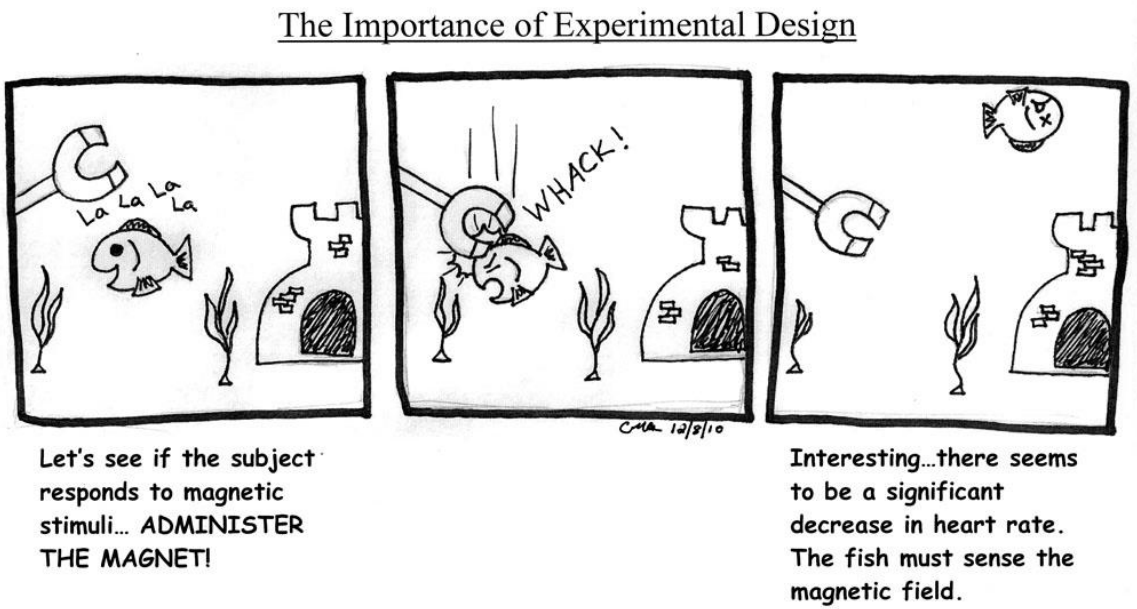
# C82MST: Practical and statistical methods

# Experimental and quasi-experimental designs:

## Establishing causal effects of independent variables

**Tobias Bast**

**School of Psychology, University of Nottingham**




From: http://www.hawaii.edu/fishlab/NearsideFrame.htm

**Feynman – Key to science**

http://youtu.be/b240PGCMwV0

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Richard Feynman
1918-1988

# Feynman – The key to science

In general, we look for a new law by the following process: First we guess it; then we compute the consequences of the guess to see what would be implied if this law that we guessed is right; then we compare the result of the computation to nature, with experiment or experience, compare it directly with observation, to see if it works. **If it disagrees with experiment, it is wrong.** In that simple statement is the key to science. It does not make any difference how beautiful your guess is, it does not make any difference how smart you are, who made the guess, or what his name is — **if it disagrees with experiment, it is wrong**.

http://youtu.be/b240PGCMwV0

# Outline

- **Experimental method in psychological research**

- **‘True’ experiments and importance of random allocation**

- **‘Quasi’-experiments**

- **Alternatives to random allocation: systematic ways to control for nuisance variables**

# Experiments

**Experiments –** involve the manipulation of a variable of interest (***independent variable*** or ***treatment***) and the measurement of the effect this has on another variable of interest (***dependent variable***) with the aim of establishing ***causality***. (Other research approaches include (cor)relational and observational/descriptive approaches.)

**Experimental design –** protocol for data collection with the aim to establish ***causality*** between treatment and changes in the dependent variable; a key consideration is how participants are allocated to treatment conditions.

**'True' experiments –** involve full control of the experimenter over allocation and scheduling of the treatment; in psychology and social sciences, 'true' experiment is often meant to refer to protocols involving ***random allocation*** of participants to treatments.

## What is the purpose of random allocation of participants to treatment conditions?

a) Practicality

b) Increase of statistical power

c) Control for nuisance/confounding variables

d) Increase in external validity

# Random allocation

•Every participant has an equal chance of being allocated to any condition.

•Purpose is to spread any potentially confounding differences ('nuisance variables') between participants evenly across treatment conditions; i.e. to ***minimise systematic differences other than the treatment***.

•Random allocation ***aids in isolating the causal effects of the treatment on the dependent variable*** (by removing systematic influence of nuisance variables/threats to internal validity).

•Particularly important where potential nuisance variables are difficult to identify.

•Procedures for random allocation may involve using a draw to allocate participants to conditions or using a random number generator.

# Exercise

**What's wrong with the following experiment?**

A final year project student wants to examine how caffeine affects performance on a memory test. All participants were asked not to consume caffeinated drinks within 24 h before the test. The first 15 participants to come to the department were given a cup of caffeinated coffee before the memory test. The next 15 participants to come to the department were given a cup of decaffeinated coffee before the memory test. The test performance of each group was used as dependent variable.

# What's wrong with this experiment?

a) Nothing.

b) Allocation of participants to groups.

c) Sample size.

d) Both b) and c).

# ‘Quasi’-experiments

•Research procedure that aims to establish causal effect between an independent variable and variations in a dependent variable, but where there is ***no full control over the allocation of participants*** to the different levels of the independent variable.

•The ***lack of random allocation*** is typically considered the demarcation from ‘true’ experiments in many areas of psychology, such as social and educational psychology.

•The lack of random allocation poses a key threat to internal validity, as it increases the ***risk that group/conditions may systematically differ with respect to factors other than the independent variable***.

# Some common 'quasi'-experimental designs

Nomenclature: X = a treatment O = observation/measurement
… = not randomly assigned

- **One group pre-post test design**

O X O

**Example**
To examine the effects of mindfulness training on mental health, researchers ask students to complete a mental health questionnaire at one time point before an extended course of mindfulness training and at one time point after this extended course of mindfulness training. The researchers found that the students' mental health was significantly improved at the second time point.

**What are the problems with this pre-post test (O X O) design – why may we see changes in the dependent variable, even though the treatment (X) had no causal effect?**

**What are the problems with a pre-post test (O X O) design?**

a) Experimenter bias.

b) Events unrelated to the treatment, happening between the two measurements, may affect the dependent variable (aka 'history').

c) Changes related to the participants between the two measures may affect the dependent variable (aka 'maturation').

d) Both b and c

**These problems compromise:**

a) the internal validity of the study.

b) the external validity of the study.

c) both a) and b).

d) None of the above.

# Some common ‘quasi’-experimental designs

Nomenclature: X = a treatment O = observation/measurement
… = not randomly assigned

- **One group pre-post test design**

O X O

- **Non-equivalent control group design**

O X O
. . . . . .
O O

- **Interrupted time-series design**

O O O O O O O X O O O O O O O O
. . . . . .
O O O O O O O O O O O O O O O O (with non-equivalent control group)

# Systematic approaches to control for nuisance variables

- **Blocking**
- **Matching**
- **Counterbalancing**

# Blocking

Experiment is arranged in blocks, within which a known nuisance variable is held constant (e.g., experimenter; sex of participant; time of testing) while the independent factor is allowed to vary; within each block, each level of the independent variable is tested equally often and participants can be randomly allocated to the levels of the independent variable (***randomised block design***). Blocking separates the effect of a known nuisance variable from the effect of the variable of interest.

***Example***: Blocking by experimenter in a lab class experiment

| ***Block*** | ***Subject*** | ***Treatment*** |
|---|---|---|
| Experimenter 1 | 1 | A |
| | 2 | A |
| | 3 | B |
| | 4 | B |
| Experimenter 2 | 5 | B |
| | 6 | A |
| | 7 | B |
| | 8 | A |
| Etc. | | |

# Matching

Groups are 'matched'/comparable with respect to specific individual difference (e.g., education; sex; age; pre-treatment performance; smoking status) to minimise the influence of this nuisance variable on the dependent variable; in a ***matched-subjects/matched case-control design***, every participant is directly compared with another participant who is matched according to a relevant nuisance variable. Compare: http://www.bmj.com/content/309/6962/1128

***Example***: Differences in hippocampal size between depressed and control subjects (YI Sheline et al., 1994, Proc Nat Acad Sci USA 93:3908-3913)

Table 1. Comparison of depressed subjects and matched controls

| | Depressed subjects ($n$ = 10) | | Matched controls ($n$ = 10) | | Paired $t$ test | |
|---|---|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD | $t$ | $p$ |
| Age, yr | 68.5 | 10.4 | 68.0 | 9.5 | 0.7 | 0.52 |
| Education, yr | 14.3 | 2.9 | 13.6 | 2.9 | 1.1 | 0.31 |
| Height, cm | 161.2 | 5.6 | 162.1 | 6.1 | 0.5 | 0.73 |
| Cortisol, μg/liter | | | | | | |
| Baseline | 20.8 | 10.2 | 23.8 | 9.5 | 0.6 | 0.55 |
| DST | 2.7 | 1.2 | 2.0 | 0.9 | 0.8 | 0.44 |
| HRSD | 6.0 | 5.4 | 1.9 | 1.7 | 2.8 | 0.02 |
| Days depressed | 1293 | 1067 | — | — | — | — |
| Race, no. | | | | | | |
| White | 9 | (90%) | 10 | (100%) | — | — |
| Black | 1 | (10%) | 0 | (0%) | — | — |
| Hippocampal gray matter volume, $mm^3$ | | | | | | |
| Left | 2159 | 301 | 2544 | 333 | 4.1 | 0.003 |
| Right | 2283 | 324 | 2577 | 259 | 2.8 | 0.02 |
| Total cerebral volume, $mm^3 \times 10^3$ | | | | | | |
| | 1167 | 133 | 1159 | 104 | 0.2 | 0.83 |
| Number of LSF in hippocampus | | | | | | |
| Left large* LSF | 22 | 14 | 6 | 7 | 3.3 | 0.01 |
| Right large* LSF | 23 | 16 | 9 | 9 | 2.4 | 0.04 |

LSF, low signal foci; DST, dexamethasone suppression test (see text).
*Large defined as ≥4.5 mm diameter.

# Counterbalancing

In within-subjects designs, the confounding effects of nuisance variables, such as testing order, can be ***countered*** by ***balancing*** them across the different levels of the independent variable.

***Examples***:

1. Counterbalancing of testing order across two conditions A and B: One half of the subjects is first tested in condition A, then in B; the other half is first tested in B, then in A. So, any difference in the dependent variable between conditions A and B cannot be accounted for by testing order.

2. ***Latin Square Designs***: used when there are more than 2 conditions.
   - N conditions, A, B, C, . . ., N, can be arranged such that each condition occupies each rank in the testing order equally often.
   - To do this, sequences of A to N are arranged in a Latin Square consisting of N columns and N rows, with each row and each column containing each condition exactly once, e.g. for three conditions A to C:

     A B C
     B C A
     C A B

   - Then, the same number of participants is allocated to each of the different testing orders specified by the different rows (i.e., the number of participants must be a multiple of the different testing orders, which equals the number of conditions).

# Exercise

a) Work out the testing sequences for an experiment with five conditions A, B, C, D and E according to a Latin Square design to avoid confounding effects of testing order.

b) Which numbers of participants may, in principle, be suitable for such a Latin Square design: 5, 10, 12, 15, 35, 36, 40, 99, 10 000

# In a nutshell

• Aim of the experimental approach is to establish **causal relationships** between an independent variable (‘treatment’, ‘manipulation’) and a dependent variable.

• A key consideration in experimental design are **nuisance variables or third factors** that may lead us to falsely conclude a causal relationship between independent and dependent variables.

• One key approach to control for nuisance variables is **random allocation** to treatments; alternatives include **blocking, counterbalancing and matching**.

• In psychology and social sciences, random allocation is often considered the demarcation criterion between ‘true’ experiments and ‘quasi’-experiments.

## Suggested reading

A book on research methods, for example:

A Field & G Hole (2003) How to design and report experiments. Sage, London. Especially chapter 3.

DG Elmes, BH Kantowitz, HL Roediger III (any recent edition) Research methods in psychology. West Publishing Company, St. Paul.

## Further specialised reading

E Ferguson & P Bibby (2004) The design and analysis of quasi-experimental field research. In: GM Breakwell (ed.), Doing social psychology research, Chapter 3, p. 93-127.

MFW Festing, P Overend, RG Das, MC Borja, M Berdoy (2002) The design of animal experiments – reducing the use of animals in research through better experimental design. Royal Society of Medicine Press, London.

DT Campbell, JC Stanley (1966) Experimental and quasi-experimental designs for research. Houghton Miflin Company, Boston.

# Some questions for revisions

•What are the key features differentiating the experimental approach from correlational and observational research approaches in psychology?

• In general:
How can we ensure the internal validity of our research (i.e., that any changes observed in our dependent variables are due to changes in our independent variable, rather than other factors)?

• More specifically:
-What is the purpose of random allocation?
- Are there alternatives to random allocation? Consider pros and cons.