

Psychology 205: Research Methods in Psychology

Methods in Differential Psychology

Department of Psychology
Northwestern University
Evanston, Illinois USA



NORTHWESTERN
UNIVERSITY

November, 2011

Outline

- 1 The two disciplines of psychology
- 2 Data = Model + Residual
- 3 Types of designs
- 4 Methods of analysis

Two fundamentally different approaches to psychological research

- Experimental versus differential
 - Experimental approaches associated with cognitive, neurobiology, social.
 - Correlational primarily with personality/clinical/life span developmental.
- Some combine both approaches:
 - Experimental clinical/personality examines interactions of situation with individual differences.
 - Experimental developmental examines changes in response time capabilities with age.
- Should not confuse lab based data collection with experimental – experimental requires random assignment to conditions.

Differential Psychology

- ① Emphasis on dispersion (variance and covariance) rather than central tendency (mean and median)
- ② Studies of Ability, Interests/Values, Temperament
- ③ Prediction of performance in college/universities, professions, military
- ④ Correlations with marriage, divorce, employment, health and mortality
- ⑤ Correlations with psychopathology
 - Understanding the risk for depression in terms of cognitive processing
 - Understanding the genetic and environmental basis of schizophrenia

Data = Model + Residual

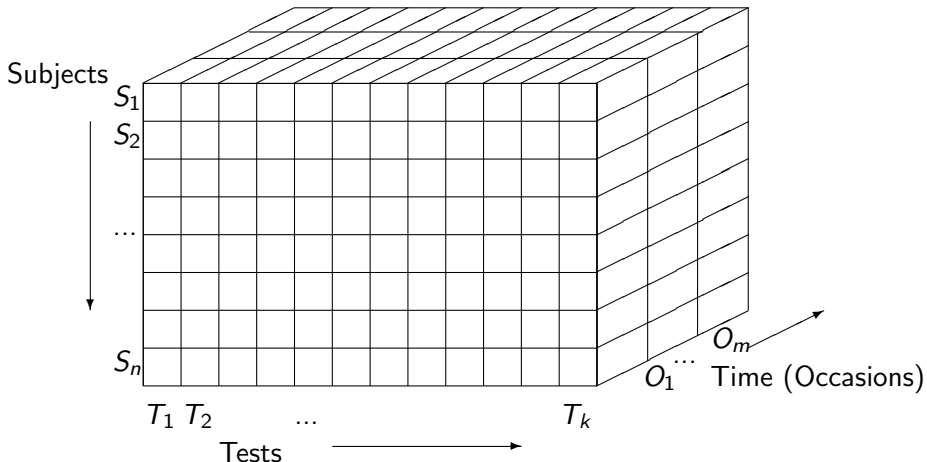
All studies have three questions:

- 1 What are the data?
 - Collected experimentally
 - Collected observationally
- 2 How do we model them?
- 3 How large is the residual variance?

Theory of Data + the Data Box

- Theory of data
 - Sets of Objects and of People
 - Two types of comparisons (Proximity and order)
 - Comparisons of preferences: Do you like X? is a proximity measure.
 - $|o_i - p_j| < \epsilon$ (Do you like something?)
 - $|o_i - p_j| < |o_k - p_l|$ (do you like something more than I like something else?)
 - Comparisons of order: Can you do X?
 - $o_i < p_j$ Are you better than an item?
 - $o_i < o_j$ Is this object more than that object?
- The Data Box: comparisons across time and space
 - People over tests, people over time
 - Tests over people, tests over time
 - etc.

The data box: Subjects x Measures x Time



Cattell (1946, 1966)

Methods of data collection

- ① Self report – how do you *normally* think, feel, act, do
 - Traits/states (normally versus in the moment)
 - adjectives/sentences/narratives
- ② Other report (peer, supervisor, subordinate)
- ③ Ability tests – what is the *best* you can do
- ④ Behavioral observation - what is the subject doing?
- ⑤ Physiology
 - above the neck EEG/MRI/fMRI/PET
 - below the neck HR/SC/BP/blood/urine
- ⑥ Telemetric
 - active: web/big EAR/text messaging
 - passive: appearance of webpages, facebook
- ⑦ National and international surveys
- ⑧ Animal – lesion/drug/observation

Types of designs used in differential psychology

- 1 Experimental
 - Lab based
 - Field based
 - Quasi-experimental designs
 - Do not have random assignment
 - Need to consider many alternative threats to validity
- 2 Observational
 - Cross sectional
 - Longitudinal (both short term as well as long term)

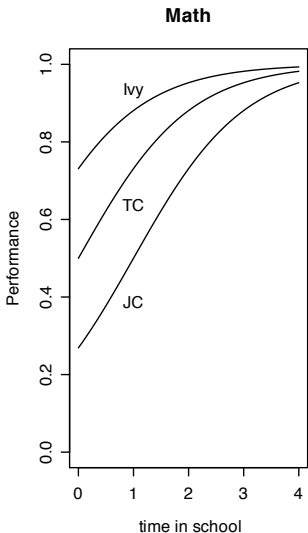
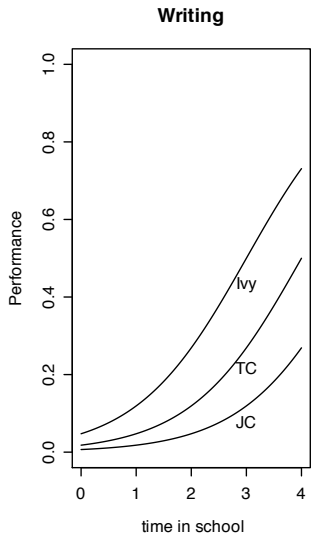
Longitudinal

- 1 Brief within subject
 - Within subject correlations of mood and behavior over time
 - Using diaries, PDAs, cell phones, text messaging, big EAR
- 2 Long term within subject
 - Terman and the “termites”
 - Berkeley/Oakland Growth Studies
 - Scottish Mental Health Study
 - Seattle Longitudinal Study
 - Dunedin longitudinal study
 - Study of Mathematically Precocious Youth

Methods of Analysis

- 1 Issues in scaling – what do the numbers mean?
 - Do the measures have linear or just ordinal properties?
 - Are we measuring the same thing at age 11 as age 77?
 - Can we adapt prior measures for current theory?
- 2 The correlation coefficient and its many forms
 - Y varies as X
 - Is Y a continuous variable?
 - is X a continuous variable?
 - what is the underlying relationship?

Issues in scaling and scaling artifacts



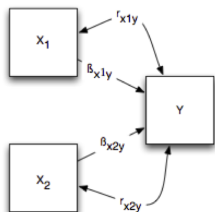
The many forms of the correlation coefficient

Table: A number of correlations are Pearson r in different forms, or with particular assumptions. If $r = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \sum y_i^2}}$, then depending upon the type of data being analyzed, a variety of correlations are found.

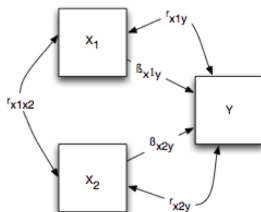
Coefficient	symbol	X	Y	Assumptions
Pearson	r	continuous	continuous	
Spearman	ρ (ρ)	ranks	ranks	
Point bi-serial	r_{pb}	dichotomous	continuous	
Phi	ϕ	dichotomous	dichotomous	
Bi-serial	r_{bis}	dichotomous	continuous	normality
Tetrachoric	r_{tet}	dichotomous	dichotomous	bivariate normality
Polychoric	r_{pc}	categorical	categorical	bivariate normality

Correlation and the multiple R

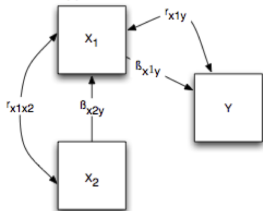
Independent Predictors



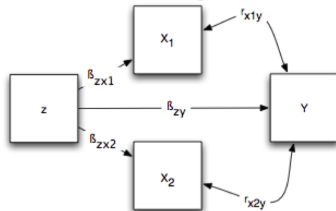
Correlated Predictors



Suppression



Missing variable



Many other statistics (F, d, b) may be expressed in terms of r

Table: Alternative Estimates of effect size. Using the correlation as a scale free estimate of effect size allows for combining experimental and correlational data in a metric that is directly interpretable as the effect of a standardized unit change in x leads to r change in standardized y.

Statistic	Estimate	r equivalent	as a function of r
Pearson correlation	$r_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y}$	r_{xy}	
Regression	$b_{y \cdot x} = \frac{C_{xy}}{\sigma_x^2}$	$r = b_{y \cdot x} \frac{\sigma_y}{\sigma_x}$	$b_{y \cdot x} = r \frac{\sigma_x}{\sigma_y}$
Cohen's d	$d = \frac{X_1 - \bar{X}_2}{\sigma_x}$	$r = \frac{d}{\sqrt{d^2 + 4}}$	$d = \frac{2r}{\sqrt{1 - r^2}}$
Hedge's g	$g = \frac{X_1 - X_2}{s_x}$	$r = \frac{g}{\sqrt{g^2 + 4(df/N)}}$	$g = \frac{2r\sqrt{df/N}}{\sqrt{1 - r^2}}$
t - test	$t = \frac{d\sqrt{df}}{2}$	$r = \sqrt{t^2 / (t^2 + df)}$	$t = \sqrt{\frac{r^2 df}{1 - r^2}}$
F-test	$F = \frac{d^2 df}{4}$	$r = \sqrt{F / (F + df)}$	$F = \frac{r^2 df}{1 - r^2}$
Chi Square		$r = \sqrt{\chi^2 / n}$	$\chi^2 = r^2 n$
Odds ratio	$d = \frac{\ln(OR)}{1.81}$	$r = \frac{\ln(OR)}{1.81\sqrt{(\ln(OR)/1.81)^2 + 4}}$	$\ln(OR) = \frac{3.62r}{\sqrt{1 - r^2}}$
<i>r</i> equivalent	r with probability p	$r = r_{equivalent}$	

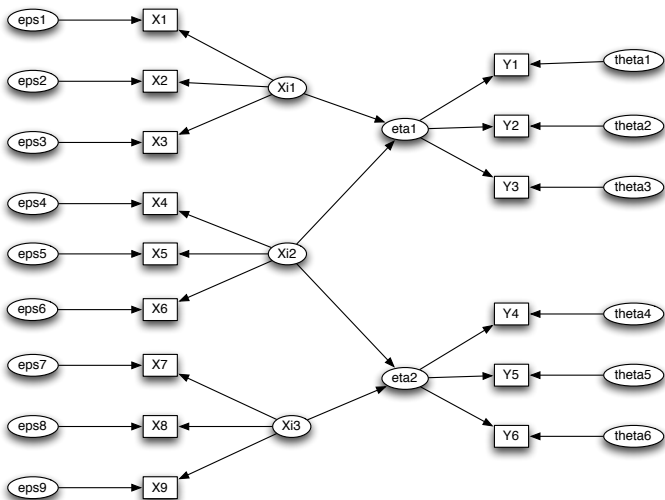
Data quality

- ① Reliability (Do multiple measures measure the same thing?)
 - Estimates of reliability
 - internal consistency ($\alpha, \omega_h, \omega_t, \gamma$)
 - stability
 - alternate form
 - parallel tests
- ② Validity: Does the arrow hit the target?
 - Concurrent Validity
 - Predictive Validity
 - Construct Validity
 - convergent (do 2 measures of the same thing measure agree?)
 - discriminant (do 2 measures of different things not agree?)
 - incremental (does it make a difference?)

Advanced methods in differential psychology

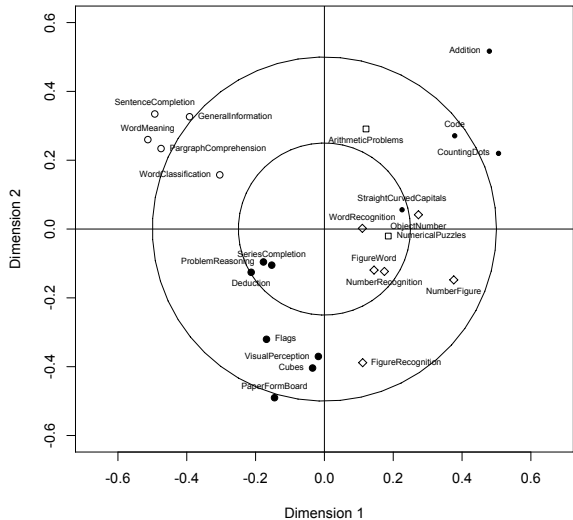
- 1 Factor analysis, cluster analysis, multidimensional scaling
 - These are attempts at reducing the complexity of the measures to a more meaningful number.
 - Instead of measuring 24 emotion words, are we just measuring 2 dimensions of emotion?
 - Instead of measuring 100 adjectives, are we measuring just “Big 5” dimensions of personality?
 - Do all cognitive tests measure a general factor of ability?
- 2 Structural Equation Modeling: Can we apply regression to latent variables?
- 3 Multilevel modeling: Measures are nested within levels:
 - students within classrooms within schools
 - emotions within people over time
 - people within couples
- 4 Computer modeling and simulation

Advanced methods in differential psychology: EFA, CFA, SEM



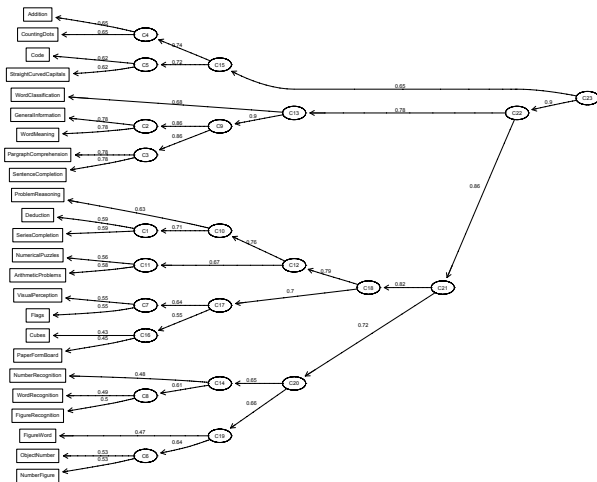
24 mental measures in a multidimensional space

Multidimensional Scaling of 24 ability tests



24 mental measures in a hierarchical space

ICLUST of 24 mental tests



Advanced methods in differential psychology within a structural framework: scaling, reliability, validity, EFA, CFA, PC, MDS, SEM

