

# An introduction to Psychometric Theory

William Revelle

Department of Psychology  
Northwestern University  
Evanston, Illinois USA



March, 2017

## Outline of Day 1/part 1

### What is psychometrics?

- Goals and Requirements

### Conceptual overview

- Theory: the organization of Observed and Latent variables

  - Data and scaling

- Correlation, Regression, Multiple Regression

### Overview

- A latent variable approach to measurement

  - Types of validity; What are we measuring

  - Structural Equation Models

  - Individual differences

### Two disciplines of scientific psychology

- Two cultures

- Two tribes within the scientific culture

### Theory testing

- The process of theory testing

### What is R? Where did it come from, why use it?

- Installing R on your computer and adding packages

## What is psychometrics?

*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)

## What is psychometrics?

*The character which shapes our conduct is a definite and durable 'something', and therefore . . . it is reasonable to attempt to measure it. (Galton, 1884)*

*The history of science is the history of measurement" (J. M. Cattell, 1893)*

*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)*

## What is psychometrics?

*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)*

Psychometrics: the assigning of numbers to observed psychological phenomena and to unobserved concepts. Evaluation of the fit of theoretical models to empirical data.

## 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.
4. To instill an appreciation of and an interest in the principles and methods of psychometric theory.
5. This course is not designed to make you into an accomplished psychometrist (one who gives tests) nor is it designed to make you a skilled psychometrician (one who constructs tests)
6. It will give you limited experience with psychometric computer programs (although all of the examples will use R, it not necessary to learn R).

## Requirements

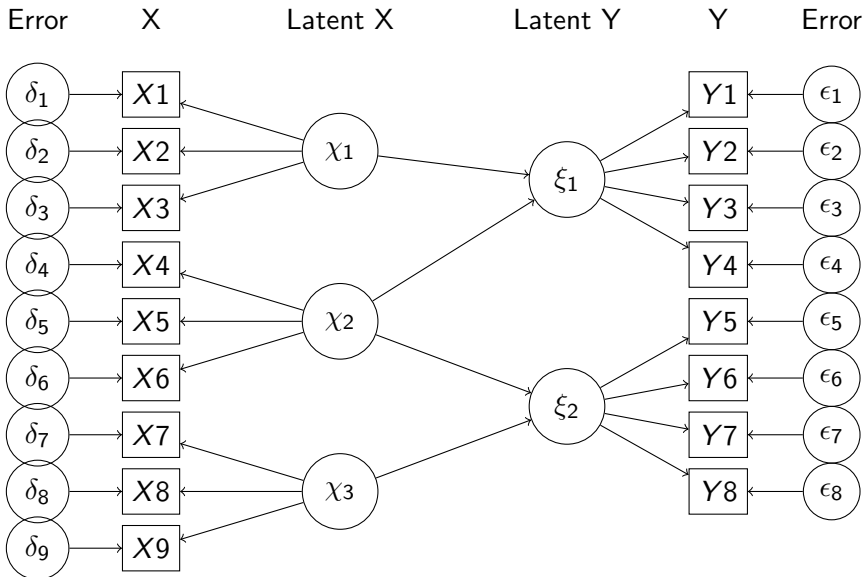
1. Asking questions!
2. Objective Midterm exam
3. Objective Final exam
4. Final paper applying principles from the course to a problem of interest to you.
5. Sporadic applied homework and data sets

## Texts (required and recommended supplements)

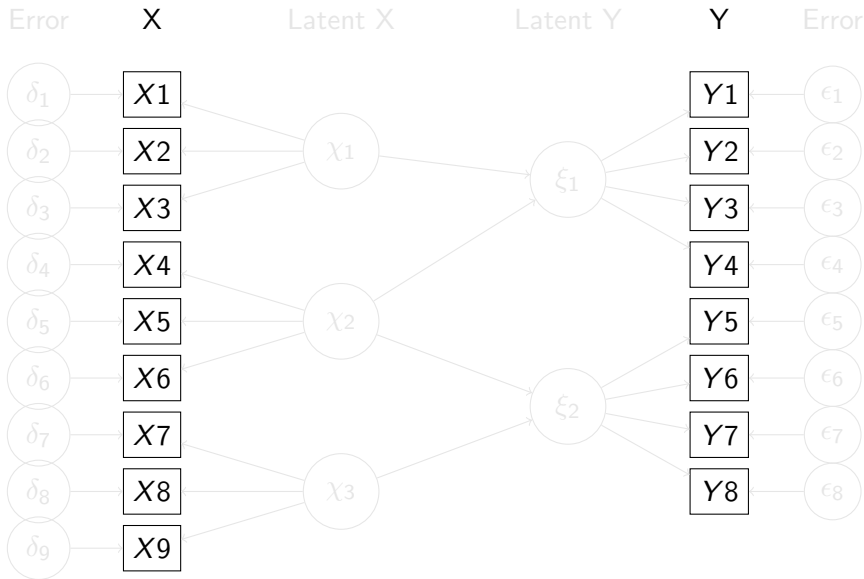
1. Revelle, W. (in prep).  
<http://www.personality-project.org/r/book/>[An introduction to psychometric theory with applications in R]. Springer Available on line and linked throughout the syllabus.
2. Loehlin (2004) Useful reading on SEM (suggested to read)
3. Nunnally & Bernstein (1994) An old classic but dreadfully out of date.



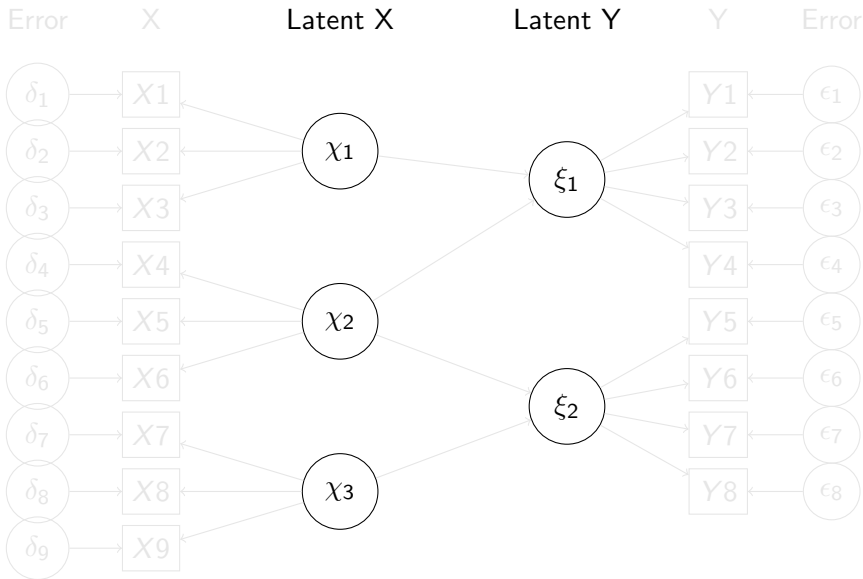
## Psychometric Theory: A conceptual Syllabus



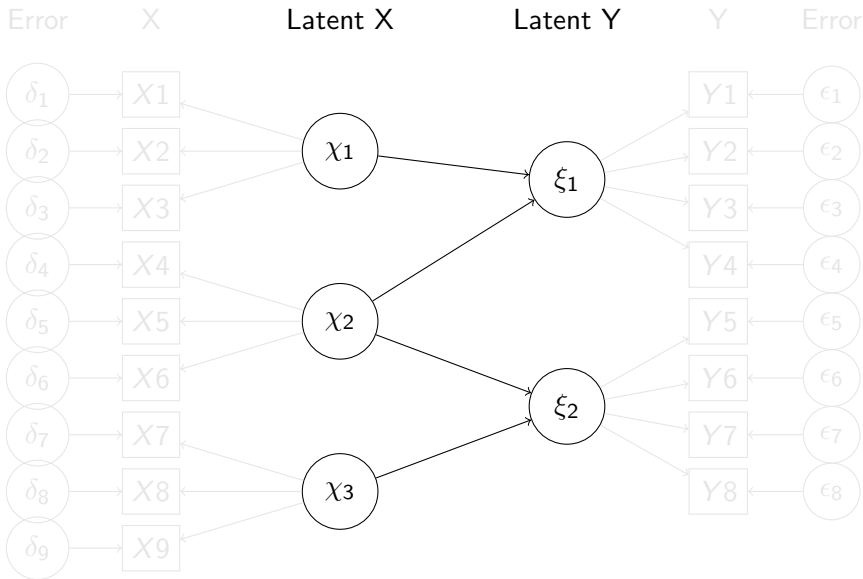
## Observed Variables



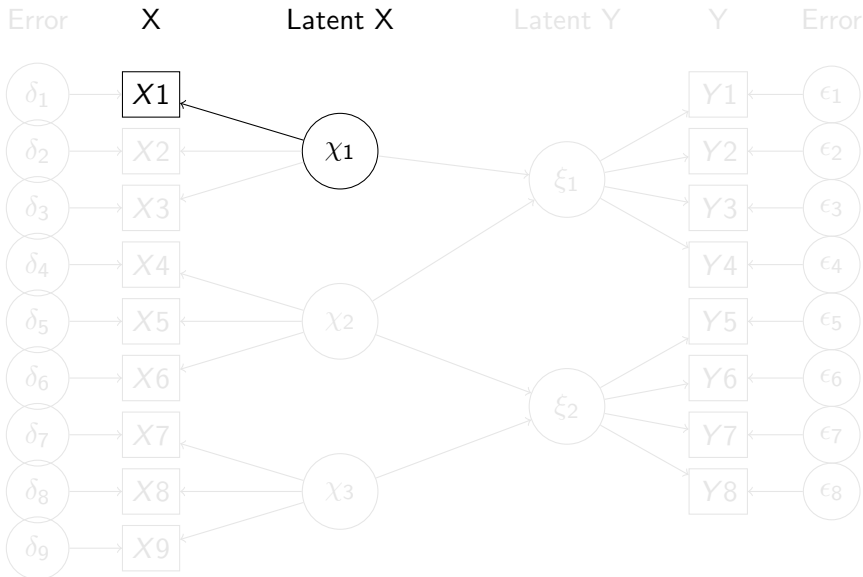
## Latent Variables



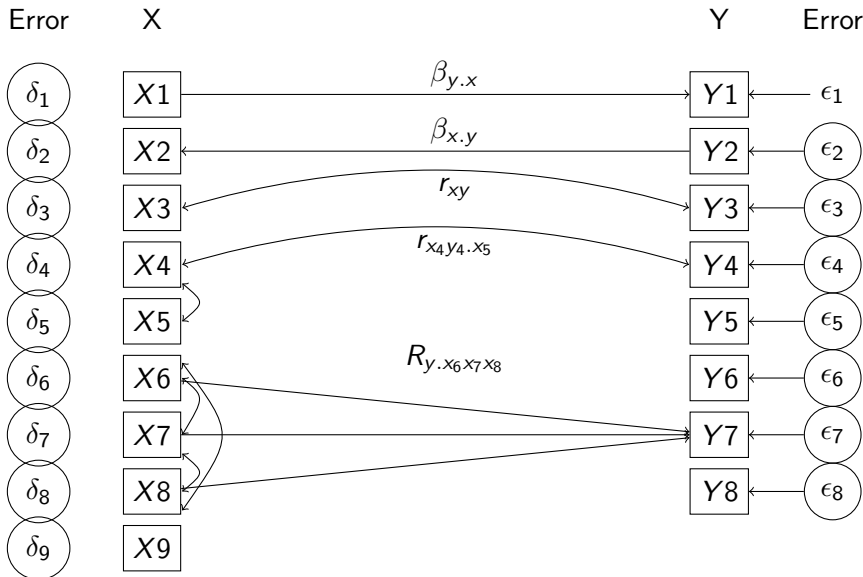
## Theory



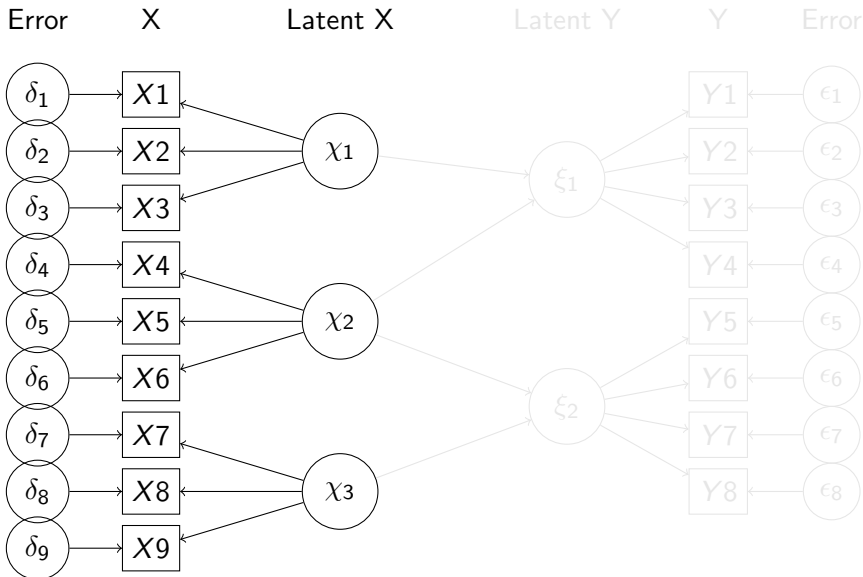
## A theory of data and fundamentals of scaling



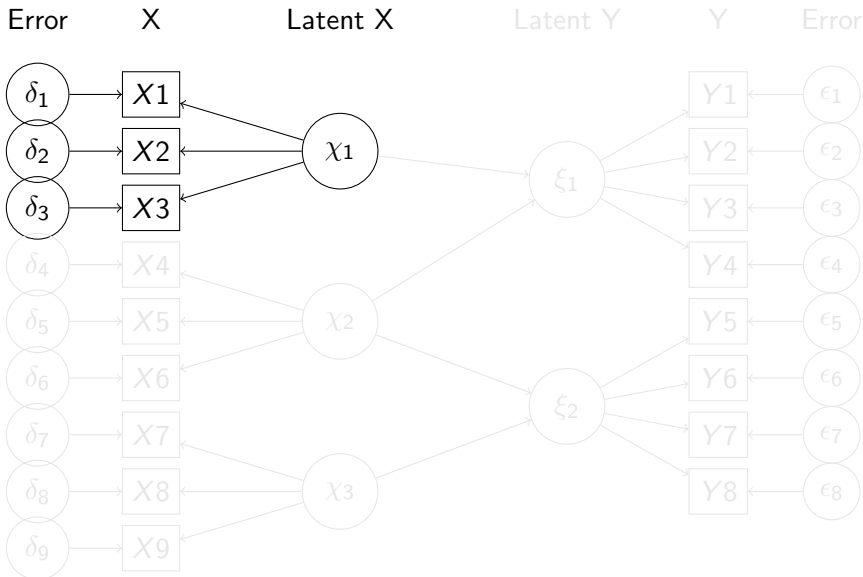
## Correlation, Regression, Partial Correlation, Multiple Regression



## Measurement: A latent variable approach.

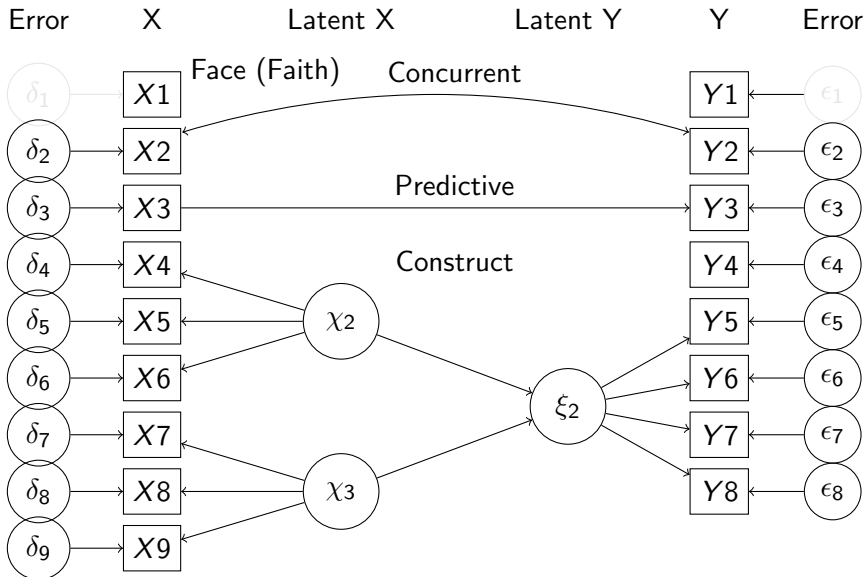


## Reliability: How well does a test reflect one latent trait?

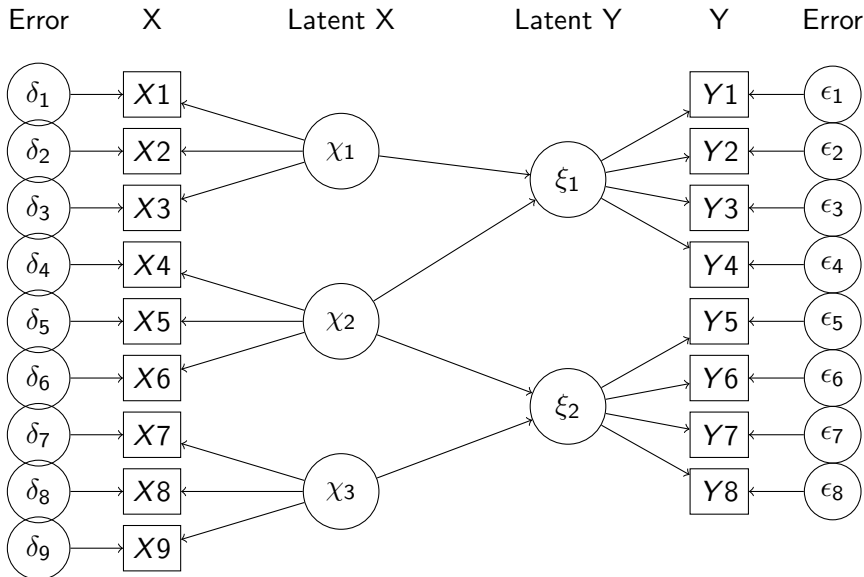




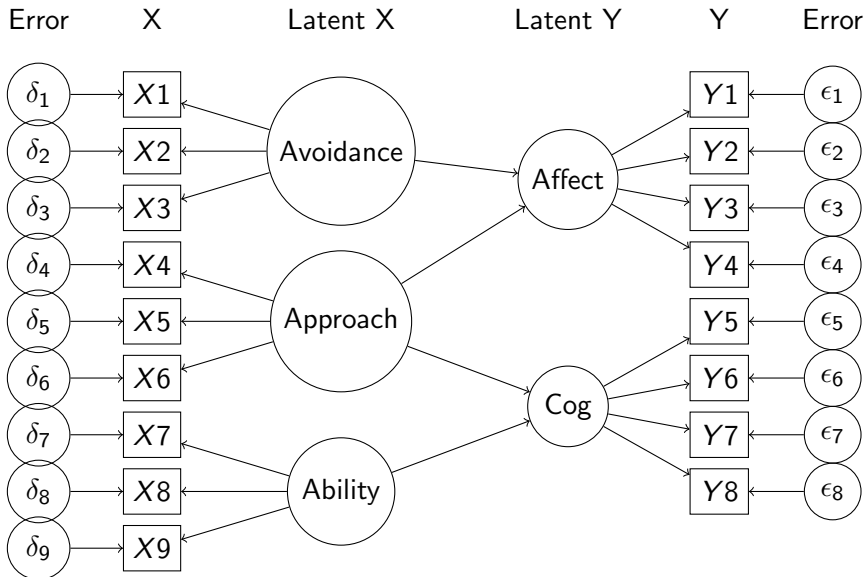
## Face, Concurrent, Predictive, Construct



## Psychometric Theory: Data, Measurement, Theory



## Psychometric Theory: Data, Measurement, Theory



## Overview

1. Overview: An introduction to Psychometric Theory
  - What is Psychometrics?
  - What is R?
2. Weeks 1-2: Theory of Data, Issues in Scaling.
3. Weeks 2-3: More than you ever wanted to know about correlation.
  - A review (?) of linear algebra.
4. Weeks 4-5 Dimension reduction through factor analysis, principal components analyze and cluster analysis.
5. Week 6-7 Classical Test Theory and Item Response Theory.
6. Week 8-9 Structural Equation Modeling and applied scale construction.

## Eysenck and the process of science

Prologue: two broad themes to be discussed and interwoven

1. The two disciplines of scientific psychology
  - 1.1 Two broad cultures of intellectual activity (Snow, 1959)
  - 1.2 Two broad cultures of psychology (Kimble, 1984)
  - 1.3 Two disciplines within scientific psychology (Cronbach, 1957, 1975) and (Eysenck, 1966, 1987a, 1997).
2. The process of theory construction and validation
  - 2.1 Science from hunch to law (Eysenck, 1976, 1985)
  - 2.2 Good theories as alive and generative: the example of theories of Extraversion.

I will emphasize the power of integrating psychometric and experimental techniques in a programmatic study of personality and individual differences.

## The two cultures of intellectual inquiry

C.P. Snow (1959) considered two cultures of intellectual inquiry:

*“I believe the intellectual life of the whole of western society is increasingly being split into two polar groups.”*  
.. *“I felt I was moving among two groups—comparable in intelligence, identical in race, not grossly different in social origin, earning about the same incomes, who had almost ceased to communicate at all, who in intellectual, moral and psychological climate had so little in common ... one might have crossed an ocean.”*

## Kimble and the two cultures of psychology

Just as Snow considered the scientific versus humanistic cultures of English and American society, so did Kimble (1984) consider two cultures of psychology: the scientific and the humanistic.

*“The remaining points of disagreement involve the items asking about most important values (scientific vs. human), source of basic knowledge (objectivism vs. intuitionism), and generality of laws (nomothetic vs. idiographic).”*

## Two competing tribes/paradigms within scientific psychology

But even within the culture of scientific psychology, we have two competing tribes who differ in their basic paradigmatic view of how to do science: the correlational vs. experimental paradigms discussed by Cronbach (1957, 1975) and Eysenck (1966, 1987a, 1997). Both pleaded for an integration of the two tribes. Neither was overly successful.

Others who have tried to reconcile these differences include Vale & Vale (1969), and Underwood (1975).

In a prior review Revelle & Oehlberg (2008) we reported that this dichotomy still continues. Today I will try to go beyond this dichotomy by showing how theory development and theory testing requires a mixture of the inductive power of correlations with the deductive power of experimental techniques. For we as individual differences psychologists are most able to unify the two disciplines.



## The conventional dichotomy of research paradigms in psychology ala Cronbach (1957, 1975) and Eysenck (1966, 1987a, 1997)

### Correlational

1. Influential founders
  - 1.1 Galton (1886)
  - 1.2 Pearson (1896)
  - 1.3 Spearman (1904)
2. Measurement of variances and covariances
  - 2.1 bivariate  $r$ ,  $\phi$ ,  $YuleQ$
  - 2.2 multivariate  $R$ , factor analysis, principal components
  - 2.3 General Linear Model and its extension to multi-level modeling
3. Addresses threats to validity by statistical “control”

### Experimental

1. Influential founders
  - 1.1 Wundt (1904)
  - 1.2 Gossett (Student, 1908)
  - 1.3 Fisher (1925)
2. Measurement of central tendencies
  - 2.1 bivariate  $t$  and  $F$
  - 2.2 multivariate MANOVA
  - 2.3 General Linear Model and its extension to multi-level modeling
3. Addresses threats to validity by randomization

## Two disciplines: two viewpoints

**Table:** The naive perspective from both sides—the other side is easy, why don't they just do it right? Our variables are complicated, well articulated, theirs are simple, just use any one.

Individual Differences

Experimental

---

Personality  
Ability

Task Performance

## The experimentalist's challenge: what to measure

### Measures

1. Giant 3
  - EPI
  - EPQ
2. Big 5
  - NEO-PI-R
  - IPIP B5
  - IPIP NEO
  - BFI
  - TIPI
3. Beyond the Big 5
  - HEXACO
  - IPIP HEXACO
  - BFAS
  - SAPA 3-6-12
  - ICAR-IQ
  - ...

### Constructs

1. Extraversion
  - but which one? Costa vs. Goldberg
2. Neuroticism
3. Agreeableness
4. Conscientiousness
5. Openness-Intellect
  - but is it openness or is it intellect?
6. Honesty/Humility
7. Impulsivity
8. Sociability
9. Trust
10. ...

## The challenge for individual difference researchers: what constructs to measure

### Memory

1. Working memory
2. Iconic memory
3. Short Term memory
4. Long Term memory
5. Semantic memory
6. Episodic memory
7. Procedural memory
8. Autobiographical memory
9. False memory
10. Recall
11. Recognition

### Attention

1. Sustained Attention
2. Allocation of Attention
3. Capturing Attention
4. Breadth of Attention
5. Local/Global Attention
6. Paying Attention

## The experimentalist's challenge: how to analyze, what to report

### Analysis

1. Dimension Reduction
  - Principal Components
  - EFA
  - CFA
2. Structure
  - Path Analysis
  - SEM
  - Latent Growth Curves
3. Reliability analysis
  - Internal Consistency
  - Alternate Form
  - Test-Retest
4. Item Response Theory

### Statistics

1. Measures of association
  - Pearson  $r$ , Spearman  $\rho$
  - $\phi$  or  $Yule_Q$
  - $r_{tetrachoric}$ ,  $r_{polychoric}$
2. Goodness of fit
  - $\chi^2$  or  $\chi^2$  difference
  - RMSEA or RMSR
  - Tucker-Lewis
  - BIC or AIC
3. Reliability
  - $\alpha$
  - $\beta$
  - $\omega_h$
  - $\omega_t$

## The challenge for individual difference researchers: which paradigm to use

### Memory

1. Reaction time
  - Sternberg Memory Scanning
  - Ratcliff choice
  - Jacoby identification
2. Accuracy
3. Serial anticipation
4. Free recall
5. Cued recognition

### Attention

1. Posner letter search
2. Erickson flanker task
3. Vigilance
4. dot probe
5. emotional "Stroop"
6. Eye tracking
7. Reaction Time

## The extra subtleties of design

### Personality

1. Item wording
2. Response alternatives
3. Appropriate sample size
4. Subject selection
  - restriction of range
5. generalization of subject characteristics

### Experimental

1. number of practice trials
2. Inter Stimulus Interval
3. Stimulus Onset Asynchrony
4. Type of randomization/counterbalancing
  - block randomization
  - complete randomization
  - counterbalancing
5. Data trimming procedures
6. Power/p-hacking

## Scientific progress and levels of theory

Eysenck (1976, 1985); Eysenck & Eysenck (1985)

### 1. Hunch

- observations
- deduction

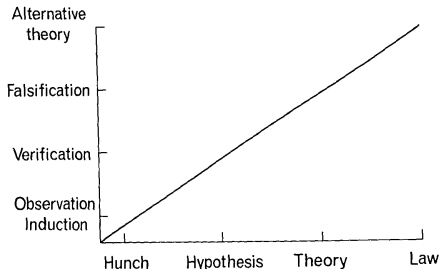
### 2. Hypothesis

- hypothesis development
- hypothesis verification

### 3. Theory

- Weak theory – confirmation studies
- Strong theory –disconfirmation studies

### 4. Law





## Eysenck, Lakatos, Popper and Kuhn

Eysenck (1983, 1985, 1987b, 1988); Eysenck & Eysenck (1985) followed Lakatos (1968) in suggesting that disconfirmation studies did not lead to theory rejection until a better theory was supplied.

*“Purely negative, destructive criticism, like ‘refutation’ or demonstration of an inconsistency does not eliminate a programme. Criticism of a programme is a long and often frustrating process and one must treat budding programmes leniently. One can, of course, undermine a research-programme but only with dogged patience. It is usually only constructive criticism which, with the help of rival research programmes can achieve major successes; but even so, dramatic, spectacular results become visible only with hindsight and rational reconstruction.”*  
(Lakatos, 1968, p 183)

## Eysenck's theory as an adaptive and changing theory of personality

Eysenck (1983) thought that the building of paradigmatic personality research required critical analysis of theory and welcomed the publications of some of his strongest critics (e.g., Gray, 1981).

*“the existence of anomalies should be no bar to the acceptance of the paradigm; the existence of such anomalies should merely act as a spur for the puzzle-solving capacities of ordinary science.”*

Indeed, in his presidential address to this society, Eysenck (1983) spent much of the time discussing Gray's criticisms and then cheerfully announced that Gray was going to replace him at the Maudsley!

## Use R



## R: Statistics for all of us

1. What is it?
2. Why use it?
3. Common (mis)perceptions of R
4. Examples for psychologists
  - graphical displays
  - basic statistics
  - advanced statistics
  - Although programming is easy in R, that is beyond the scope of today

## R: What is it?

1. R: An international collaboration
2. R: The open source - public domain version of S+
3. R: Written by statistician (and all of us) for statisticians (and the rest of us)
4. R: Not just a statistics system, also an extensible language.
  - This means that as new statistics are developed they tend to appear in R far sooner than elsewhere.
  - R facilitates asking questions that have not already been asked.
5. R: encourages publications of "Reproducible Research"
  - integrate data, code, text into one document
  - Sweave and knitr

## Statistical Programs for Psychologists

- General purpose programs
  - R
  - S+
  - SAS
  - SPSS
  - STATA
  - Systat
- Specialized programs
  - Mx
  - EQS
  - AMOS
  - LISREL
  - MPlus
  - Your favorite program

## Statistical Programs for Psychologists

- General purpose programs
  - R
  - \$+
  - \$\$
  - \$P\$\$
  - \$TATA
  - \$y\$at
- Specialized programs
  - Mx (OpenMx is part of R)
  - EQ\$
  - AMO\$
  - LI\$REL
  - MPlu\$
  - Your favorite program

## R: A way of thinking

- “R is the lingua franca of statistical research. Work in all other languages should be discouraged.” (Jan de Leeuw , 2003)
- “This is R. There is no if. Only how.” (Simon ‘Yoda’ Blomberg, 2005)
- “Overall, SAS is about 11 years behind R and S-Plus in statistical capabilities (last year it was about 10 years behind) in my estimation.” (Frank Harrell, 2003)
- “I quit using SAS in 1991 because my productivity jumped at least 20% within one month of using S-Plus.” (Frank Harrell, 2003)

Taken from the R.-fortunes (selections from the R.-help list serve)



## More fortunes

“You must realize that R is written by experts in statistics and statistical computing who, despite popular opinion, do not believe that everything in SAS and SPSS is worth copying. Some things done in such packages, which trace their roots back to the days of punched cards and magnetic tape when fitting a single linear model may take several days because your first 5 attempts failed due to syntax errors in the JCL or the SAS code, still reflect the approach of “give me every possible statistic that could be calculated from this model, whether or not it makes sense”. The approach taken in R is different. The underlying assumption is that the user is thinking about the analysis while doing it. ” (Douglas Bates, 2007)

## R is open source, how can you trust it?

- Q: “When you use it [R], since it is written by so many authors, how do you know that the results are trustable?”
- A: “The R engine [...] is pretty well uniformly excellent code but you have to take my word for that. Actually, you don't. The whole engine is open source so, if you wish, you can check every line of it. If people were out to push dodgy software, this is not the way they'd go about it.” (Bill Venables, 2004)
- “It's interesting that SAS Institute feels that non-peer-reviewed software with hidden implementations of analytic methods that cannot be reproduced by others should be trusted when building aircraft engines.” – Frank Harrell (in response to the statement of the SAS director of technology product marketing: “We have customers who build engines for aircraft. I am happy they are not using freeware when I get on a jet.”) R-help (January 2009)

## What is R?: Technically

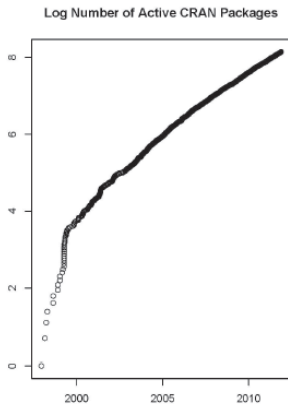
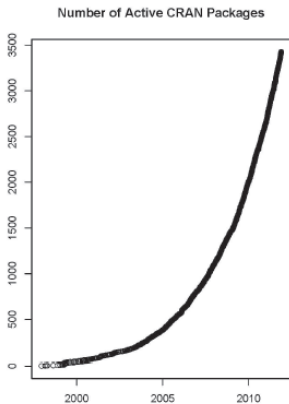
- R is an open source implementation of S (S-Plus is a commercial implementation)
- R is available under GNU Copy-left
- The current version of R is 3.3.3
- R is a group project run by a core group of developers (with new releases  $\approx$  semiannually)
- R 3.4.0 is to be released sometime in late April, 2017

(Adapted from Robert Gentleman)

## R: A brief history

- 1991-93: Ross Ihaka and Robert Gentleman begin work on R project for Macs at U. Auckland (S for Macs).
- 1995: R available by ftp under the General Public License.
- 96-97: mailing list and R core group is formed.
- 2000: John Chambers, designer of S joins the Rcore (wins a prize for best software from ACM for S)
- 2001-2016: Core team continues to improve base package with a new release every 6 months (now more like yearly).
- Many others contribute “packages” to supplement the functionality for particular problems.
  - 2003-04-01: 250 packages
  - 2004-10-01: 500 packages
  - 2007-04-12: 1,000 packages
  - 2009-10-04: 2,000 packages
  - 2011-05-12: 3,000 packages
  - 2012-08-27: 4,000 packages
  - 2014-05-16: 5,547 packages (on CRAN) + 824 bioinformatic packages on BioConductor
  - 2015-05-20 6,678 packages (on CRAN) + 1,024 bioinformatic packages + ?,000s on GitHub/R-Forge
  - 2016-03-21 8,120 packages (on CRAN) + 1,104 bioinformatic packages + ?,000s on GitHub/R-Forge
  - 2017-03-26 10,327 packages (CRAN) + 1,296 on BioConductor + ?,000s on GitHub

## Has R grown too much? Exponential growth rate continues



See also <http://r4stats.com/articles/popularity/>

## Misconception: R is hard to use

1. R doesn't have a GUI (Graphical User Interface)
  - Partly true, many use syntax
  - Partly not true, GUIs exist (e.g., R Commander, R-Studio)
  - Quasi GUIs for Mac and PCs make syntax writing easier
2. R syntax is hard to use
  - Not really, unless you think an iPhone is hard to use
  - Easier to give instructions of 1-4 lines of syntax rather than pictures of what menu to pull down.
  - Keep a copy of your syntax, modify it for the next analysis.
3. R is not user friendly: A personological description of R
  - R is introverted: it will tell you what you want to know if you ask, but not if you don't ask.
  - R is conscientious: it wants commands to be correct.
  - R is not agreeable: its error messages are at best cryptic.
  - R is stable: it does not break down under stress.
  - R is open: new ideas about statistics are easily developed.

## Misconceptions: R is hard to learn

1. With a brief web based tutorial  
<http://personality-project.org/r>, 2nd and 3rd year undergraduates in psychological methods and personality research courses are using R for descriptive and inferential statistics and producing publication quality graphics.
2. More and more psychology departments are using it for graduate and undergraduate instruction.
3. R is easy to learn, hard to master
  - R-help newsgroup is very supportive
  - Multiple web based and pdf tutorials see (e.g., <http://www.r-project.org/>)
  - Short courses using R for many applications
4. Books and websites for SPSS and SAS users trying to learn R (e.g., <http://oit.utk.edu/scc/RforSAS&SPSSusers.pdf> by Bob Muenchen).

## Ok, how do I get it: Getting started with R

- Download from R Cran (<http://cran.r-project.org/>)
  - Choose appropriate operating system and download compiled R
- Install R (current version is 3.3.3 ) with 3.4.0 coming this spring
- Start R
- Add useful packages (just need to do this once)
  - `install.packages("ctv")` #this downloads the task view package
  - `library(ctv)` #this activates the ctv package
  - `install.views("Psychometrics")` #among others
  - Take a 5 minute break
- Activate the package(s) you want to use today (e.g., *psych*)
  - `library(psych)` #necessary for most of today's examples
  - `library(sem)` #will be used for a few examples
- Use R
- (See detailed tutorial at <https://personality-project.org/r/r.guide.html#gettingstarted>)



## Start up R and get ready to play (Mac version)

```
R Under development (unstable) (2017-03-05 r72309) -- "Unsuffered Consequences"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (64-bit)
```

```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
```

```
Natural language support but running in an English locale
```

```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
```

```
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

```
[R.app GUI 1.69 (7325) x86_64-apple-darwin13.4.0]
```

```
[Workspace restored from /Users/WR/.RData]
```

```
[History restored from /Users/WR/.Rapp.history]
```

```
Good afternoon Bill.
```

```
Are you ready to have fun?
```

## Annotated installation guide: don't type the >

- > `install.packages("ctv")`
  - Install the task view installer package. You might have to choose a “mirror” site.
- > `library(ctv)`
  - Make it active
- > `install.views("Psychometrics")`
  - Install all the packages in the “Psychometrics” task view. This will take a few minutes.
- #or just install a few packages*
  - Or, just install one package (e.g., psych)
- > `install.packages("psych")`
  - as well as a few suggested packages that add functionality for factor rotation, multivariate normal distributions, etc.
- > `install.packages("GPArotation")`
- > `install.packages("MASS")`
- > `install.packages("mvtnorm")`

## Check the version number for R (should be $\geq 3.3.3$ and for psych ( $\geq 1.7.3$ ))

```
> library(psych)
```

```
> sessionInfo()
```

```
R Under development (unstable) (2017-03-05 r72309)
```

```
Platform: x86_64-apple-darwin13.4.0 (64-bit)
```

```
Running under: macOS Sierra 10.12.3
```

```
Matrix products: default
```

```
BLAS: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRblas.0.dyl
```

```
LAPACK: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRlapack.d
```

```
locale:
```

```
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
```

```
attached base packages:
```

```
[1] stats      graphics  grDevices  utils      datasets  methods   base
```

```
other attached packages:
```

```
[1] psych_1.7.3.21
```

```
loaded via a namespace (and not attached):
```

```
[1] compiler_3.4.0  tools_3.4.0      parallel_3.4.0  foreign_0.8-67  nlme_3.1-13
```

## R is extensible: The use of “packages”

- More than 10,327 packages are available for R (and growing daily)
- Can search all packages that do a particular operation by using the `sos` package
  - `install.packages("sos")` #if you haven't already
  - `library(sos)` # make it active once you have it
    - `findFn("X")` #will search a web data base for all packages/functions that have "X"
    - `findFn("principal components analysis ")` #will return 2,554 matches and reports the top 400 and download 385 links to 174 packages
    - `findFn("Item Response Theory")` # will return 503 matches with 326 links in 76 packages
    - `findFn("INDSCAL ")` # will return 20 matches in 5 packages.
- `install.packages("X")` will install a particular package (add it to your R library – you need to do this just once)
- `library(X)` #will make the package X available to use if it has been installed (and thus in your library)

## A small subset of very useful packages

- General use
  - core R
  - MASS
  - lattice (core)
  - lme4 (core)
  - psych
  - Zelig
- Special use
  - ltm
  - sem
  - lavaan
  - OpenMx
  - GPArotation
  - mvtnorm
  - > 10,327 known
  - + ?
- General applications
  - most descriptive and inferential stats
  - Modern Applied Statistics with S
  - Lattice or Trellis graphics
  - Linear mixed-effects models
  - Personality and psychometrics
  - General purpose toolkit
- More specialized packages
  - Latent Trait Model (IRT)
  - SEM and CFA (multiple groups)
  - SEM and CFA (multiple groups )
  - SEM and CFA (multiple groups +)
  - Jennrich + Browne rotations
  - Multivariate distributions
  - Thousands of more packages on CRAN
  - Code on webpages/journal articles

## Implementations

1. Base R in the Unix/Linux/Mac X11 framework
2. Base R on the Mac/PC
  - Mac has prompts at bottom of window
3. Graphical User Interfaces
  - R Commander
  - R studio as a convenient shell

# R Commander (by John Fox) has a basic GUI

The image displays the R Commander interface, which includes a 3D scatter plot of prestige (y-axis, 0-80) against education (x-axis, 0-14) and income (z-axis, 0-30000). The data points are colored by type, and a blue surface represents the fitted model. Below the 3D plot are three 2D effect plots: 'education effect plot', 'income effect plot', and 'type effect plot'. The 'education effect plot' shows prestige vs education with a solid black line and two dashed red lines. The 'income effect plot' shows prestige vs income with a solid black line and two dashed red lines. The 'type effect plot' shows prestige vs type with a solid black line and error bars for each category.

The R Commander window shows the following code in the R Script pane:

```

rattle.device(theme="cool.vistabg")
plot(allEffects(LinearModel1), ask=FALSE)
library(r2lmp, pos=4)

library(mgcv, pos=4)
> acatter3d(prestige~education+income, data=Prestige, fit="additive",
+ residuals=TRUE, bg="white", axis.scales=TRUE, grid=TRUE, ellipsoid=TRUE)
  
```

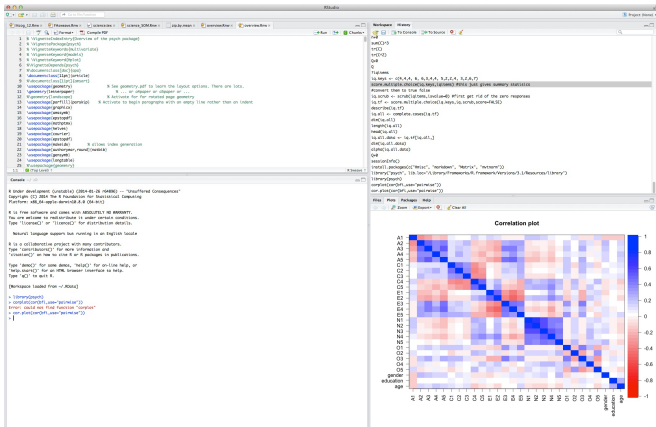
The Linear Model dialog box is open, showing the model formula `prestige ~ education + income + type` and the following options:

- Operation (click to formula): `+`, `*`, `/`, `%>`, `~`, `-`, `()`
- Splines/Polynomials (select variable and click): B-spline, natural spline, orthogonal polynomial, raw polynomial
- df for splines: 5
- deg for polynomial: 2
- Subset expression: <all valid cases>

The Prestige data table is visible at the bottom of the window:

	education	income	women	prestige	conserv	type
gov.administrators	13.11	12351	13.16	68.8	1113	prof
general.managers	12.26	23879	4.02	69.1	1130	prof
mechanics	12.77	9271	15.70	63.4	1171	prof
publishing.officers	11.42	8365	9.31	56.8	1178	prof
chemists	14.42	8403	11.68	73.5	2111	prof
physicists	15.44	11030	9.13	77.6	2113	prof
biologists	15.09	8238	25.65	72.3	2143	prof
scholars	15.44	14143	2.69	78.1	2141	prof
civil.engineers	14.52	11377	1.03	73.1	2143	prof
mining.engineers	14.64	11023	0.94	65.8	2153	prof
scavengers	12.39	5902	1.91	62.0	2161	prof
draughtsmen	12.30	7059	7.83	60.0	2163	prof
computer.programmers	13.83	8425	15.33	53.8	2183	prof
accountants	14.44	8048	37.31	62.2	2311	prof
psychologists	14.36	7405	48.28	74.9	2315	prof
social.workers	14.21	6336	54.77	55.1	2331	prof
lawyers	15.77	8243	5.13	82.3	2343	prof
librarians	14.15	6112	77.10	58.1	2351	prof
vocational.counsellors	15.22	9593	34.89	58.3	2391	prof
bankers	14.50	4888	4.14	72.8	2511	prof
university.teachers	15.37	12480	19.59	84.6	2711	prof
primary.school.teachers	13.42	5648	83.78	59.6	2731	prof
secondary.school.teachers	15.08	8034	46.80	66.1	2733	prof
chiropractors	14.00	74308	10.54	67.3	2733	prof

# RStudio (particularly nice for PCs)





## Using R

1. Install the relevant packages (just once!)
  - Either one at a time, or by using a “task view”
2. Make the packages you want to use “active” by library(package name) e.g., library(psych)
  - For each session
  - Can be automatized
3. Use the functions in a package
  - To see all functions in a package go to the index of the package or use the objects function: e.g., objects(package:psych)
  - Apply a function to data
  - All functions require an object to act upon. Most require this in parentheses. All functions return an object. This may be saved for later.
    - `function(object) #apply the function to the object, show result`
    - `sqrt(2)`
    - `result <- function(object) #apply the function to object, save result`
    - `answer <- alpha(ability) #lots and lots of output is saved`

## What makes R useful?

1. The basic philosophy of open source allows one to see the code and modify it.
2. Object oriented meaning that all functions return objects that can be acted upon by other functions.
  - e.g., output of `mean` and `sd` is included in `describe` or `describeBy`
  - Output of `cor` is used in `fa`, `alpha`, etc.
  - Can add other functions (e.g. `tetrachoric` to do correlations instead of `cor`)
3. R is vectorized, so that if it works for  $X$ , it will work for  $\vec{X}$

## Basic R commands – remember don't enter the >

R is just a fancy calculator. Add, subtract, sum, products, group

```
> 2 + 2
```

```
[1] 4
```

```
> 3^4
```

```
[1] 81
```

```
> sum(1:10)
```

```
[1] 55
```

```
> prod(c(1, 2, 3, 5, 7))
```

```
[1] 210
```

It is also a statistics table ( the normal distribution, the t distribution, F,  $\chi^2$ , ...)

```
> pnorm(q = 1) #probability of normal value > 1
```

```
[1] 0.8413447
```

```
> pt(q = 2, df = 20) #probability of t > 2 with 20 df
```

```
[1] 0.9703672
```

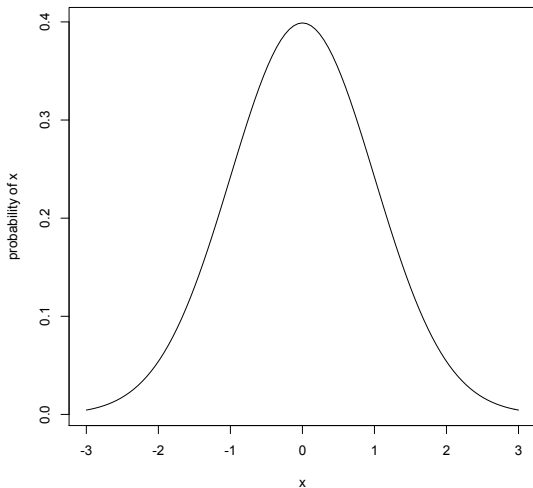
## R is a set of distributions. Don't buy a stats book with tables!

**Table:** To obtain the density, prefix with  $d$ , probability with  $p$ , quantiles with  $q$  and to generate random values with  $r$ . (e.g., the normal distribution may be chosen by using `dnorm`, `pnorm`, `qnorm`, or `rnorm`.)

Distribution	base name	P 1	P 2	P 3	example application
<i>Normal</i>	<code>norm</code>	mean	sigma		Most data
<i>Multivariate normal</i>	<code>mvnorm</code>	mean	$r$	sigma	Most data
<i>Log Normal</i>	<code>lnorm</code>	log mean	log sigma		income or reaction time
<i>Uniform</i>	<code>unif</code>	min	max		rectangular distributions
<i>Binomial</i>	<code>binom</code>	size	prob		Bernuilli trials (e.g. coin flips)
<i>Student's t</i>	<code>t</code>	df		nc	Finding significance of a t-test
<i>Multivariate t</i>	<code>mvt</code>	df	corr	nc	Multivariate applications
<i>Fisher's F</i>	<code>f</code>	df1	df2	nc	Testing for significance of F test
$\chi^2$	<code>chisq</code>	df		nc	Testing for significance of $\chi^2$
<i>Exponential</i>	<code>exp</code>	rate			Exponential decay
<i>Gamma</i>	<code>gamma</code>	shape	rate	scale	distribution theory
<i>Hypergeometric</i>	<code>hyper</code>	$m$	$n$	$k$	
<i>Logistic</i>	<code>logis</code>	location	scale		Item Response Theory
<i>Poisson</i>	<code>pois</code>	lambda			Count data
<i>Weibull</i>	<code>weibull</code>	shape	scale		Reaction time distributions
<i>Cauchy</i>	<code>cauchy</code>	location	scale	log	infinite variance!

## R can draw distributions

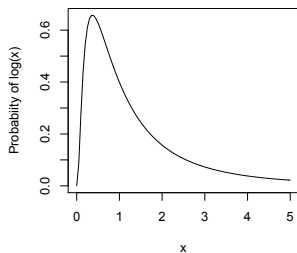
**A normal curve**



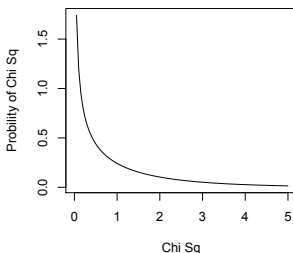
```
curve(dnorm(x),-3,3,ylab="probability of x",main="A normal curve")
```

## R can draw more interesting distributions

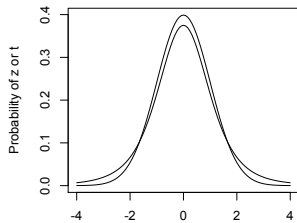
### Log normal



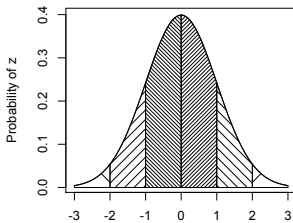
### Chi Square distribution



### Normal and t with 4 df



### The normal curve



## R is also a graphics calculator

The first line draws the normal curve, the second prints the title, the next lines draw the cross hatching.

```
op <- par(mfrow=c(2,2))          #set up a 2 x 2 graph
curve(dnorm(x),-3,3,xlab="",ylab="Probability of z")
title(main="The normal curve",outer=FALSE)
xvals <- seq(-3,-2,length=100)
dvals <- dnorm(xvals)
polygon(c(xvals,rev(xvals)),c(rep(0,100),rev(dvals)),density=2,angle=-45)
xvals <- seq(-2,-1,length=100)
dvals <- dnorm(xvals)
polygon(c(xvals,rev(xvals)),c(rep(0,100),rev(dvals)),density=14,angle=45)
xvals <- seq(-1,-0,length=100)
dvals <- dnorm(xvals)
polygon(c(xvals,rev(xvals)),c(rep(0,100),rev(dvals)),density=34,angle=-45)
xvals <- seq(2,3,length=100)
dvals <- dnorm(xvals)
polygon(c(xvals,rev(xvals)),c(rep(0,100),rev(dvals)),density=2,angle=45)
xvals <- seq(1,2,length=100)
dvals <- dnorm(xvals)
polygon(c(xvals,rev(xvals)),c(rep(0,100),rev(dvals)),density=14,angle=-45)
xvals <- seq(0,1,length=100)
dvals <- dnorm(xvals)
polygon(c(xvals,rev(xvals)),c(rep(0,100),rev(dvals)),density=34,angle=45)
curve(dlnorm(x),0,5,ylab='Probabiity of log(x)',main='Log normal')
curve(dchisq(x,1),0,5,ylab='Probility of Chi Sq',xlab='Chi Sq',main='Chi Square distribution')
curve(dnorm(x),-4,4,ylab='Probability of z or t',xlab='z or t',main='Normal and t with 4 df')
curve(dt(x,4),add=TRUE)
op <- par(mfrow=c(1,1))
```

## Example data sets built into many packages

Table: Some of the 48 data sets in the psych package

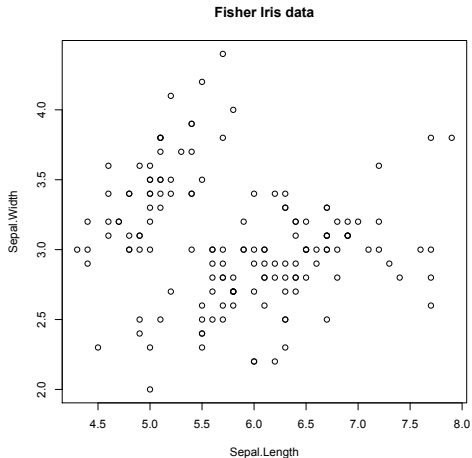
Name	Content description
veg	Thurstone's Vegetables
cities	Airplane distances for 11 US cities
galton	Francis Galton's original data set of heights
cushny	The original t-test data from "student" (Gossett)
ability	16 ability items from SAPA
bfi	25 Big Five items + gender, age, education from SAPA
sat.act	Test scores, gender, age and education
Thurstone	9 ability variables from Thurstone
msq	75 mood items from the PMC lab
neo	Correlation matrix of the 30 NEO-PI-R facets

```
data() #to see all available
```

```
data(package="psych") # to see all psych data sets
```

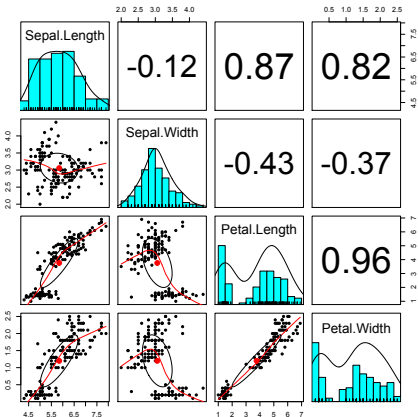


## A simple scatter plot using `plot` shows Fisher's Iris data set



```
plot(iris[1:2], xlab="Sepal.Length", ylab="Sepal.Width",  
     ,main="Fisher Iris data")
```

## A scatter plot matrix with loess regression using `pairs.panels` shows more information than a simple scatter plot

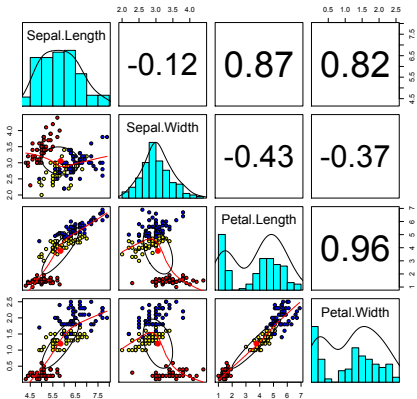


1. Correlations above the diagonal
2. Diagonal shows histograms and densities
3. scatter plots below the diagonal with correlation ellipse
4. locally smoothed (loess) regressions for each pair

```
pairs.panels(iris[1:4])
```

## A better SPLOM with colors for groups using pairs.panels

Fisher Iris data by Species



1. Correlations above the diagonal
2. Diagonal shows histograms and densities
3. scatter plots below the diagonal with correlation ellipse
4. locally smoothed (loess) regressions for each pair
5. optional color coding of grouping variables.

```

pairs.panels(iris[1:4],bg=c("red","yellow","blue")
[iris$Species],pch=21,main="Fisher Iris data by
Species")

```

## Using R for psychological statistics: Basic statistics

### 1. Writing syntax

- For a single line, just type it
- Mistakes can be redone by using the up arrow key
- For longer code, use a text editor (built into some GUIs)

### 2. Data entry

- Using built in data sets for examples
- Copying from another program (using the "clipboard")
- Reading a text or csv file
- Importing from SPSS or SAS
- Simulate it (using various simulation routines)

### 3. Descriptives

- Graphical displays
- Descriptive statistics
- Correlation

### 4. Inferential

- the t test
- the F test
- the linear model

## Data entry overview

1. Using built in data sets for examples
  - `data()` will list  $> 100$  data sets in the `datasets` package as well as all sets in loaded packages.
  - Most packages have associated data sets used as examples
  - *psych* has  $> 40$  example data sets
2. Copying from another program
  - use copy and paste into R using `read.clipboard` and its variations
3. Reading a text or csv file
  - read a local or remote file using the `read.file` function
4. Importing from SPSS or SAS
5. Simulate it (using various simulation routines)

## Examples of built in data sets from the psych package

```
> data(package="psych")
```

Bechtoldt

Seven data sets showing a bifactor solution.

Dwyer

8 cognitive variables used by Dwyer for an exam

Reise

Seven data sets showing a bifactor solution.

all.income (income)

US family income from US census 2008

bfi

25 Personality items representing 5 factors

blot

Bond's Logical Operations Test - BLOT

burt

11 emotional variables from Burt (1915)

cities

Distances between 11 US cities

epi.bfi

13 personality scales from the Eysenck Personality  
and Big 5 inventory

affect

Two data sets of affect and arousal scores as a  
personality and movie conditions.

galton

Galton's Mid parent child height data

income

US family income from US census 2008

iqitems

16 multiple choice IQ items

msq

75 mood items from the Motivational State Questi  
for N = 3896

neo

NEO correlation matrix from the NEO\_PI\_R manual

sat.act

3 Measures of ability: SATV, SATQ, ACT

Thurstone

The classic Thurstone 9 variable problem

veg (vegetables)

Paired comparison of preferences for 9 vegetable

## Reading data from another program –using the clipboard

1. Read the data in your favorite spreadsheet or text editor
2. Copy to the clipboard
3. Execute the appropriate `read.clipboard` function with or without various options specified

```
my.data <- read.clipboard()    #assumes headers and tab or space delimited
my.data <- read.clipboard.csv() #assumes headers and comma delimited
my.data <- read.clipboard.tab() #assumes headers and tab delimited
                                (e.g., from Excel)
my.data <- read.clipboard.lower() #read in a matrix given the lower
my.data <- read.clipboard.upper() #           or upper off diagonal
my.data <- read.clipboard.fwf() #read in data using a fixed format width
                                (see read.fwf for instructions)
```

4. `read.clipboard()` has default values for the most common cases and these do not need to be specified. Consult `?read.clipboard` for details.

## Reading from a local or remote file

- Perhaps the standard way of reading in data is using the read command.
  - First must specify the location of the file
  - Can either type this in directly or use the file.choose function
  - The file name/location can be a remote URL
- Two examples of reading data

```
file.name <- file.choose() #this opens a window to allow you find the file
my.data <- read.table(file.name) #
my.data <- read.file() #opens a search window, reads the file according to
datafilename="http://personality-project.org/r/datasets/R.appendix1.data"
data.ex1=read.table(datafilename,header=TRUE) #read the data into a table

> dim(data.ex1) #what are the dimensions of what we read?
[1] 18 2

> describe(data.ex1) #do the data look right?
      var  n mean  sd median trimmed  mad min max range skew kurtosi
Dosage*   1 18  1.89 0.76     2   1.88 1.48  1  3   2 0.16  -1.1
Alertness 2 18 27.67 6.82    27  27.50 8.15 17 41  24 0.25  -0.6
```



## read a “foreign” file e.g., an SPSS sav file

`read.spss` reads a file stored by the SPSS save or export commands.

```
read.spss(file, use.value.labels = TRUE, to.data.frame = FALSE,
          max.value.labels = Inf, trim.factor.names = FALSE,
          trim_values = TRUE, reencode = NA, use.missings = to.data.frame)
```

**file** Character string: the name of the file or URL to read.

**use.value.labels** Convert variables with value labels into R factors with those levels?

**to.data.frame** return a data frame? Defaults to FALSE, probably should be TRUE in most cases.

**max.value.labels** Only variables with value labels and at most this many unique values will be converted to factors if `use.value.labels = TRUE`.

**trim.factor.names** Logical: trim trailing spaces from factor levels?

**trim\_values** logical: should values and value labels have trailing spaces ignored when matching for `use.value.labels = TRUE`?

**use.missings** logical: should information on user-defined missing values be used to set the corresponding values to NA?

## Simulate data

For many demonstration purposes, it is convenient to generate simulated data with a certain defined structure. The *psych* package has a number of built in simulation functions. Here are a few of them.

### 1. Simulate various item structures

[`sim.congeneric`](#) A one factor congeneric measure model

[`sim.items`](#) A two factor structure with either simple structure or a circumplex structure.

[`sim.rasch`](#) Generate items for a one parameter IRT model.

[`sim.irt`](#) Generate items for a one-four parameter IRT Model

### 2. Simulate various factor structures

[`sim.simplex`](#) Default is a four factor structure with a three time point simplex structure.

[`sim.hierarchical`](#) Default is 9 variables with three correlated factors.

## Get the data and look at it

Read in some data, look at the first and last few cases, and then get basic descriptive statistics. For this example, we will use a built in data set.

```
> my.data <- epi.bfi
> headtail(my.data)
```

	epiE	epiS	epiImp	epilie	epiNeur	bfagree	bfcon	bfext	bfneur	bfopen	bdi	traitanx	stateanx
1	18	10	7	3	9	138	96	141	51	138	1	24	22
2	16	8	5	1	12	101	99	107	116	132	7	41	40
3	6	1	3	2	5	143	118	38	68	90	4	37	44
4	12	6	4	3	15	104	106	64	114	101	8	54	40
...	...	...	...	...	...	...	...	...	...	...	...	...	...
228	12	7	4	3	15	155	129	127	88	110	9	35	34
229	19	10	7	2	11	162	152	163	104	164	1	29	47
230	4	1	1	2	10	95	111	75	123	138	5	39	58
231	8	6	3	2	15	85	62	90	131	96	24	58	58

epi.bfi has 231 cases from two personality measures

## Now find the descriptive statistics for this data set

```
> describe(my.data)
```

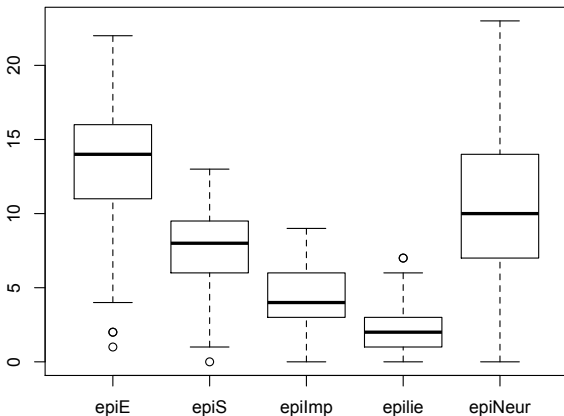
	var	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
epiE	1	231	13.33	4.14	14	13.49	4.45	1	22	21	-0.33	-0.01	0.27
epiS	2	231	7.58	2.69	8	7.77	2.97	0	13	13	-0.57	0.04	0.18
epiImp	3	231	4.37	1.88	4	4.36	1.48	0	9	9	0.06	-0.59	0.12
epilie	4	231	2.38	1.50	2	2.27	1.48	0	7	7	0.66	0.30	0.10
epiNeur	5	231	10.41	4.90	10	10.39	4.45	0	23	23	0.06	-0.46	0.32
bfagree	6	231	125.00	18.14	126	125.26	17.79	74	167	93	-0.21	-0.22	1.19
bfcon	7	231	113.25	21.88	114	113.42	22.24	53	178	125	-0.02	0.29	1.44
bfext	8	231	102.18	26.45	104	102.99	22.24	8	168	160	-0.41	0.58	1.74
bfneur	9	231	87.97	23.34	90	87.70	23.72	34	152	118	0.07	-0.51	1.54
bfopen	10	231	123.43	20.51	125	123.78	20.76	73	173	100	-0.16	-0.11	1.35
bdi	11	231	6.78	5.78	6	5.97	4.45	0	27	27	1.29	1.60	0.38
traitanx	12	231	39.01	9.52	38	38.36	8.90	22	71	49	0.67	0.54	0.63
stateanx	13	231	39.85	11.48	38	38.92	10.38	21	79	58	0.72	0.04	0.76

## Boxplots are a convenient descriptive device

Show the Tukey “boxplot” for the Eysenck Personality Inventory

```
boxplot(my.data[1:5]) #just the first 5 variables
```

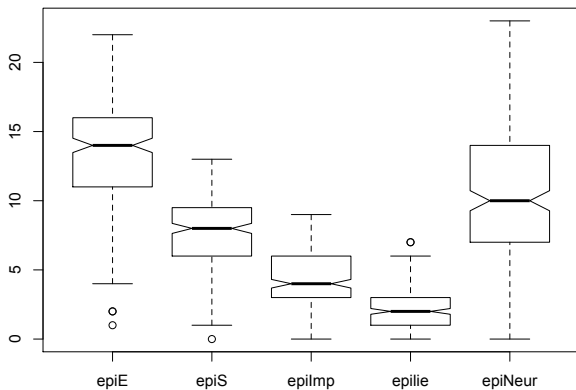
**Boxplots of EPI scales**



## Enhanced box plots are even more convenient descriptive devices

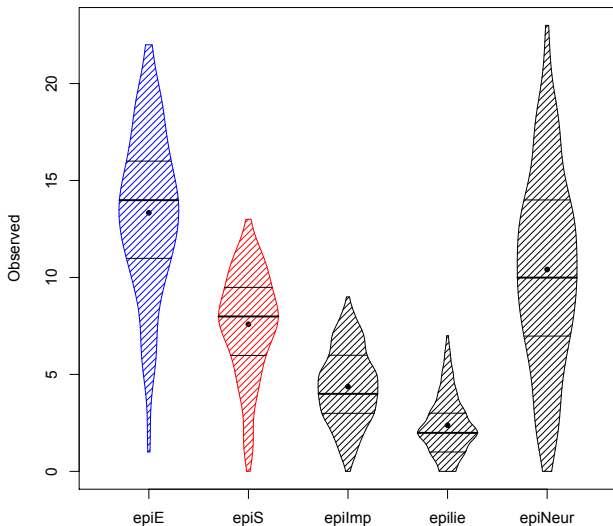
Show the Tukey “boxplot” for the Eysenck Personality Inventory  
`boxplot(my.data[1:5])` #just the first 5 variables

A notched boxplot of the epi

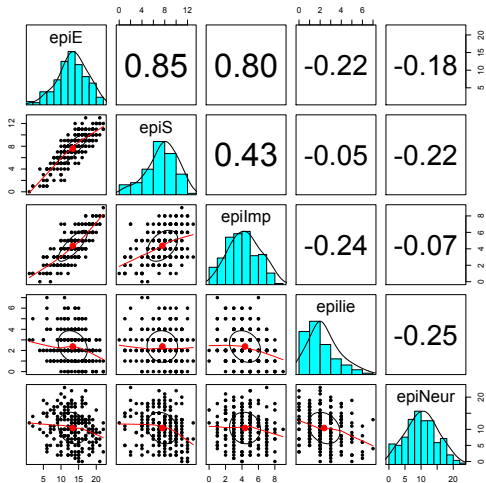


## A violin plot shows the density distribution more graphically

Density plot



## Plot the scatter plot matrix (SPLOM) of the first 5 variables using the `pairs.panels` function



Use the `pairs.panels` function from *psych*

```
pairs.panels(my.data[1:5])
```



## Find the correlations for this data set, round off to 2 decimal places

```
> round(cor(my.data, use = "pairwise"), 2)
```

	epiE	epiS	epiImp	epilie	epiNeur	bfragee	bfcon	bfext	bfneur	bfopen	
epiE	1.00	0.85	0.80	-0.22	-0.18	0.18	-0.11	0.54	-0.09	0.14	-0
epiS	0.85	1.00	0.43	-0.05	-0.22	0.20	0.05	0.58	-0.07	0.15	-0
epiImp	0.80	0.43	1.00	-0.24	-0.07	0.08	-0.24	0.35	-0.09	0.07	-0
epilie	-0.22	-0.05	-0.24	1.00	-0.25	0.17	0.23	-0.04	-0.22	-0.03	-0
epiNeur	-0.18	-0.22	-0.07	-0.25	1.00	-0.08	-0.13	-0.17	0.63	0.09	0
bfragee	0.18	0.20	0.08	0.17	-0.08	1.00	0.45	0.48	-0.04	0.39	-0
bfcon	-0.11	0.05	-0.24	0.23	-0.13	0.45	1.00	0.27	0.04	0.31	-0
bfext	0.54	0.58	0.35	-0.04	-0.17	0.48	0.27	1.00	0.04	0.46	-0
bfneur	-0.09	-0.07	-0.09	-0.22	0.63	-0.04	0.04	0.04	1.00	0.29	0
bfopen	0.14	0.15	0.07	-0.03	0.09	0.39	0.31	0.46	0.29	1.00	-0
bdi	-0.16	-0.13	-0.11	-0.20	0.58	-0.14	-0.18	-0.14	0.47	-0.08	1
traitanx	-0.23	-0.26	-0.12	-0.23	0.73	-0.31	-0.29	-0.39	0.59	-0.11	0
stateanx	-0.13	-0.12	-0.09	-0.15	0.49	-0.19	-0.14	-0.15	0.49	-0.04	0

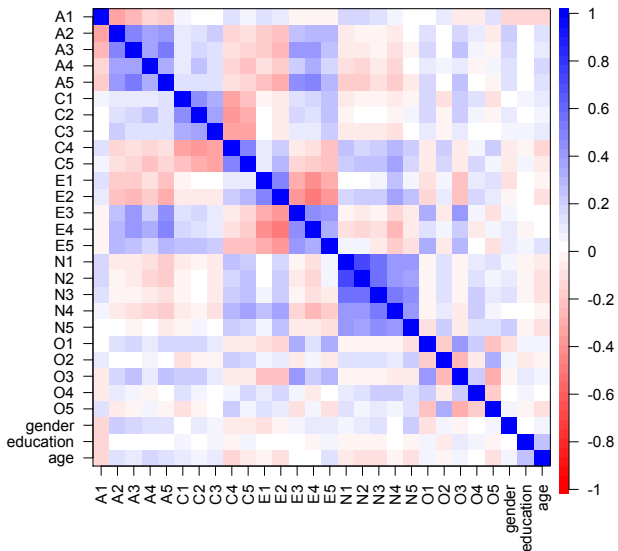
## Find the correlations using `lowerCor`

```
> lowerCor(my.data)
```

	epiE	epiS	epImp	epili	epiNr	bfagr	bfcon	bfext	bfner	bfopn	bdi	trtn
epiE	1.00											
epiS	0.85	1.00										
epiImp	0.80	0.43	1.00									
epilie	-0.22	-0.05	-0.24	1.00								
epiNeur	-0.18	-0.22	-0.07	-0.25	1.00							
bfagree	0.18	0.20	0.08	0.17	-0.08	1.00						
bfcon	-0.11	0.05	-0.24	0.23	-0.13	0.45	1.00					
bfext	0.54	0.58	0.35	-0.04	-0.17	0.48	0.27	1.00				
bfneur	-0.09	-0.07	-0.09	-0.22	0.63	-0.04	0.04	0.04	1.00			
bfopen	0.14	0.15	0.07	-0.03	0.09	0.39	0.31	0.46	0.29	1.00		
bdi	-0.16	-0.13	-0.11	-0.20	0.58	-0.14	-0.18	-0.14	0.47	-0.08	1.00	
traitanx	-0.23	-0.26	-0.12	-0.23	0.73	-0.31	-0.29	-0.39	0.59	-0.11	0.65	1.00
stateanx	-0.13	-0.12	-0.09	-0.15	0.49	-0.19	-0.14	-0.15	0.49	-0.04	0.61	0.5

## A heat map of 25 BFI items using `cor.plot`

Big 5 Inventory Items from SAPA



- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671–684.
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116–127.
- Eysenck, H. (1983). Is there a paradigm in personality research? *Journal of Research in Personality*, 17(4), 369 – 397.
- Eysenck, H. J. (1966). Personality and experimental psychology. *Bulletin of the British Psychological Society*, 19, 1–28.
- Eysenck, H. J. (1976). *The Measurement of personality*. Lancaster: MTP.
- Eysenck, H. J. (1985). The place of theory in a world of facts. In *Annals of theoretical psychology*, volume 3 (pp. 17–72). Plenum Press.
- Eysenck, H. J. (1987a). The growth of unified scientific psychology: Ordeal by quackery. In A. W. Staats & L. P. Mos (Eds.), *Annals of Theoretical Psychology*, volume 5 chapter 3, (pp. 91–113). Plenum Press.

- Eysenck, H. J. (1987b). "there is nothing more practical than a good theory" (kurt lewin) –true or false? In W. J. Baker, M. E. Hyland, H. V. Rappard, & A. W. Staats (Eds.), *Current Issues in Theoretical Psychology*, volume 40 (pp. 49–63). North Holland.
- Eysenck, H. J. (1988). The growth of a unified scientific psychology. In A. Staats & L. Mos (Eds.), *Annals of Theoretical Psychology*, volume 5 of *Annals of Theoretical Psychology* (pp. 91–113). Springer US.
- Eysenck, H. J. (1997). Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology*, 73(6), 1224–1237.
- Eysenck, H. J. & Eysenck, M. W. (1985). *Personality and Individual Differences: A natural science approach*. New York: Plenum.
- Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.

- Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246–263.
- Gray, J. A. (1981). A critique of Eysenck's theory of personality. In H. J. Eysenck (Ed.), *A Model for Personality* (pp. 246–277). Berlin: Springer.
- Kimble, G. A. (1984). Psychology's two cultures. *American Psychologist*, 39(8), 833 – 839.
- Lakatos, I. (1968). Criticism and the methodology of scientific research programmes. *Proceedings of the Aristotelian Society*, 69, pp. 149–186.
- Loehlin, J. C. (2004). *Latent variable models: an introduction to factor, path, and structural equation analysis* (4th ed.). Mahwah, N.J.: L. Erlbaum Associates.
- Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York,: McGraw-Hill.

Pearson, K. P. (1896). Mathematical contributions to the theory of evolution. iii. regression, heredity, and panmixia. *Philisopical Transactions of the Royal Society of London. Series A*, 187, 254–318.

Revelle, W. (in prep). *An introduction to psychometric theory with applications in R*. Springer.

Revelle, W. & Oehlberg, K. (2008). Integrating experimental and observational personality research – the contributions of Hans Eysenck. *Journal of Personality*, 76(6), 1387–1414.

Snow, C. P. (1959). "the rede lecture, 1959". In *The Two Cultures: and a Second Look* (pp. 1–21). Cambridge University Press.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101.

Student (1908). The probable error of a mean. *Biometrika*, 6(1), 1–25.

- Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist*, 30, 128–134.
- Vale, J. & Vale, C. (1969). Individual differences and general laws in psychology: a reconciliation. *American Psychologist*, 24(12), 1093–1108.
- Wundt, W. (1904). *Principles of Physiological Psychology* (Translated from the Fifth German Edition (1902) ed.). London: Swan Sonnenschein.