

I. Review of correlation/regression/reliability/ matrix algebra (405 in a day)

II. Basic model fitting/path analysis

III. Simple models

IV. Goodness of fit- what is it all about

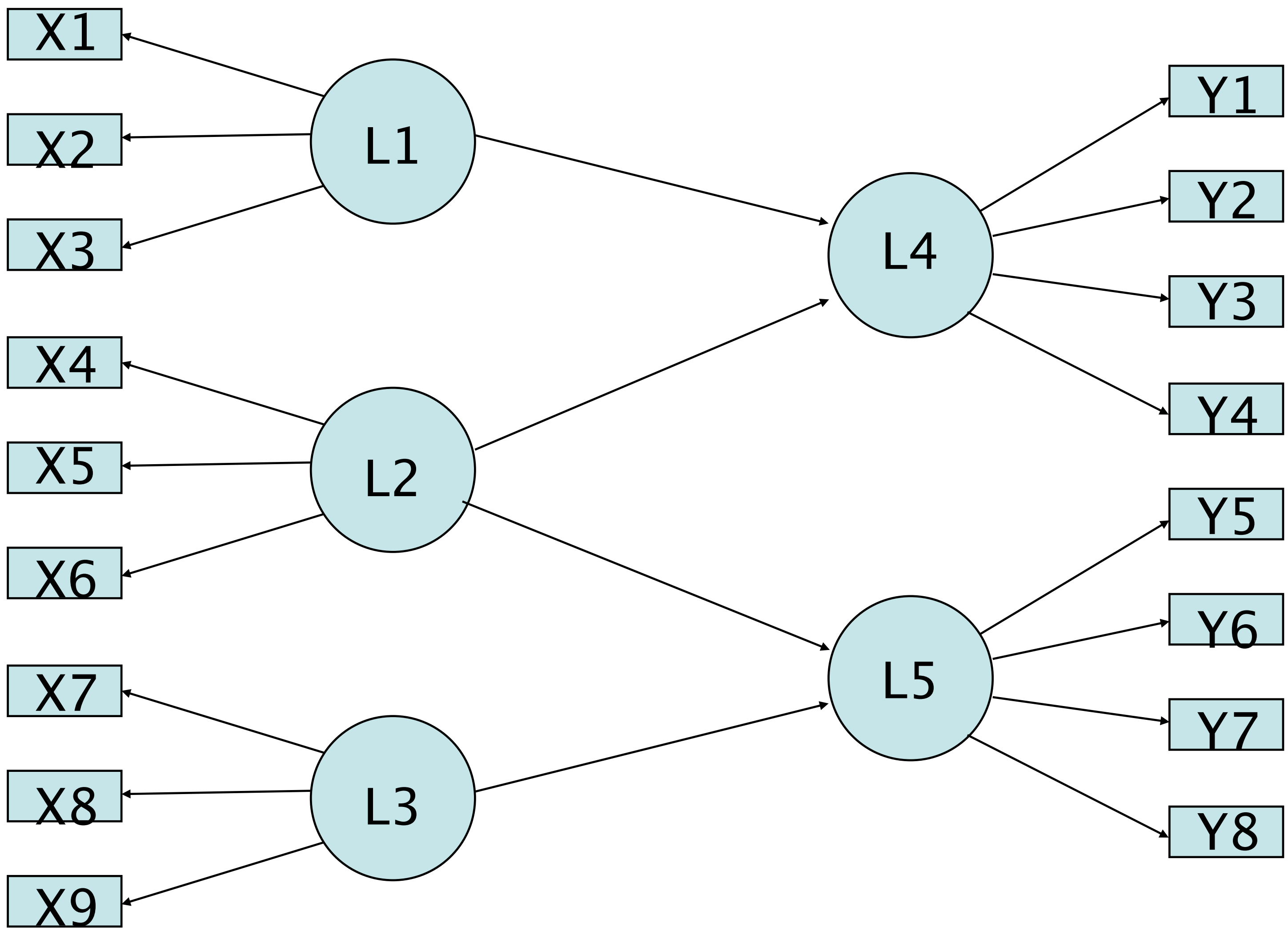
V. Exploratory Factor Analysis

VI. Confirmatory Factor Analysis

VII. Multiple groups/multiple occasions

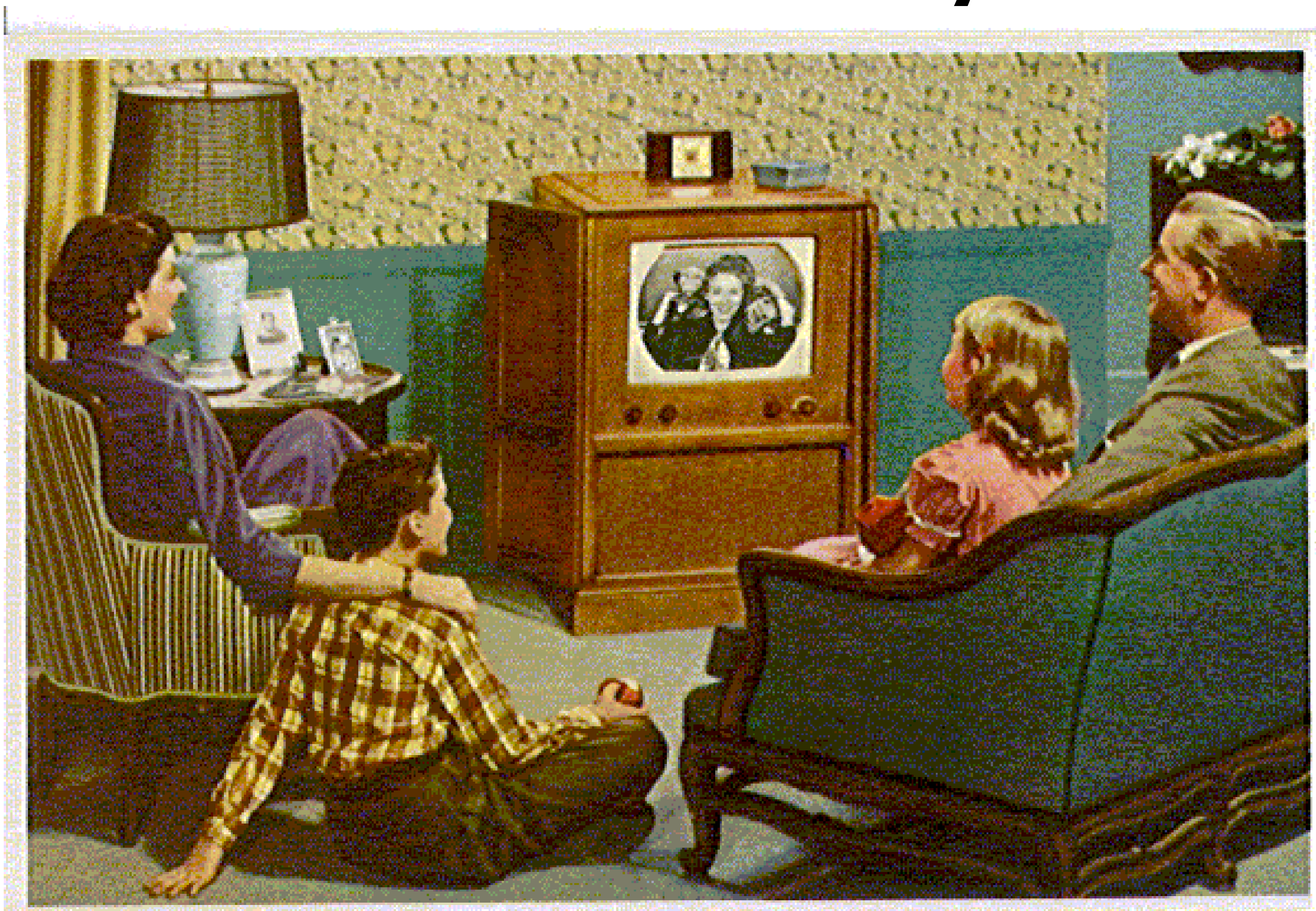
VIII. Further topics

Psychometric Theory: 405 in a day



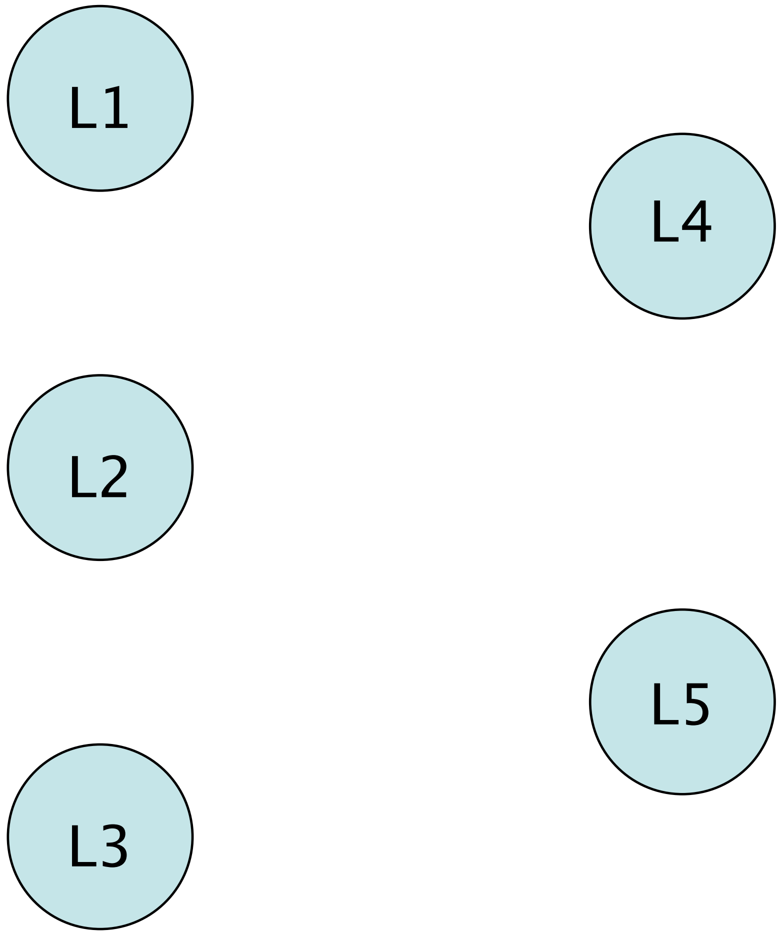
It all started with Plato

The modern day Cave?



faculty.frostberg.edu/phil/forum/TheCave.htm

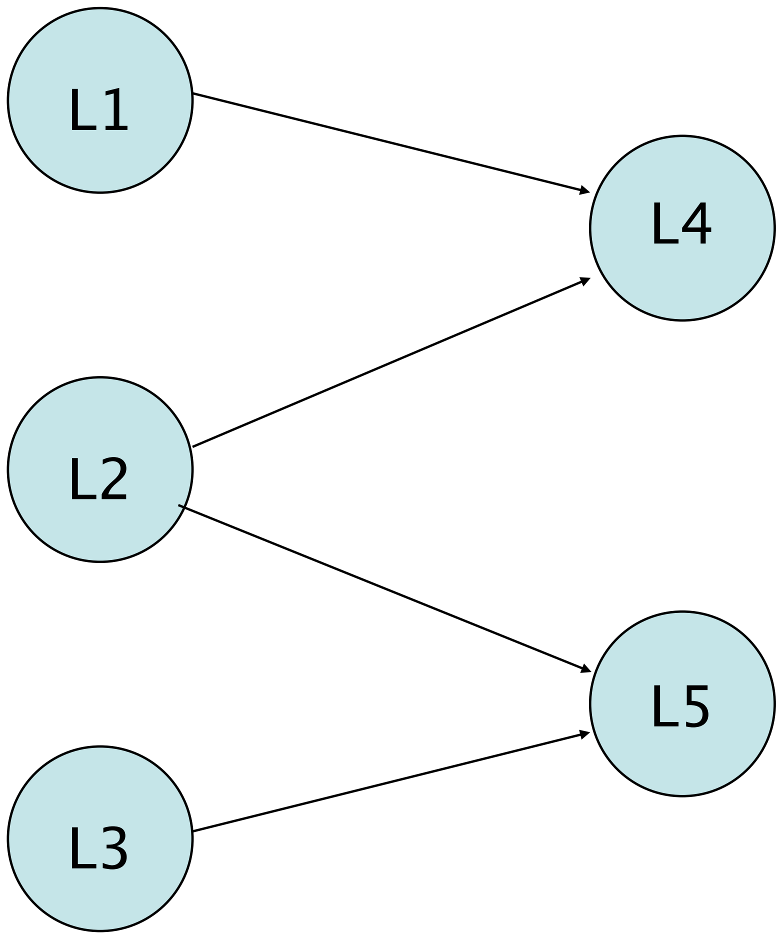
Constructs/Latent Variables



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

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

Observable or measured variables

Observed Variables

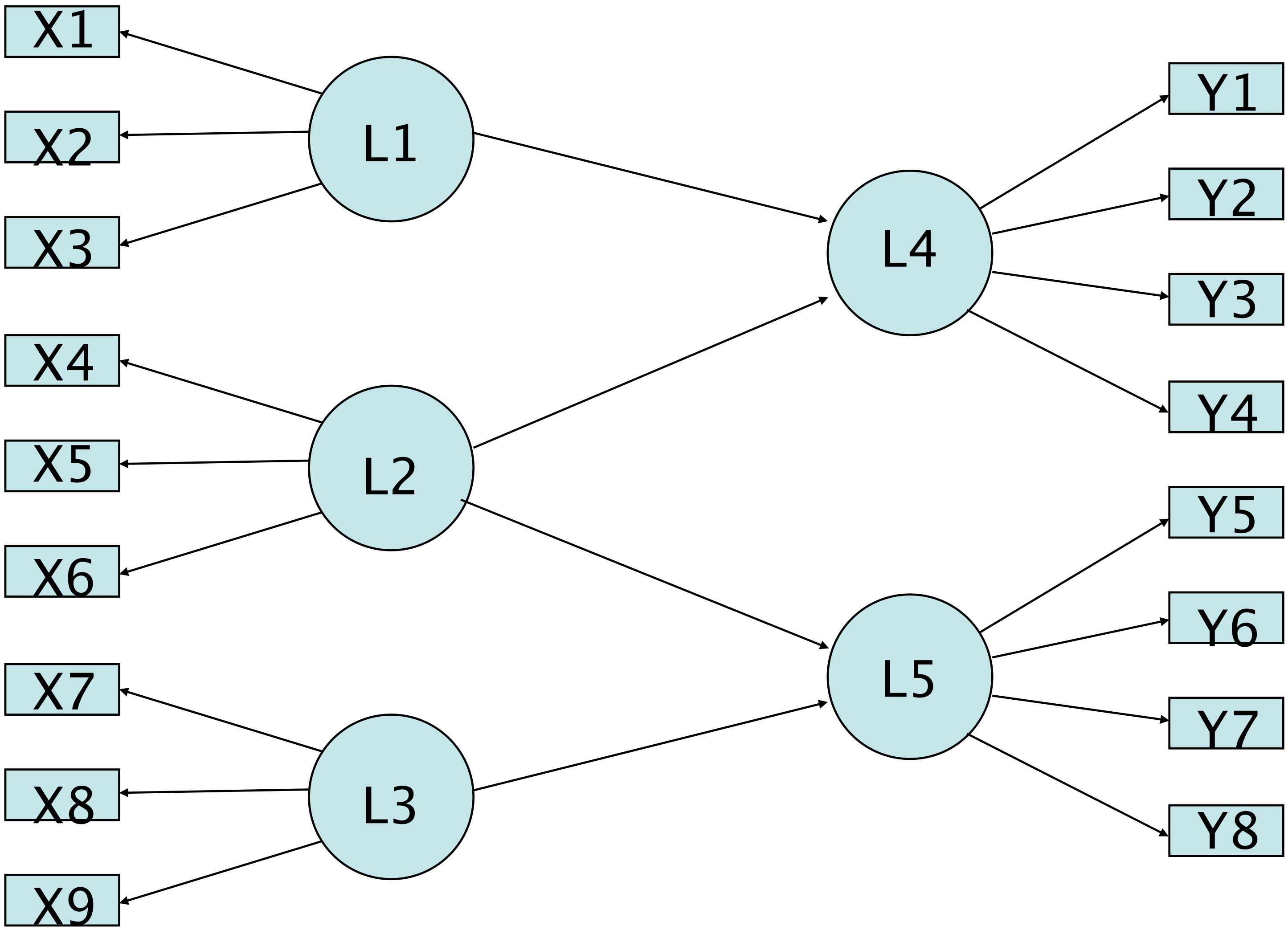
I. Psychological
   A. Item Endorsement
      1. Choice/Preference
      2. Reaction time
   B. Physiological
      1. Blood Oxygen Level Dependent Response
      2. Skin Conductance/Heart Rate

II. Sociological/economic
   A. Income
   B. Education
   C. Mortality

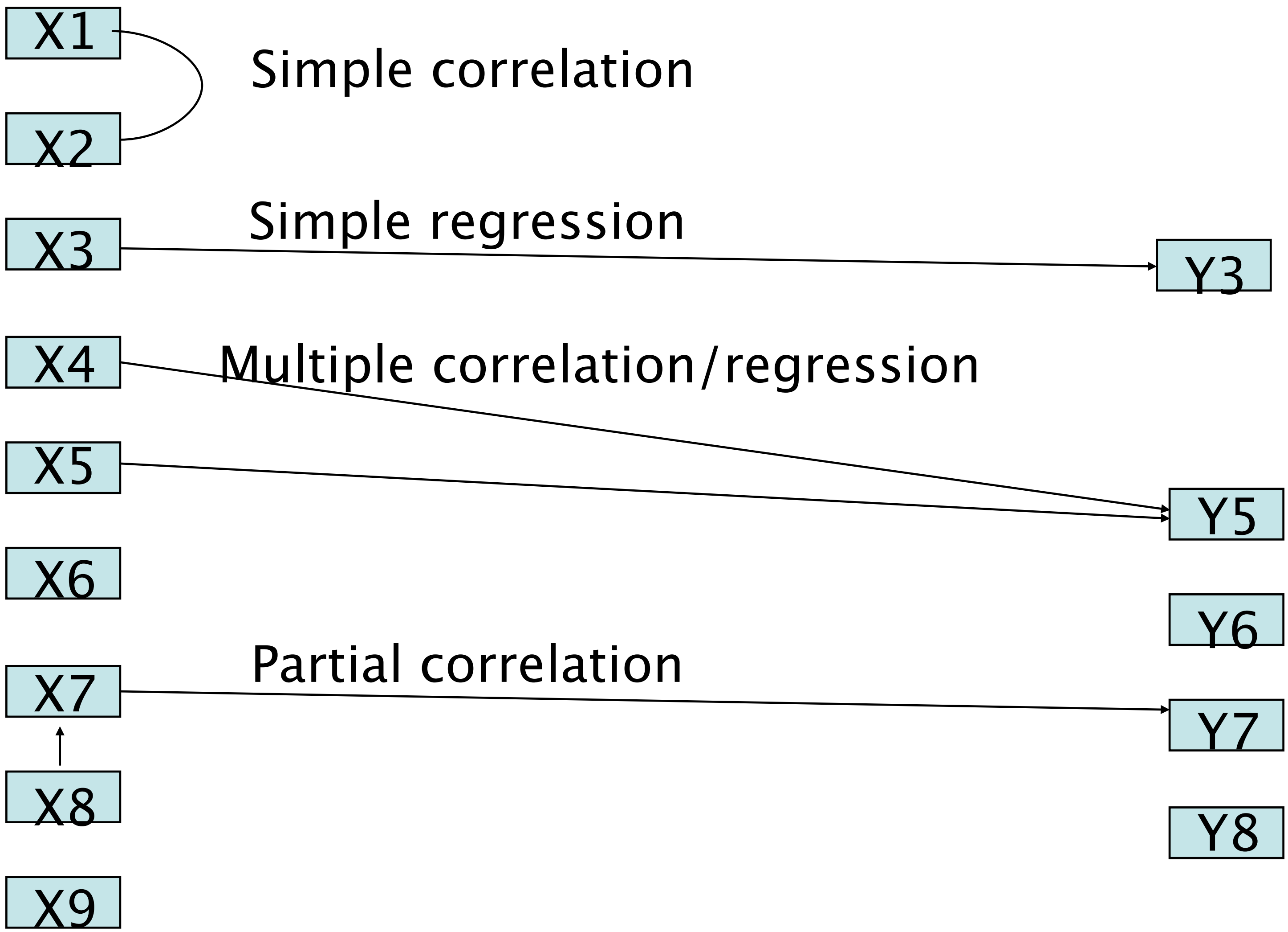
Theory development and testing

- 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

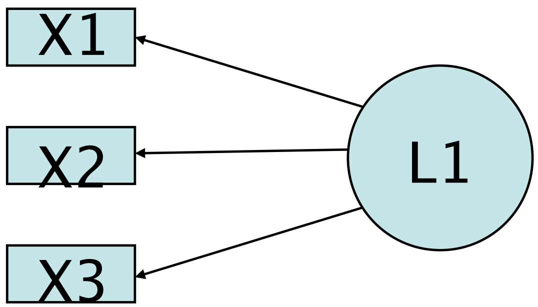
Psychometric Theory: 405 in a day



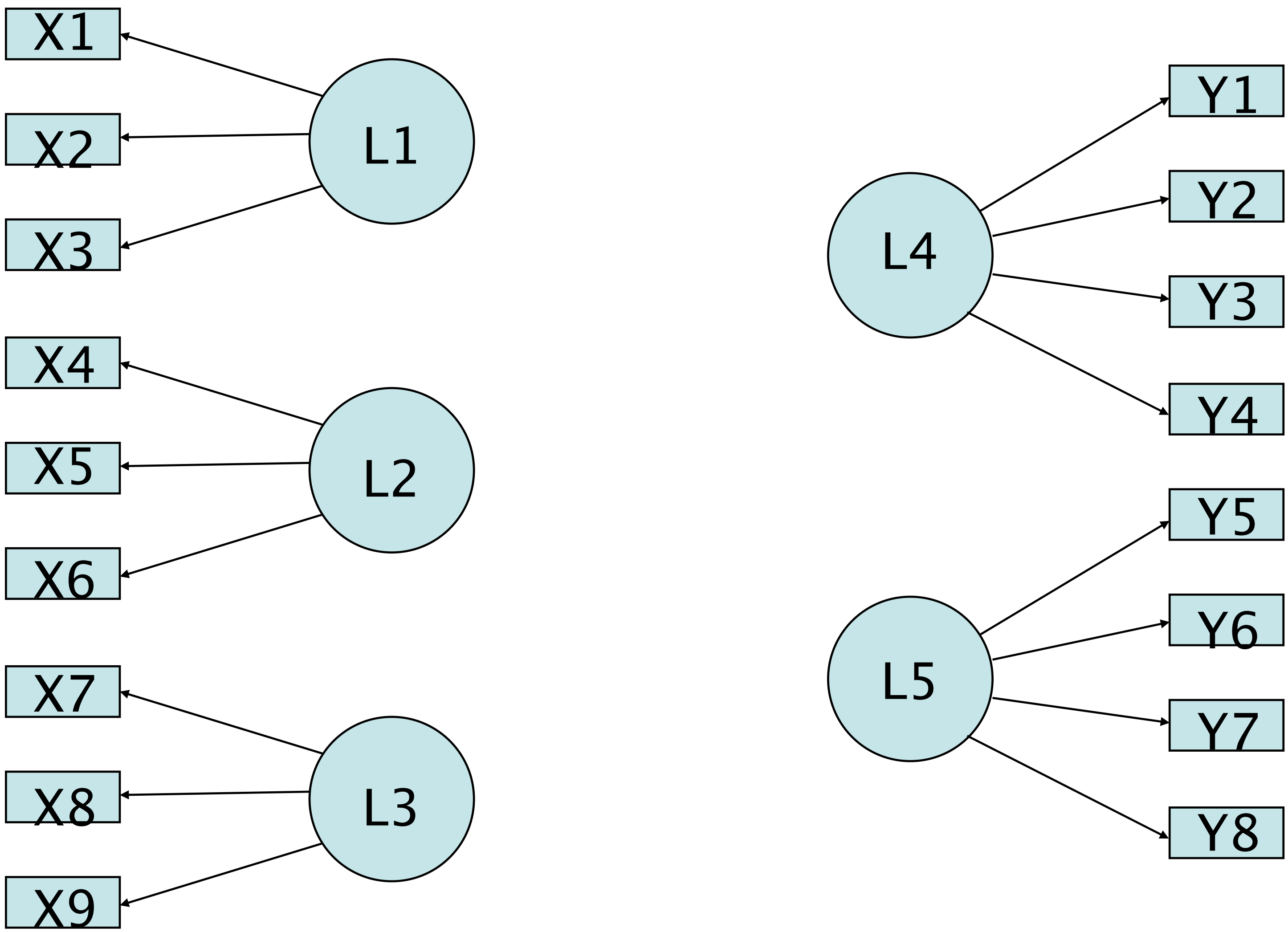
Variance, Covariance, and Correlation



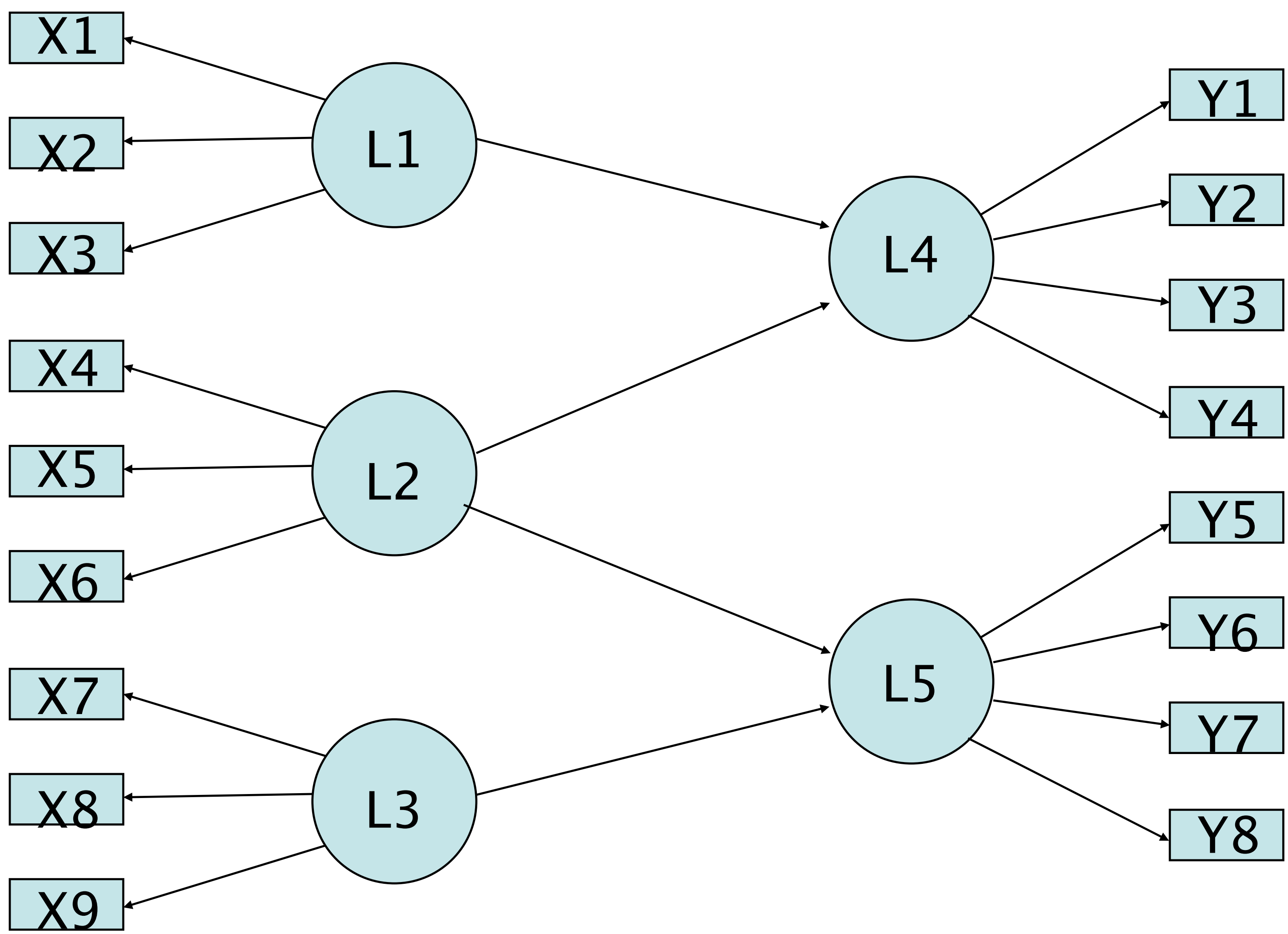
Classic Reliability Theory: How well do we measure what ever we are



Techniques of Data Reduction: Factor and Components Analysis

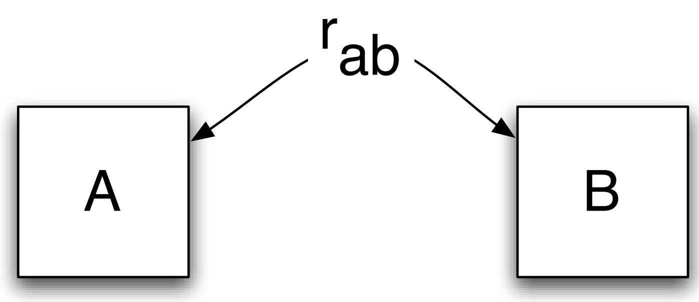


Structural Equation Modeling: Combining Measurement and Structural

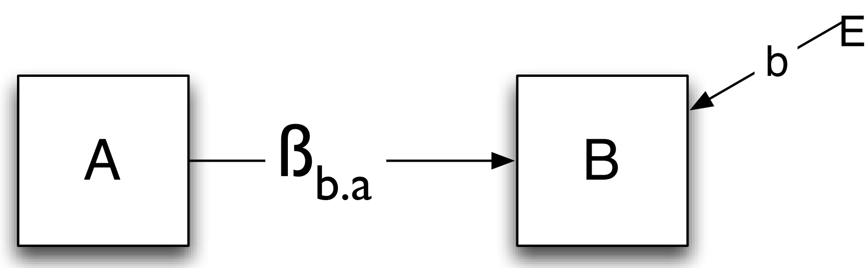


Correlation and Regression

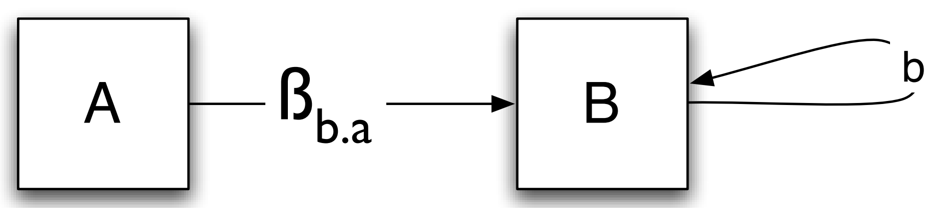
$r_{ab}= C_{ab}/\mathrm{sqrt}(V_aV_b)$



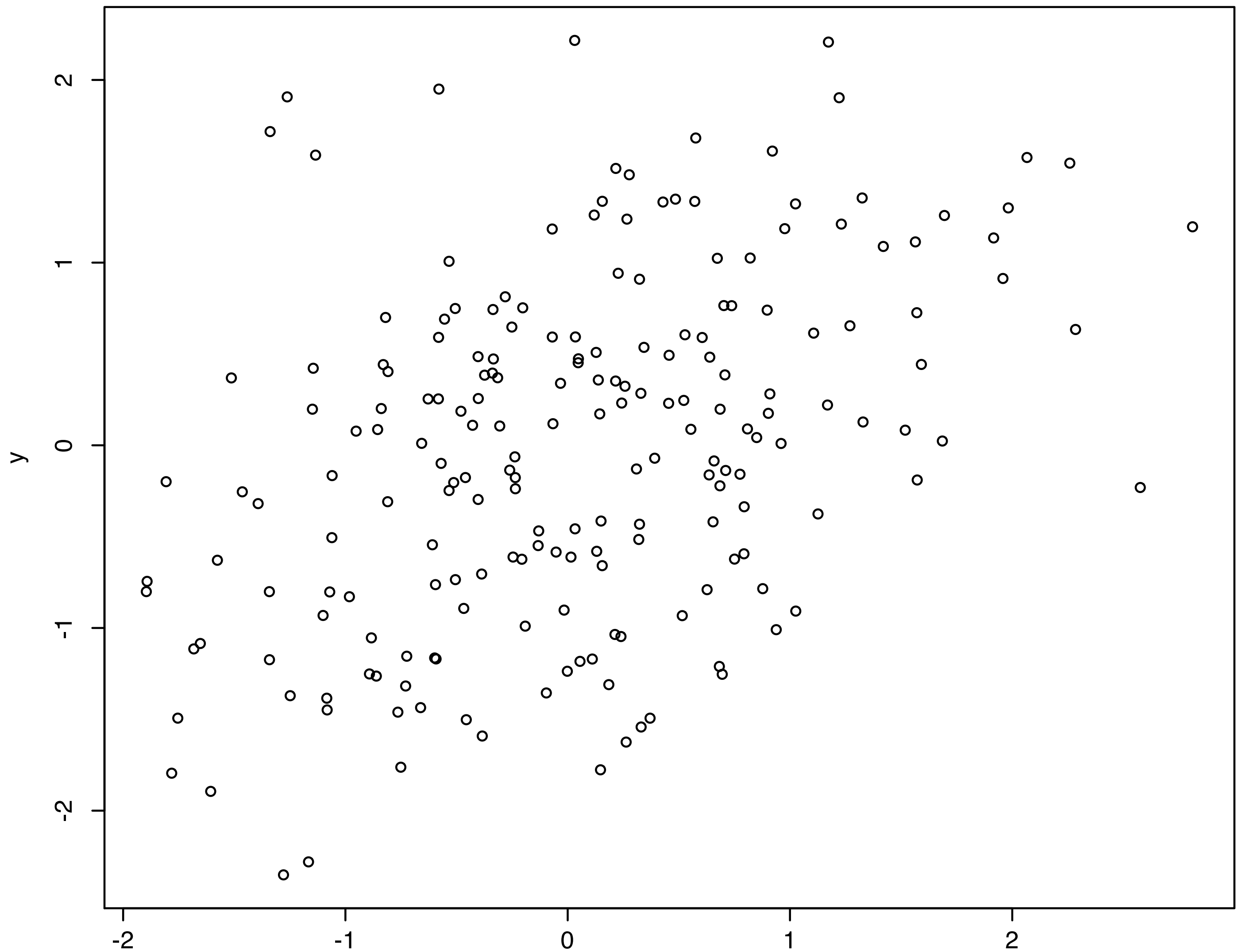
$\beta_{b.a} = C_{ab}/V_a$



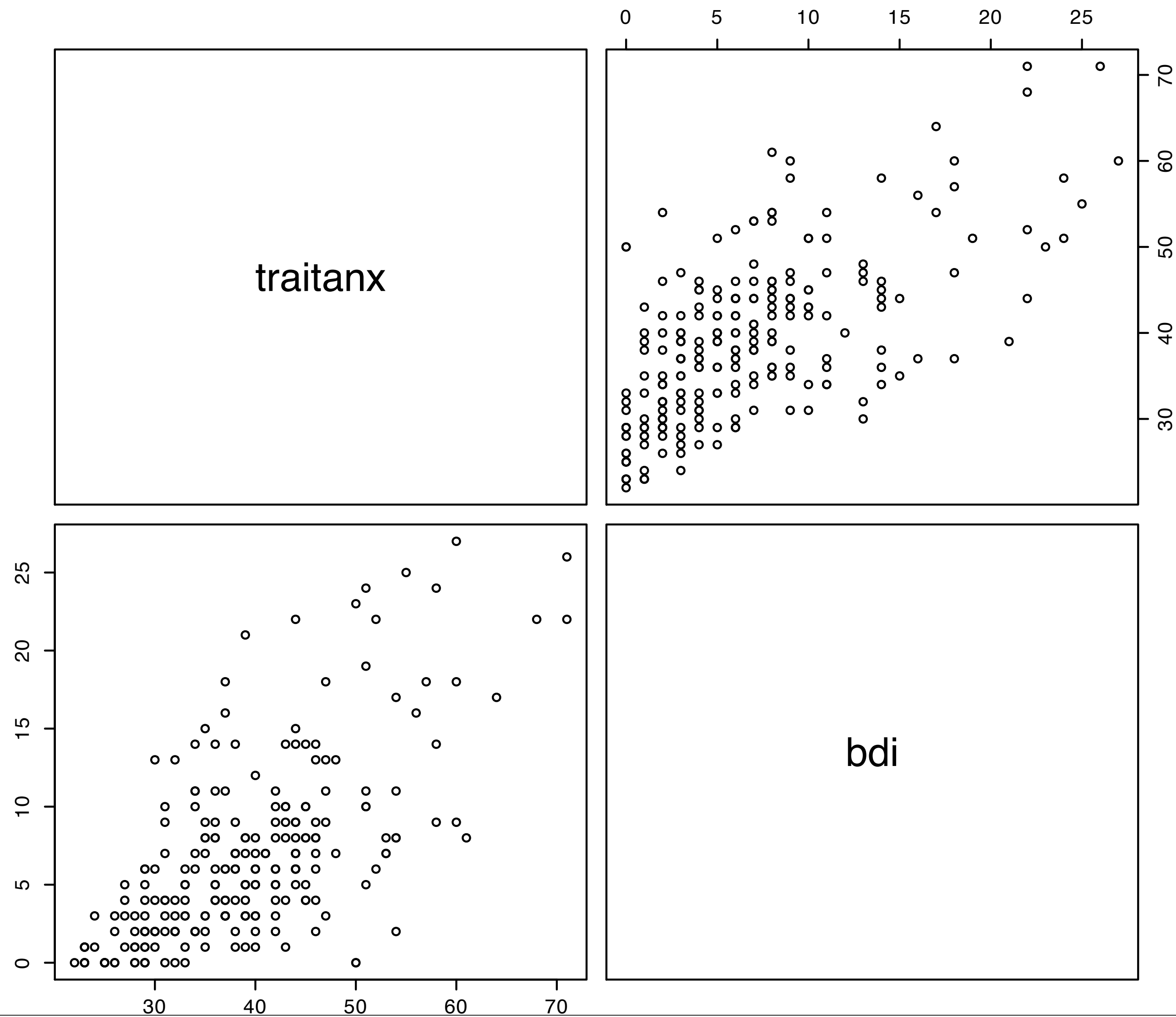
$V_b = \beta_{b.a}V_a + b^2e$



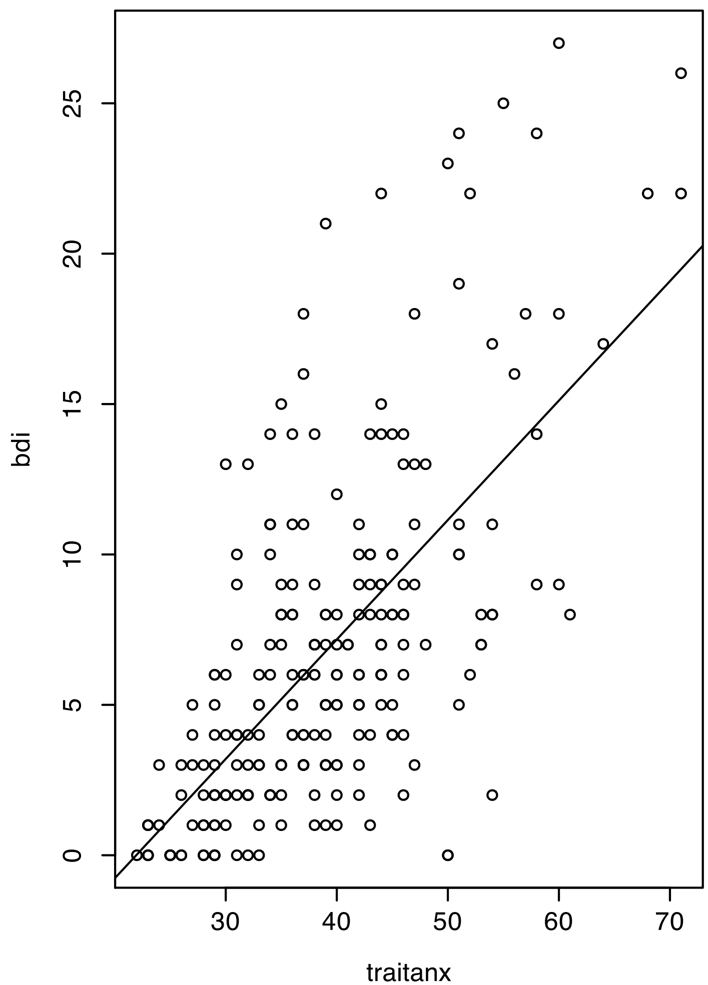
Joint distribution of X

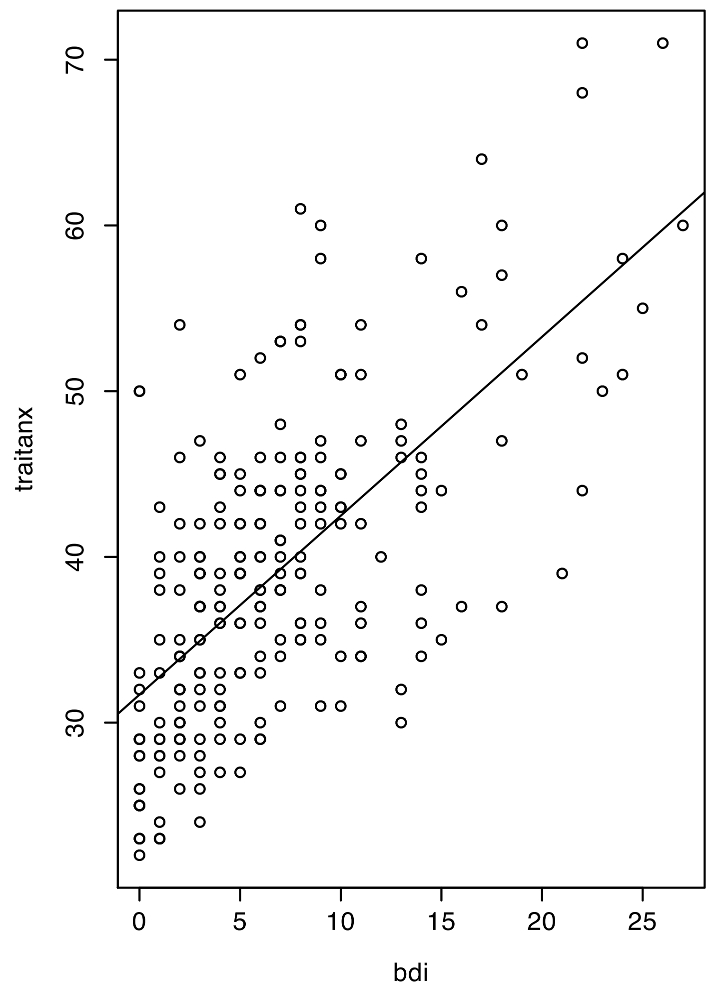


Beck Depresion x Trait Anxiety (raw)

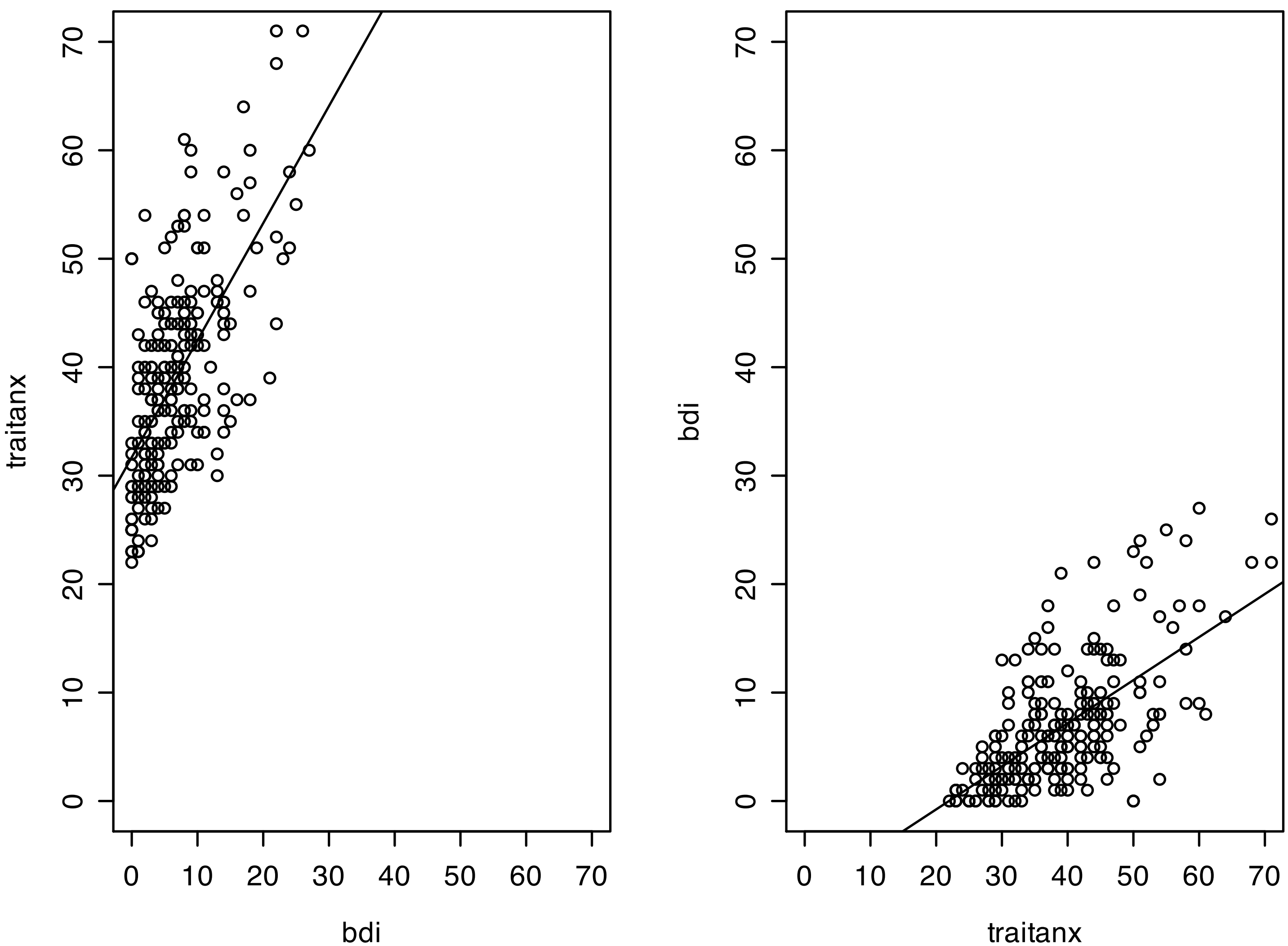


BDI x Trait Anx (raw)



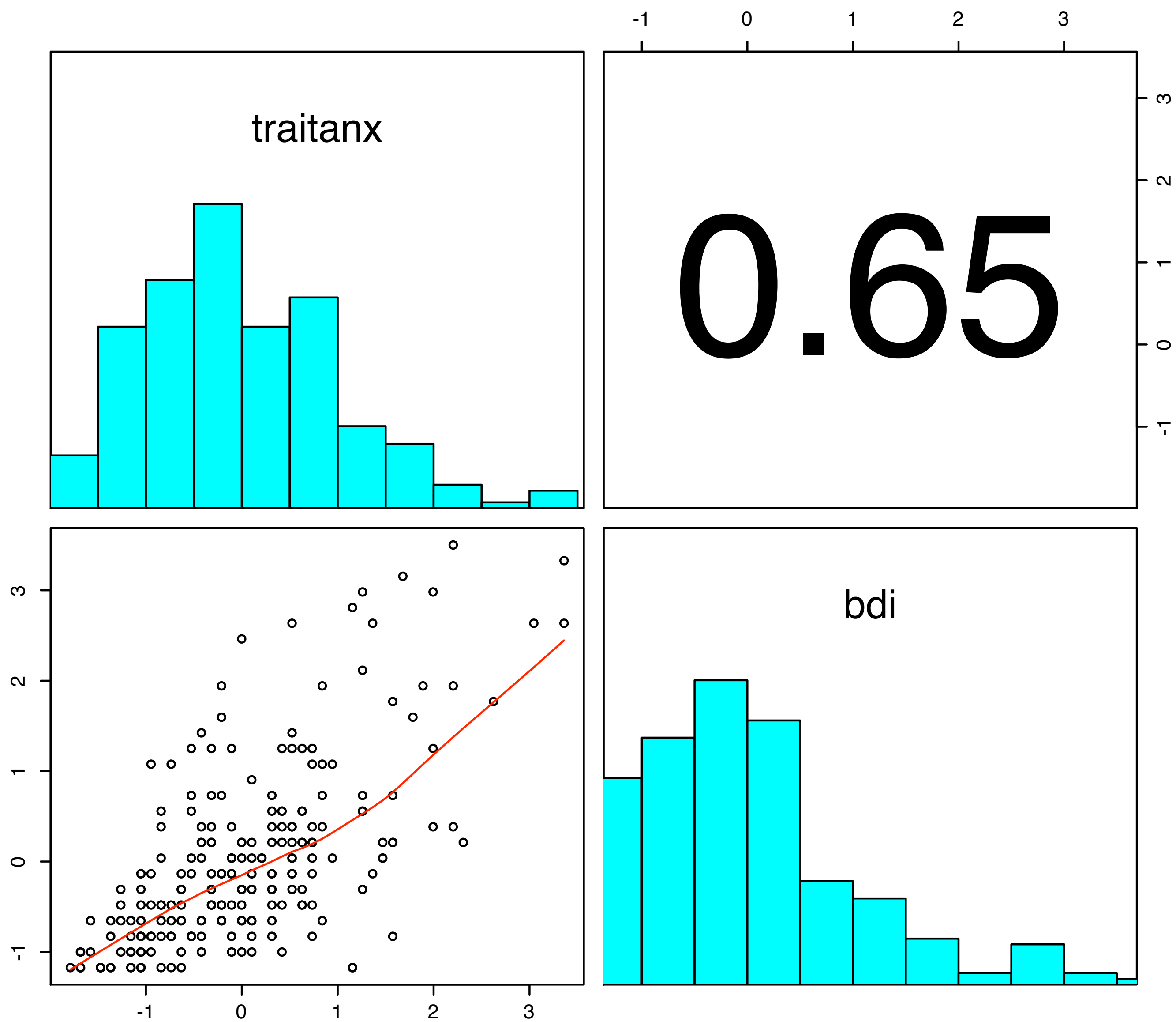


Regression lines depend upon scale

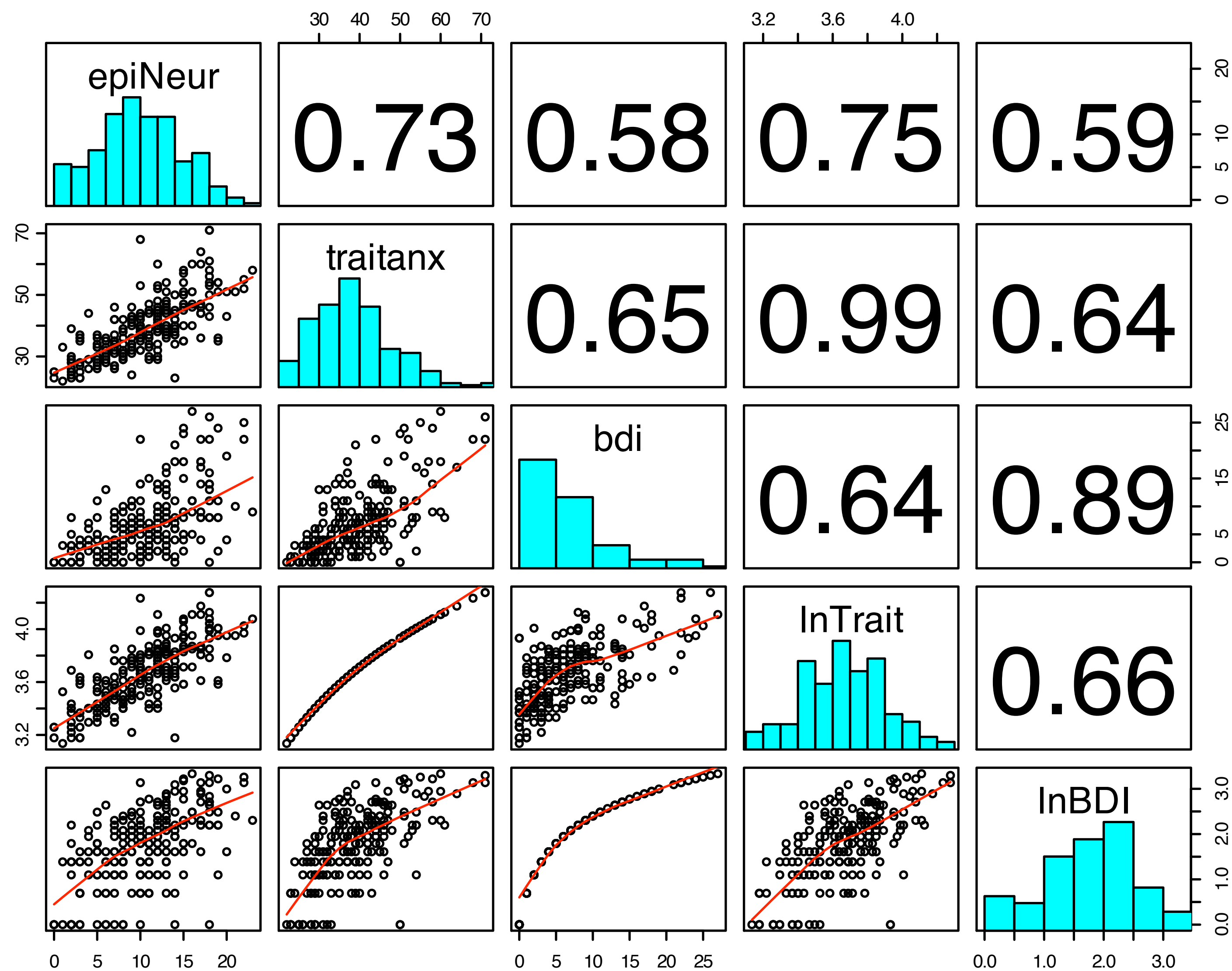


Beck Depression * Trait Anxiety

z scores



Transforming can help



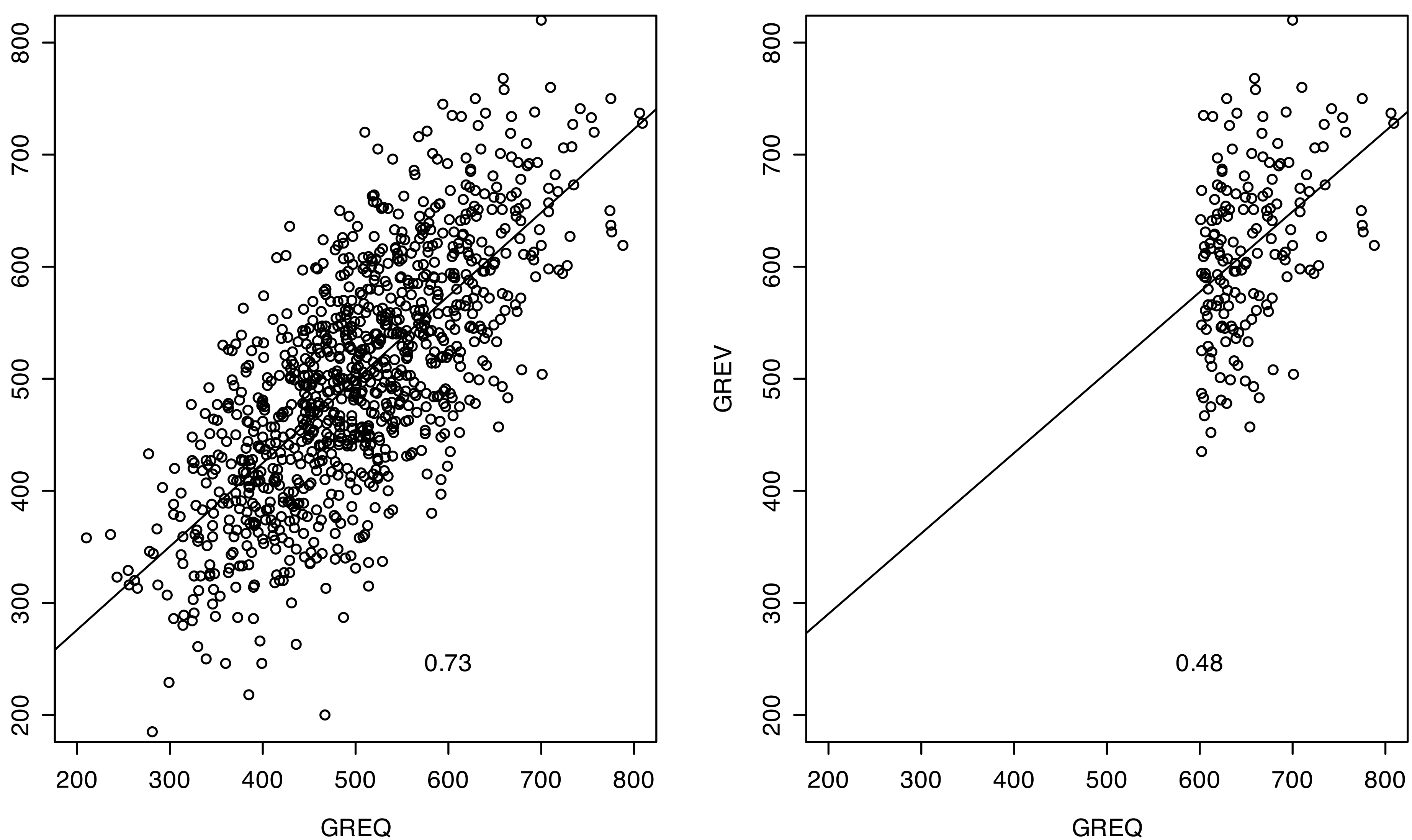
Alternative forms of r

$r = cov_{xy}/Sqrt(V_x * V_y) =$

$(\sum xy/N)(sqrt(\sum x^2/N * \sum y^2/N) = (\sum xy)(sqrt(\sum x^2 * \sum y^2)$

Correlation	X	Y
Pearson	Continuous	Continuous
Spearman	Ranks	ranks
Point biserial	Dichotomous	Continuous
Phi	Dichotomous	Dichotomous
Biserial	Dichotomous (assumed normal)	Continuous
Tetrachoric	Dichotomous (assumed normal)	Dichotomous (assumed normal
Polychoric	categorical (assumed normal)	categorical (assumed normal)

The effect of restriction of range on regression slopes vs. correlations



Caution with correlation

Consider 8 variables with means:

x1	x2	x3	x4	y1	y2	y3	y4
9.0	9.0	9.0	9.0	7.5	7.5	7.5	7.5

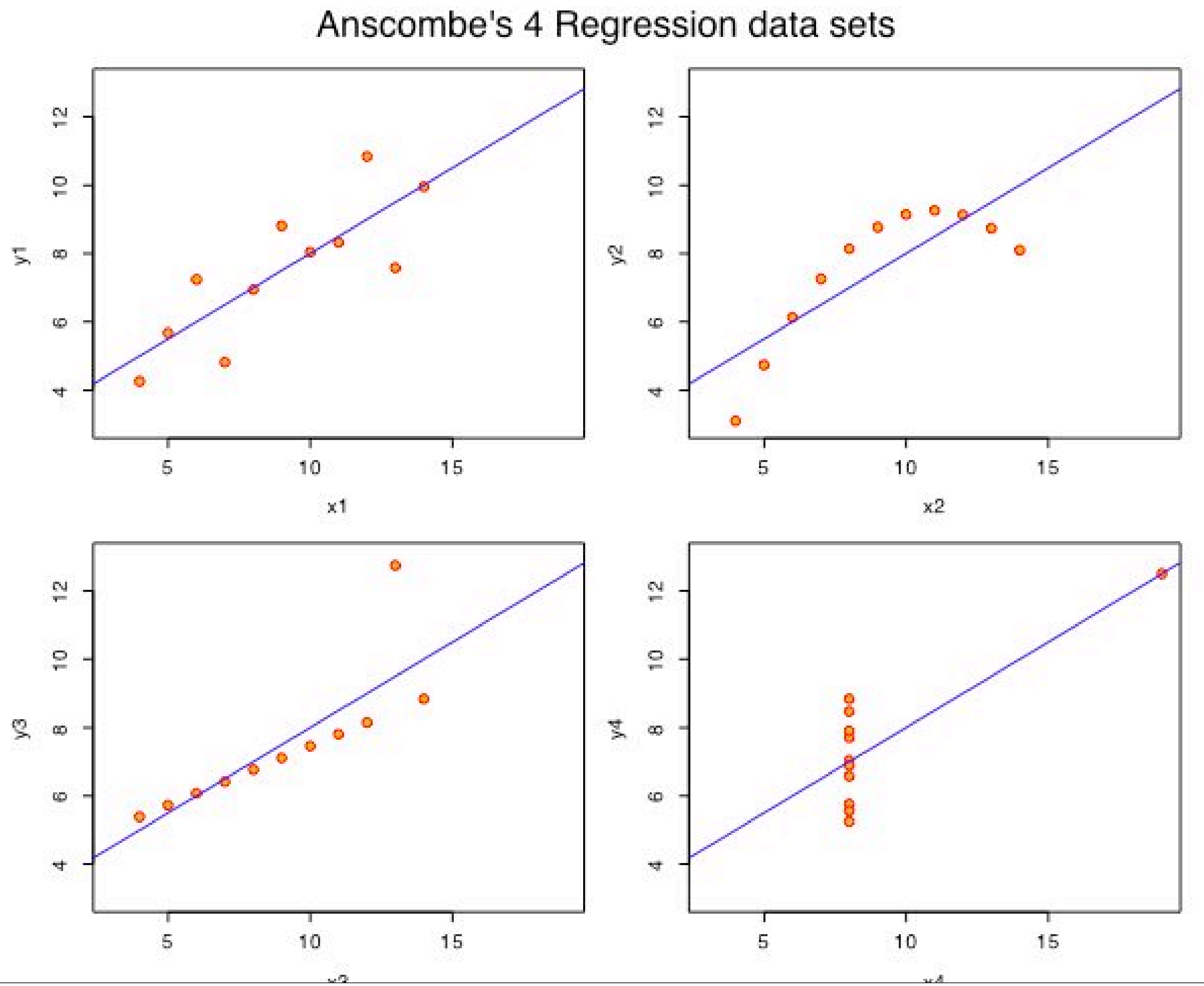
and Standard deviations

x1	x2	x3	x4	y1	y2	y3	y4
3.32	3.32	3.32	3.32	2.03	2.03	2.03	2.03

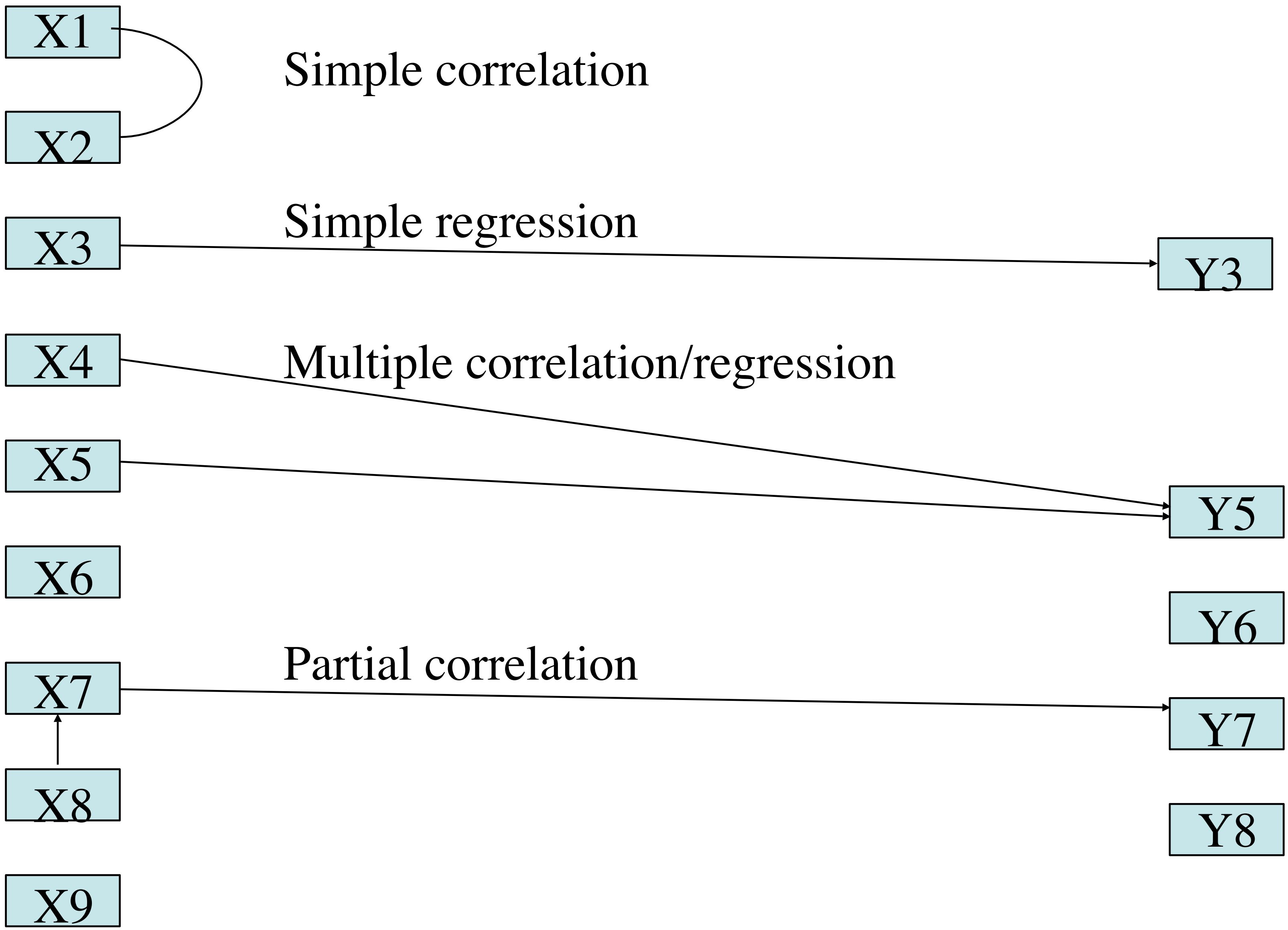
and correlations between xi and yi of

0.82 0.82 0.82 0.82

Caution with Correlation



Variance, Covariance, and Correlation



But first, a word from our sponsor:

I. Matrix algebra is the fundamental notational technique used in multiple correlation, factor analysis, and structural equation modeling

II. Although it is possible to use sem programs without understanding matrix algebra, it is much harder to do so.

III. Matrix algebra is a convenient notational system that allows us to think about data at a higher (broader) level rather than data point by data point.

Review of Matrix Algebra

I. scalers, vectors, and matrices

A. scalers: simple numbers

B. vectors: ordered sets of numbers

1. V1 = {1 2 3 4 5 6 7 8 9 10 }
2. V2 = { 11 12 13 14 15 16 17 18 19 20 }
3. V2[3] = 13

C. Matrices (vectors of vectors)

See personality-project.org/r/sem.appendix.1.pdf

Matrices

$$
{}_nX_m = \left\{ \begin{array}{c} x_{11}\ x_{12} \ldots x_{1m} \\ x_{21}\ x_{22} \ldots x_{2m} \\ \cdots\ \cdots\ \ldots\ \cdots \\ x_{n1}\ x_{n2} \ldots x_{nm} \end{array} \right\}
$$

R

	x1	x2	x3
x1	1.00	0.56	0.48
x2	0.56	1.00	0.42
x3	0.48	0.42	1.00

Xij

	V1	V2	V3	V4
S1	9	4	9	7
S2	9	7	1	8
S3	2	9	9	3
S4	8	2	9	6
S5	6	4	0	0
S6	5	9	5	8
S7	7	9	3	0
S8	1	1	9	2
S9	6	4	4	9
S10	7	5	8	6

Vector operations

I. addition

A. V3 <- V1 + V2

B. V3 = {12 14 16 18 20 22 24 26 28 30}

II. multiplication

A. element by element

1. V1*V2 = 11 24 39 56 75 96 119 144 171 200

B. inner product of vector (Sums of products)

C. outer product of vectors (matrix of products)

Inner and outer products

$$inner.product = \sum_{i=1}^{N} V1_i * V2_i$$

$${}_nX_1 *_1 Y_m =_n (XY)_m$$

Outer product (graphically)

```
> V1

 [1]  1  2  3  4  5  6  7  8  9 10

> V2

[1] 1 2 3 4

> outer.prod <- V1 %*% t(V2)
> outer.prod

      [,1] [,2] [,3] [,4]
 [1,]    1    2    3    4
 [2,]    2    4    6    8
 [3,]    3    6    9   12
 [4,]    4    8   12   16
 [5,]    5   10   15   20
 [6,]    6   12   18   24
 [7,]    7   14   21   28
 [8,]    8   16   24   32
 [9,]    9   18   27   36
[10,]   10   20   30   40
```

Matrix Operations

I. Addition/Subtraction

A. (element by element)

B. must be of same dimensions

II. Multiplication

A. ${}_{m}X_{n}\ {}_{n}Y_{p} = {}_{m}XY_{p}$ where the elements of XY, x_{ij} are the sums of the products of the elements of the ith row and jth column

B. $X\,Y \neq Y\,X$

Matrix multiplication

$${}_{m}X_{n} \; {}_{n}Y_{p} \;=\; {}_{m}XY_{p}$$

$$xy_{ij} = \sum_{k=1}^{N} x_{ik} * y_{jk}.$$

Matrix multiplication for data

one

```
1 1 1 1 1 1 1 1 1 1
```

one %*%Xij=

```
      V1 V2 V3 V4
[1,]  60 54 57 49
```

X.means <- one %*% Xij/n

```
      V1  V2  V3  V4
[1,]   6 5.4 5.7 4.9
```

Xij

	V1	V2	V3	V4
S1	9	4	9	7
S2	9	7	1	8
S3	2	9	9	3
S4	8	2	9	6
S5	6	4	0	0
S6	5	9	5	8
S7	7	9	3	0
S8	1	1	9	2
S9	6	4	4	9
S10	7	5	8	6

Deviation scores as matrix differences

```
X.diff <- Xij - one %*% X.means
```

	V1	V2	V3	V4
S1	3	-1.4	3.3	2.1
S2	3	1.6	-4.7	3.1
S3	-4	3.6	3.3	-1.9
S4	2	-3.4	3.3	1.1
S5	0	-1.4	-5.7	-4.9
S6	-1	3.6	-0.7	3.1
S7	1	3.6	-2.7	-4.9
S8	-5	-4.4	3.3	-2.9
S9	0	-1.4	-1.7	4.1
S10	1	-0.4	2.3	1.1

Covariance as matrix product

```
X.cov <- t(X.diff) %*% X.diff/(n - 1)
```

X.cov

```
      V1    V2    V3    V4
V1  7.33  0.11 -3.00  3.67
V2  0.11  8.71 -3.20 -0.18
V3 -3.00 -3.20 12.68  1.63
V4  3.67 -0.18  1.63 11.43
```

diag(X.cov)

```
   V1    V2    V3    V4
 7.33  8.71 12.68 11.43
```

Correlation = standardized covariance

	V1	V2	V3	V4
V1	0.37	0.00	0.00	0.0
V2	0.00	0.34	0.00	0.0
V3	0.00	0.00	0.28	0.0
V4	0.00	0.00	0.00	0.3

```
sdi <-
diag(1/sqrt(diag(X.cov)))
```

	V1	V2	V3	V4
V1	1.00	0.01	-0.31	0.40
V2	0.01	1.00	-0.30	-0.02
V3	-0.31	-0.30	1.00	0.14
V4	0.40	-0.02	0.14	1.00

```
X.cor <-
sdi %*% X.cov %*% sdi
```

The identity matrix

```
I <- diag(1,nrow=4)

     [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    0    1    0    0
[3,]    0    0    1    0
[4,]    0    0    0    1
```

Matrix Inverse

$$X'X^{-1} = X^{-1}X = I$$

	V1	V2	V3	V4
V1	1.00	0.01	-0.31	0.40
V2	0.01	1.00	-0.30	-0.02
V3	-0.31	-0.30	1.00	0.14
V4	0.40	-0.02	0.14	1.00

X.cor

	V1	V2	V3	V4
V1	1.44	0.15	0.58	-0.65
V2	0.15	1.12	0.40	-0.09
V3	0.58	0.40	1.36	-0.41
V4	-0.65	-0.09	-0.41	1.32

$X.cor^{-1}$

$X^{-1}X = XX^{-1} = I$

X.inv %*% X.cor

	V1	V2	V3	V4
V1	1	0	0	0
V2	0	1	0	0
V3	0	0	1	0
V4	0	0	0	1

X.cor %*% X.inv

	V1	V2	V3	V4
V1	1	0	0	0
V2	0	1	0	0
V3	0	0	1	0
V4	0	0	0	1

Multiple regresssion

I. At data level

A. $Y = X\beta + \partial$

B. $\beta = (X'X)^{-1} X'Y$

II. At structure level

A. $\beta = R^{-1} r_{xy}$

Multiple Regression: $y = xb \Rightarrow b_{xy} = R^{-1} r_{xy}$

R

	x1	x2	x3
x1	1	0	0
x2	0	1	0
x3	0	0	1

r_{xy}

	y
x1	0.8
x2	0.7
x3	0.6

R^{-1}

	x1	x2	x3
x1	1	0	0
x2	0	1	0
x3	0	0	1

$R^{-1} r_{xy}$

	y
x1	0.8
x2	0.7
x3	0.6

Multiple Regression: $y = xb \rightarrow b_{xy} = R^{-1}r_{xy}$

R

	x1	x2	x3
x1	1.00	0.56	0.48
x2	0.56	1.00	0.42
x3	0.48	0.42	1.00

r_{xy}

	y
x1	0.8
x2	0.7
x3	0.6

R^{-1}

	x1	x2	x3
x1	1.63	−0.71	−0.48
x2	−0.71	1.52	−0.30
x3	−0.48	−0.30	1.36

$R^{-1}\, r_{xy}$

	y
x1	0.52
x2	0.32
x3	0.22

Multiple Regression: $y = xb \rightarrow b_{xy} = R^{-1} r_{xy}$

R

	x1	x2	x3
x1	1.0	0.8	0.8
x2	0.8	1.0	0.8
x3	0.8	0.8	1.0

r_{xy}

	y
x1	0.8
x2	0.7
x3	0.6

R^{-1}

	x1	x2	x3
x1	3.46	−1.54	−1.54
x2	−1.54	3.46	−1.54
x3	−1.54	−1.54	3.46

$R^{-1} r_{xy}$

	y
x1	0.77
x2	0.27
x3	−0.23

Solution space is relatively flat as f(beta)

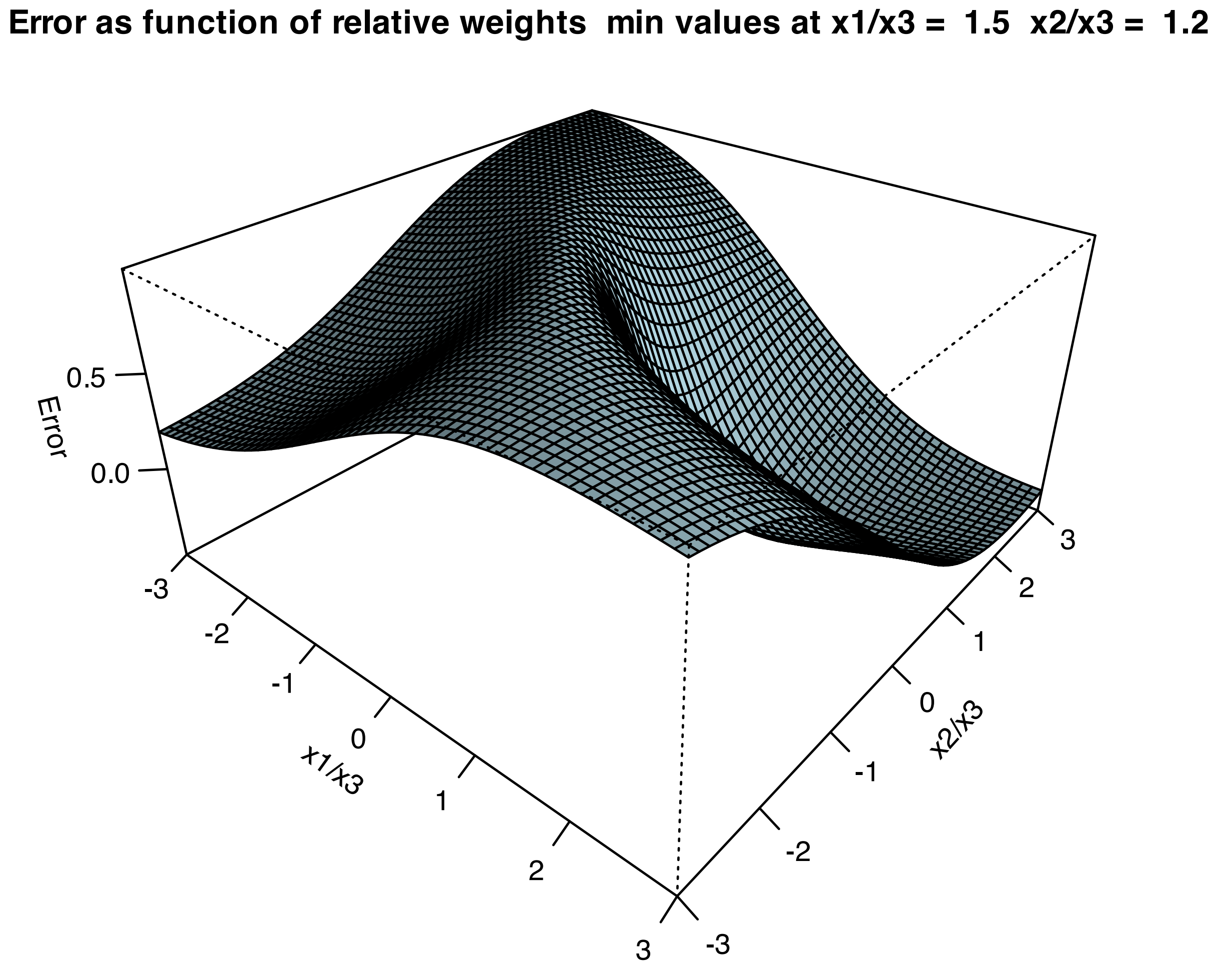
I. Although the optimal beta weights may be found precisely by multiple regression, the solution space is relatively flat and many alternative solutions are almost as good.

II. Iterative solutions can discover local minima that are far from the optimal solution

Multiple regression

	x1	x2	x3
x1	1	0	0
x2	0	1	0
x3	0	0	1

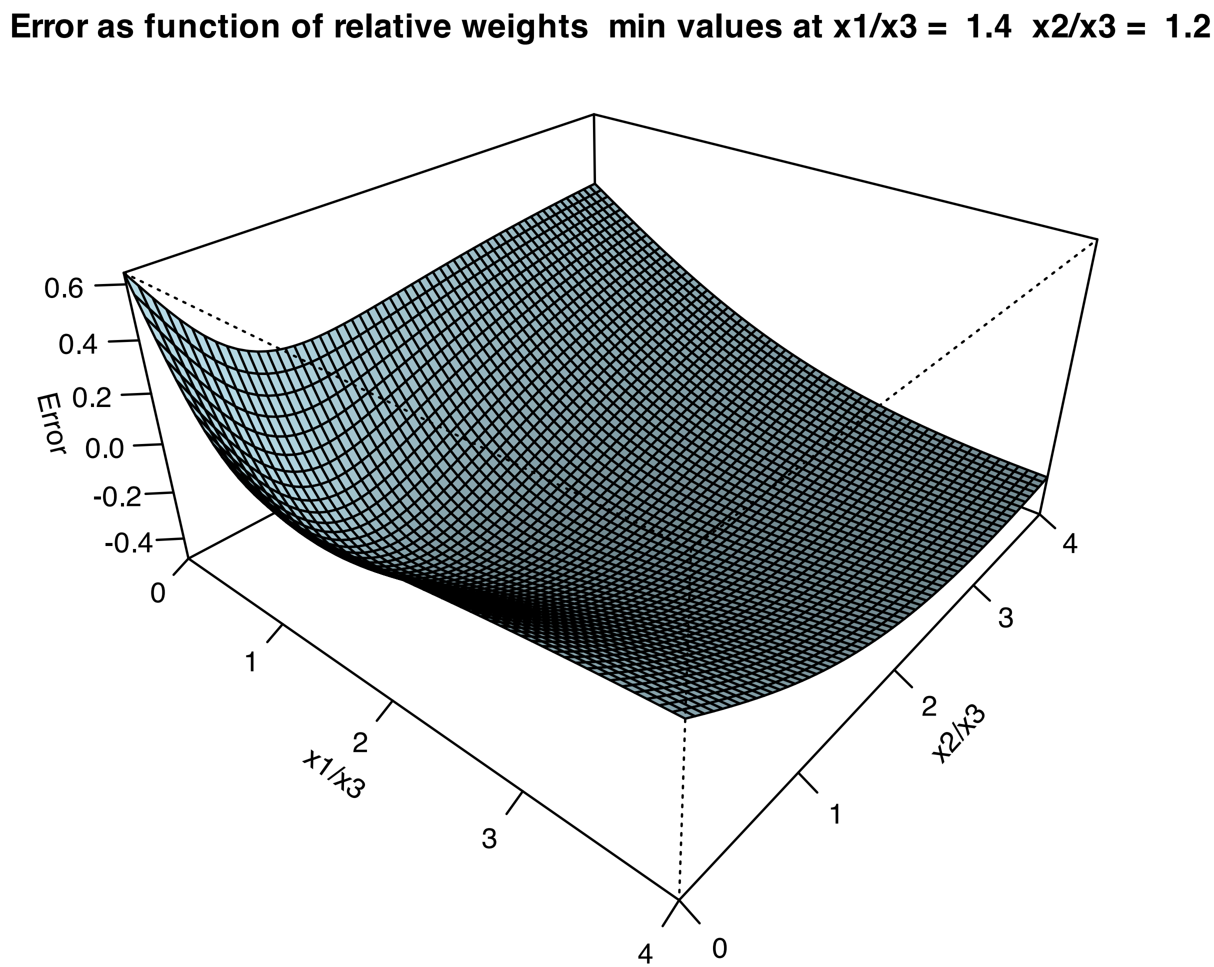
	y
x1	0.8
x2	0.7
x3	0.6



Multiple regression

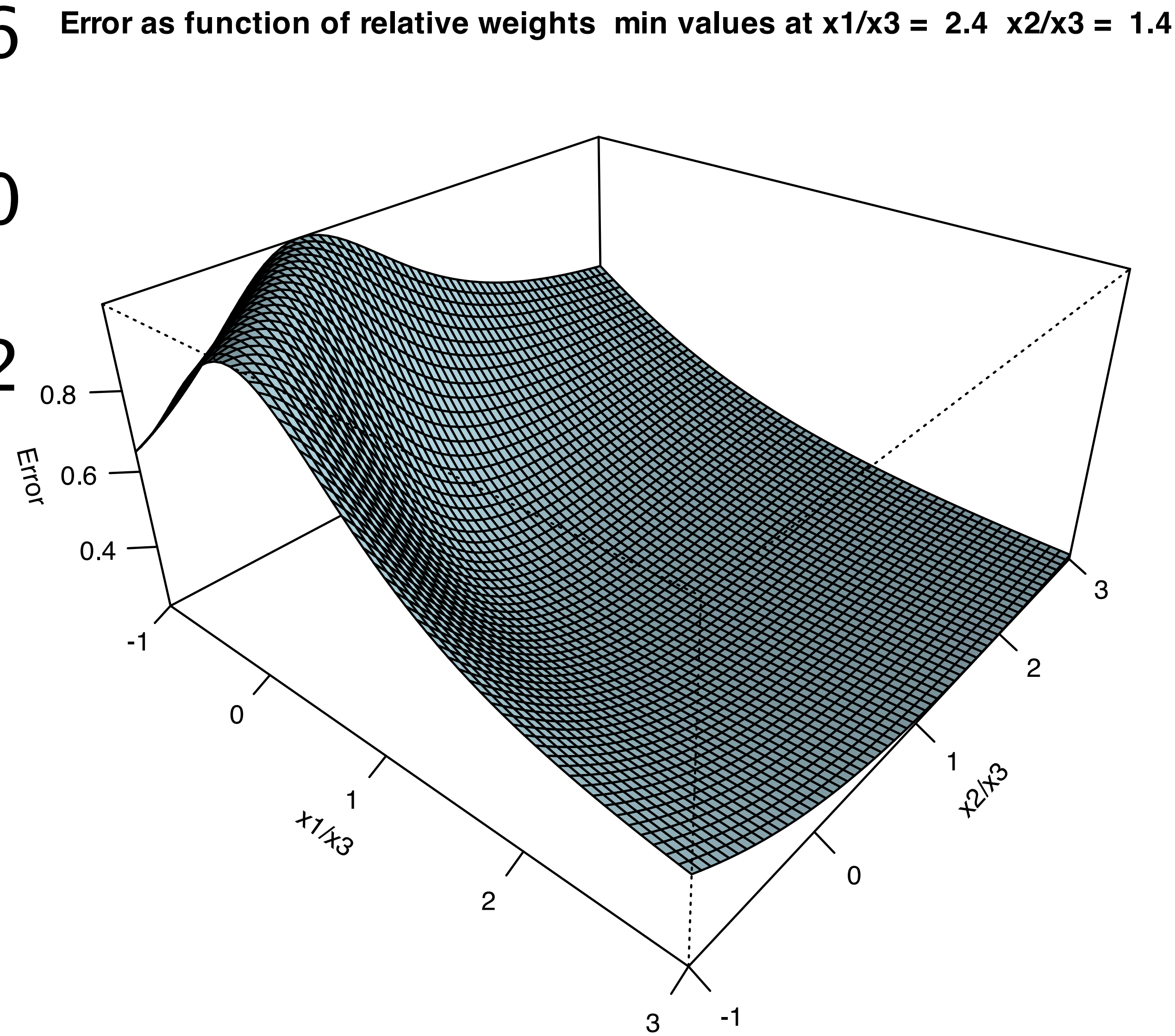
	x1	x2	x3
x1	1	0	0
x2	0	1	0
x3	0	0	1

	y
x1	0.8
x2	0.7
x3	0.6



Multiple regression

	x1	x2	x3
x1	1.00	0.56	0.48
x2	0.56	1.00	0.42
x3	0.48	0.42	1.00



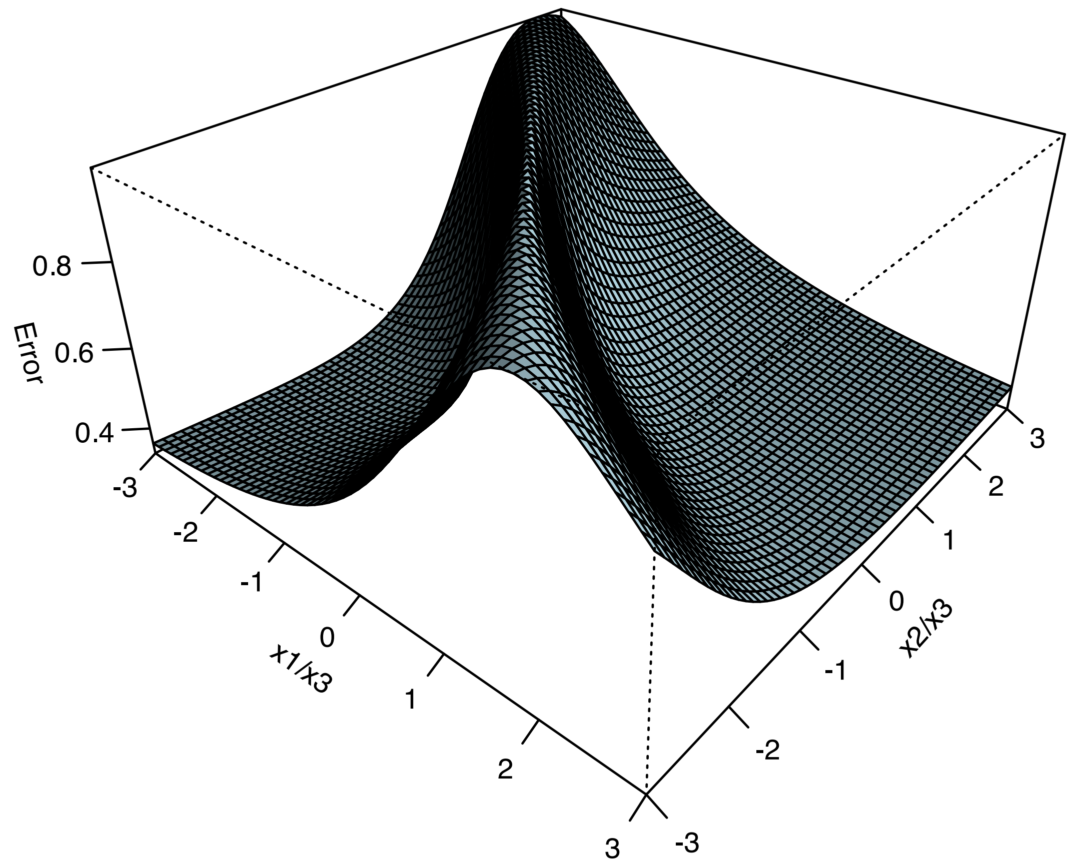
	y
x1	0.52
x2	0.32
x3	0.22

Multiple regression

	x1	x2	x3
x1	1.0	0.8	0.8
x2	0.8	1.0	0.8
x3	0.8	0.8	1.0

	y
x1	0.77
x2	0.27
x3	–0.23

Error as function of relative weights min values at x1/x3 = -2.9 x2/x3 = -1.1



Multiple Regression as a structural relation

Pattern

	X1	X2	X3
X1	1.00	0.00	0.00
X2	0.00	1.00	0.00
X3	0.00	0.00	1.00
y	0.26	0.16	0.11

Correlation

	x1	x2	x3
x1	1.00	0.56	0.48
x2	0.56	1.00	0.42
x3	0.48	0.42	1.00

=

Structure

	x1	x2	x3
X1	1.00	0.56	0.48
X2	0.56	1.00	0.42
X3	0.48	0.42	1.00
y	0.40	0.35	0.30

Pattern %*% Correlation = Structure

find the y pattern to best fit structure

Patterns, Structures, and models

I. P %*% R % P' = M

II. M - D = E

III. (M - D) %*% (M–D)' = E^2

IV. adjust elements of P and R to minimize E^2

V. In case of regression with observed variables, the beta weights minimize E^2

Pattern, structures, and models

Pattern

	X1	X2	X3
X1	1.00	0.00	0.00
X2	0.00	1.00	0.00
X3	0.00	0.00	1.00
y	0.26	0.16	0.11

	x1	x2	x3
x1	1.00	0.56	0.48
x2	0.56	1.00	0.42
x3	0.48	0.42	1.00

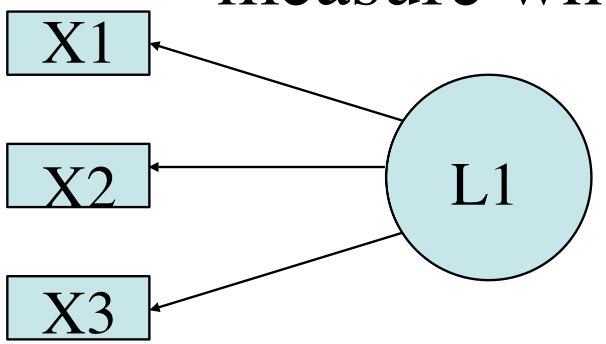
P %*% R %*% t(P)

	X1	X2	X3	y
X1	1.00	0.56	0.48	0.40
X2	0.56	1.00	0.42	0.35
X3	0.48	0.42	1.00	0.30
y	0.40	0.35	0.30	0.19

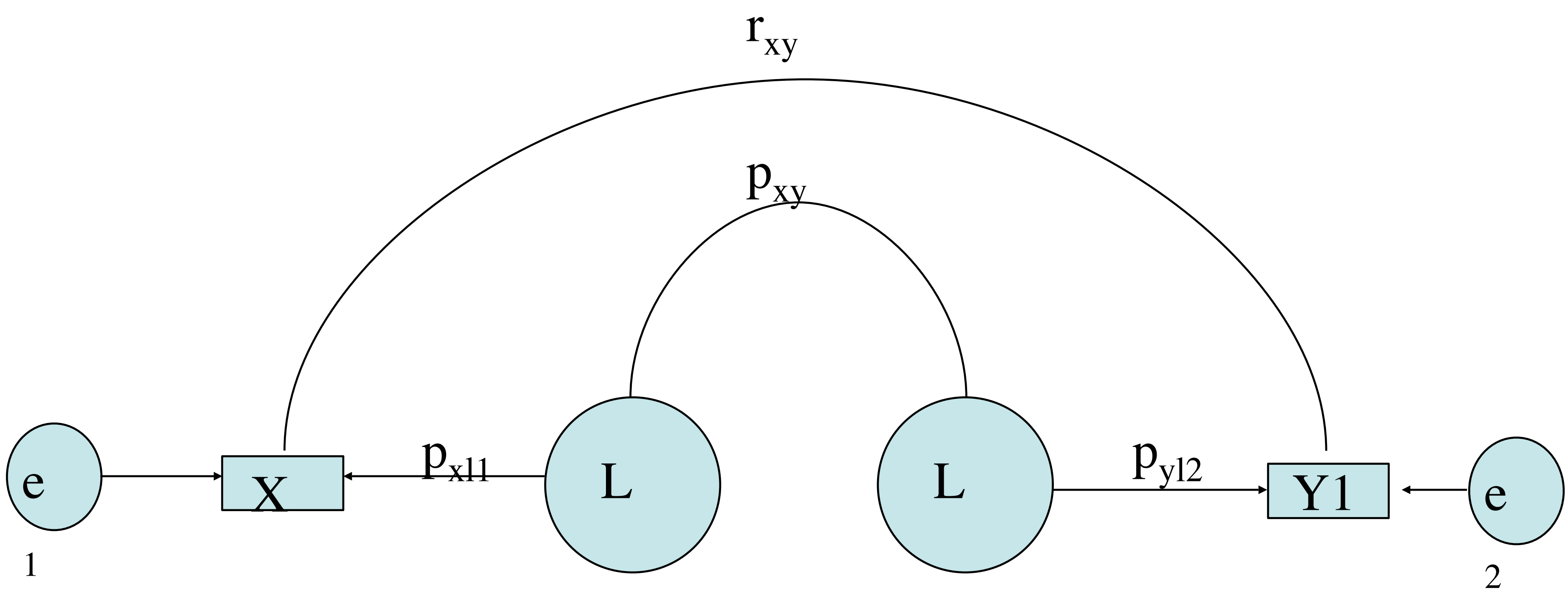
Reliabilityand latent traits

I. Classic reliability theory distinguishes between observed score and “True Score”
   A. True score as expectation given infinite replication
   B. Observed score = True Score + Error
II. True score is an unobservable, latent variable

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



Classic Reliability Theory: How well do we measure what ever we are measuring and what is the relationships between latent variables



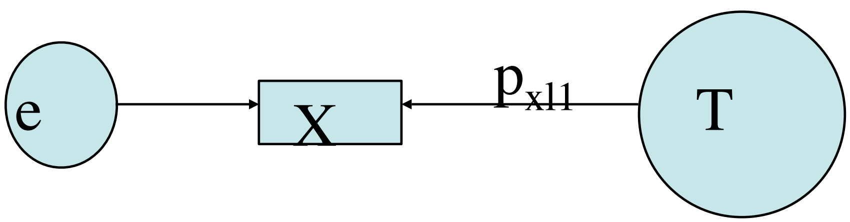
Classic Reliability Theory: How well do we measure what ever we are measuring

What is the relationship between X_1 and L_1?

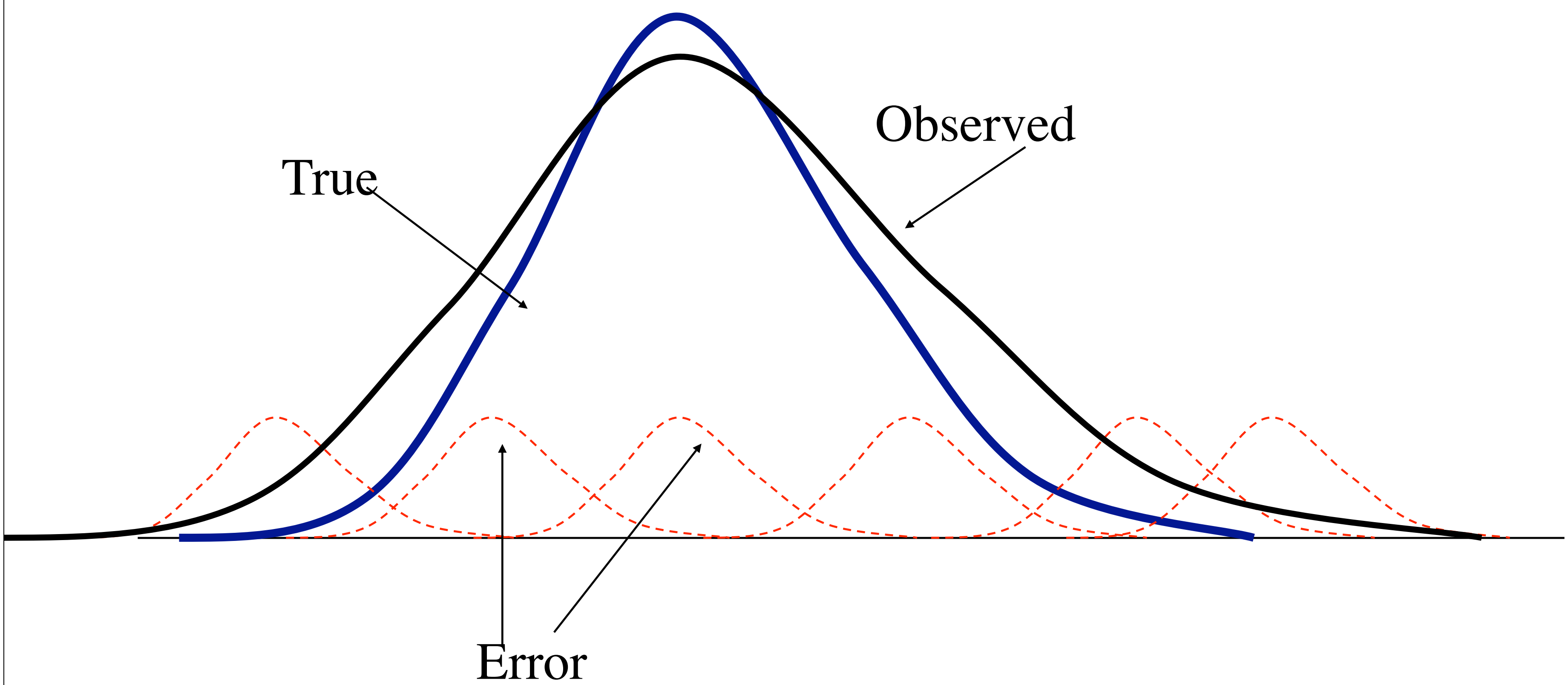
What is the variance of X_1, L_1, and E_1?

Let True Score for Subject I = expected value of X_i.

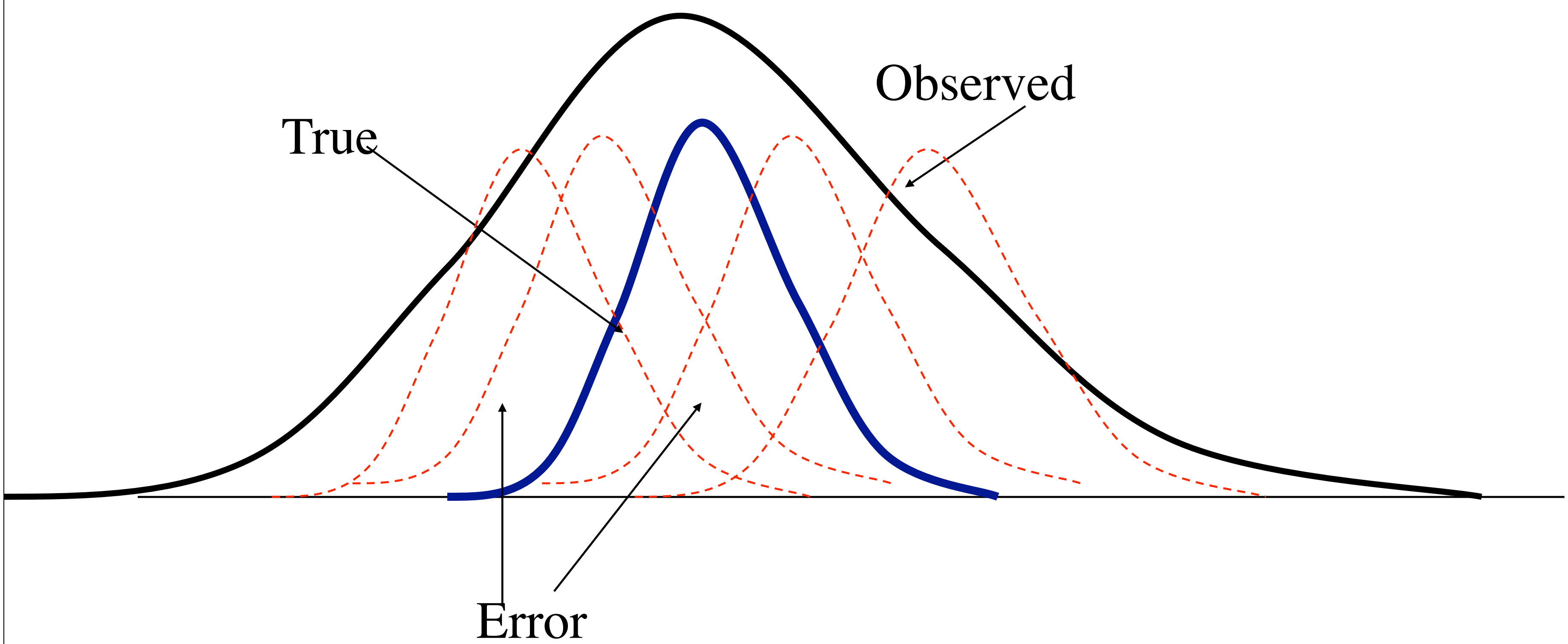
(note that this is not the Platonic Truth, but merely the average over an infinite number of trials.)



Observed= True + Error



Observed= True + Error



Observed = Truth + Error

- Define True score as expected observed score. Then Truth is uncorrelated with error, since the mean error for any True score is 0.
- Variance of Observed = Variance (T+E)=
  $V(T) + V(E) + 2Cov(T,E) = V_t+V_e$
- Covariance O,T = $Cov_{(T+E),T} = V_t$
- $p_{ot}= C_{ot}/sqrt(V_o*V_t) = V_t/ sqrt(V_o*V_t) =sqrt(V_t/ V_o)$
- $p^2_{ot} = V_t/V_o$ (the squared correlation between observed and truth is the ratio of true score variance to observed score variance)

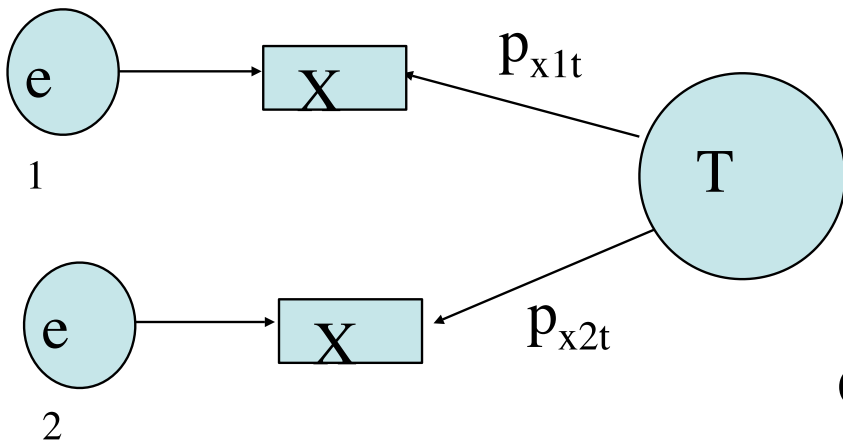
Estimating True score

- Given that $p^2_{ot} = V_t/V_o$ and p_{ot} =sqrt(V_t/ V_o), then for an observed score x, the best estimate of the true score can be found from the prediction equation:
- $z_t = p_{ox}z_x$
- The problem is, how do we find the variance of true scores and the variance of error scores?

Estimating true score: regression artifacts

- Consider the effect of reward and punishment upon pilot training:
  - From 100 pilots, reward the top 50 flyers, punish the worst 50.
  - Observation: praise does not work, blame does!
  - Explanation?

Parallel Tests



$V_{x1}=V_t+V_{e1}$
$V_{x2}=V_t+V_{e2}$

$C_{x1x2}=V_t+C_{te1}+C_{te2}+C_{e1e2}= V_t$

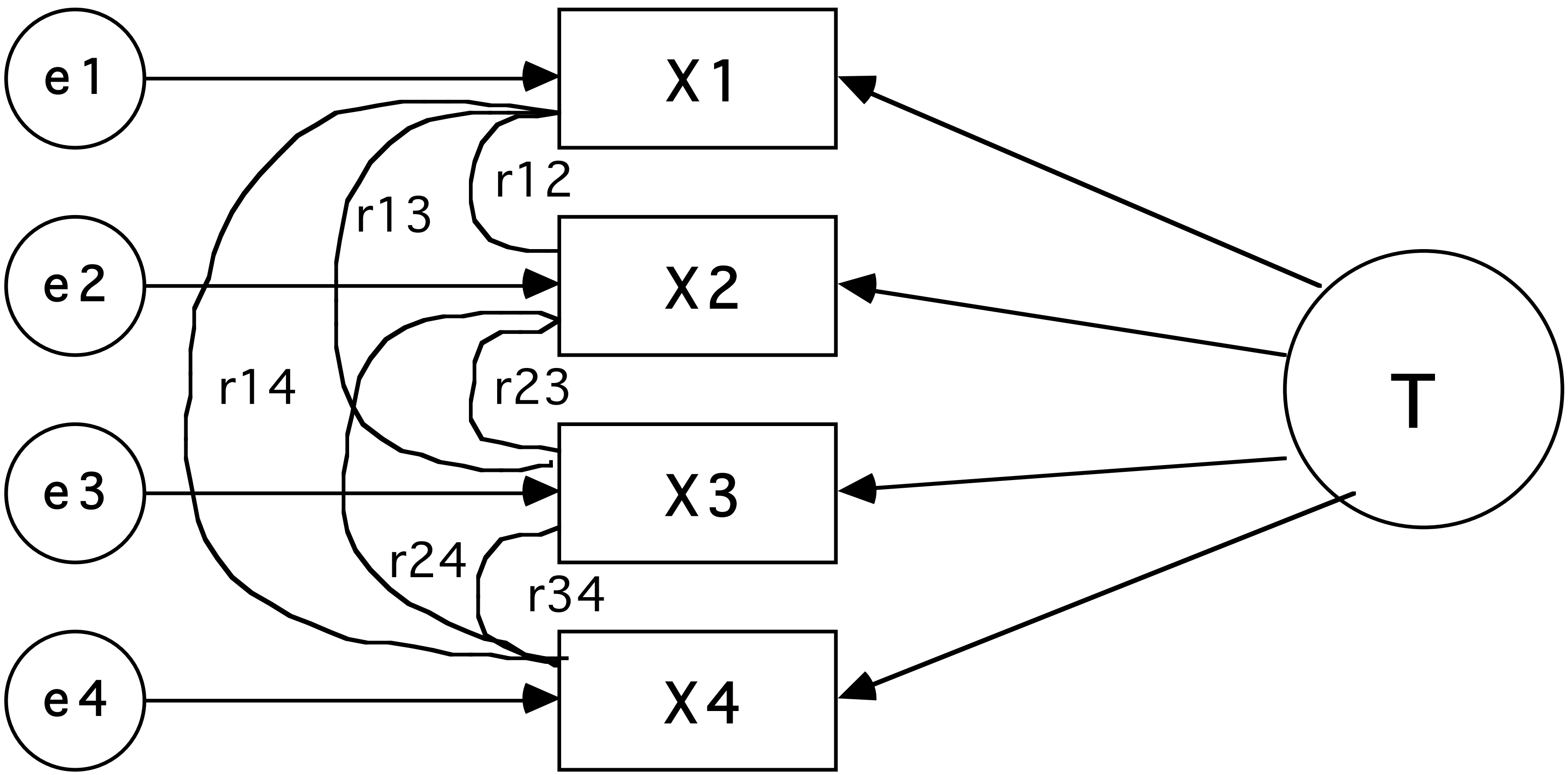
$r_{xx}=C_{x1x2}/Sqrt(V_{x1}*V_{x2}) = V_t/V_x$

The reliability of a test is the ratio of the true score variance to the observed variance = the correlation of a test with a test "just like it"

Reliability and parallel tests

- $r_{x1x2} = V_t/V_x = r_{xt}^2$
- The reliability is the correlation between two parallel tests and is equal to the squared correlation of the test with the construct. $r_{xx} = V_t/V_x$ = percent of test variance which is construct variance.
- $r_{xt} = \mathrm{sqrt}(r_{xx})$ ==> the validity of a test is bounded by the square root of the reliability.
- How do we tell if one of the two “parallel” tests is not as good as the other? That is, what if the two tests are not parallel?

Congeneric Measurement



Observed Variances/Covariances

	x_1	x_2	x_3	x_4
x_1	Vx_1			
x_2	c_{x1x2}	Vx_2		
x_3	c_{x1x3}	c_{x2x3}	Vx_3	
x_4	c_{x1x4}	c_{x3x4}	c_{x3x4}	Vx_4

Model Variances/Covariances

	x_1	x_2	x_3	x_4
x_1	V_t+Ve_1			
x_2	$c_{x1t}c_{x2t}$	V_t+ Ve_2		
x_3	$c_{x1t}c_{x3t}$	$c_{x2t}c_{x3t}$	V_t+ Ve_3	
x_4	$c_{x1t}c_{x4t}$	$c_{x3t}c_{x4t}$	$c_{x3t}c_{x4t}$	V_t+ Ve_4

Observed and modeled Variances/Covariances

	x_1	x_2	x_3	x_4
x_1	Vx_1			
x_2	c_{x1x2}	Vx_2		
x_3	c_{x1x3}	c_{x2x3}	Vx_3	
x_4	c_{x1x4}	c_{x3x4}	c_{x3x4}	Vx_4

	x_1	x_2	x_3	x_4
x_1	V_t+Ve_1			
x_2	$c_{x1t}c_{x2t}$	$V_t+ Ve_2$		
x_3	$c_{x1t}c_{x3t}$	$c_{x2t}c_{x3t}$	$V_t+ Ve_3$	
x_4	$c_{x1t}c_{x4t}$	$c_{x3t}c_{x4t}$	$c_{x3t}c_{x4t}$	$V_t+ Ve_4$

Estimating parameters of the model

1. Variances: V_t, Ve_1, Ve_2, Ve_3, Ve_4
2. Covariances: Ctx_1, Ctx_2, Ctx_3, Ctx_4
3. Parallel tests: 2 tests, 3 equations, 5 unknowns, assume $Ve_1 = Ve_2$, $Ctx_1 = Ctx_2$
4. Tau Equivalent tests: 3 tests, 6 equations, 7 unknowns, assume
    1. $Ctx_1 = Ctx_2 = Ctx_3$ but allow unequal error variance
5. Congeneric tests: 4 tests, 10 equations, 9 unknowns!