

Introduction to research methods in Psychology

Research methods

Dr Douglas Elliffe

HSB 326

Sometimes I'm there, sometimes in my lab (HSB 319)

d.elliffe@auckland.ac.nz

e-mail is definitely the best way to get hold of me

Research methods

Three lectures, 4/5/10 March

Recommended reading

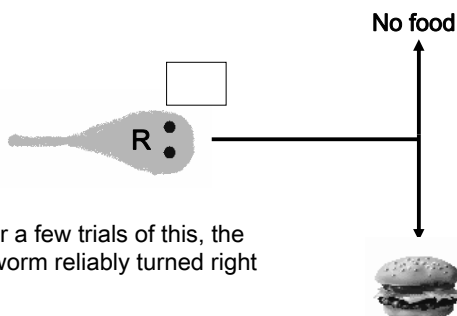
- Textbook Chapter 2
- Laboratory One in your lab manual (week starting 31 March)
- One page summary handout on Cecil
- These lecture .ppt slides are also on Cecil – two files

I'd concentrate on these last two more than the textbook

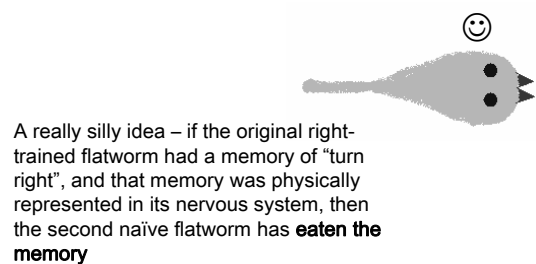
A research cautionary tale

- Not examinable, just a story (though it's a true one)
- About 40 years ago, lots of interest in the physical representation of memory in the brain
- The search for the **engram**
- Researchers did an experiment with flatworms, or planaria, to try to prove that memory was stored physically in the brain
- Would memory transfer from one flatworm to another?

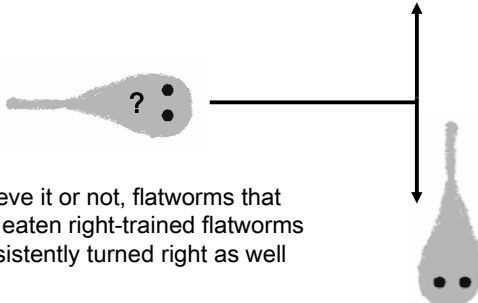
Training a flatworm to turn right ...



Feed the right-trained flatworm to a new flatworm



So test the second flatworm in the maze ...



Believe it or not, flatworms that had eaten right-trained flatworms consistently turned right as well

But ...

- Since the original flatworm experiment, there have been many attempts to **replicate** it (do it again and see if we get the same result)
- **None of them have been successful**
- What was wrong with the original study? Was it:
 - Well conducted but a fluke result?
 - The role of **statistics** is to allow for this possibility – that's why you have to take a Statistics course to major in Psychology (see also my third lecture)
 - Flawed because it was badly conducted?
 - The role of **research design** is to prevent this, and that's what these lectures are about

Basics

- Research is process of asking questions about the world
- We all do this – we're all researchers really
- Good research design is a formalised way of asking questions so that we're sure the answers will mean something
- Most of you won't go on to be **producers** of research (although I hope you will), but you are all already **consumers** of it
- I'd like to make you more sophisticated, cautious, sceptical, suspicious consumers

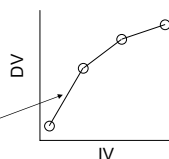
Basics

- Same principles apply in all research (physics, biology, history)
- But there's a particular ethical obligation to make sure that psychological research is done well
 - Our subjects are usually people, and we have a duty not to waste their time
 - Results of psychological research are often applied to how people are treated, so we'd better get it right
- Most of a Stage 1 course in anything is about learning the words used in the field so you can learn the content more efficiently, so ...

Terminology – DV and IV

- **Dependent variable (DV)**
 - Measure of the behaviour we're interested in
 - Goes on *y* axis of a graph of our results
 - e.g., reaction time in mental rotation experiment
- **Independent variable (IV)**
 - Variable manipulated by an experimenter to see if it affects the DV
 - Goes on *x* axis of a graph
 - e.g., rotation from vertical of target in same experiment

One possible relationship between IV and DV



Terminology – operational definitions

- Both our IV and our DV must be operationally defined
- **Operational definition**
 - A statement of the operations carried out by the researcher to measure the DV or manipulate the IV
 - The Method section of a research report is about operational definitions
 - Helps other researchers **replicate** (repeat) our study
 - Helps us be objective in what we conclude about our research – avoid bias

Terminology – validity and reliability

- Dependent variables need to be **valid** and **reliable**
- **Validity**
 - Does the DV, as operationally defined, measure what we want it to?
 - e.g., is IQ score a valid measure of intelligence?
 - (Gleitman *et al*/call this **internal** validity)
- **Reliability**
 - Will you get the same results if you measure the same variable again? – replication
 - Good operational definitions help with this
- Measures are sometimes reliable but not valid
 - e.g., shoe size as a measure of IQ
- They probably can't be valid if they're not reliable

Terminology – sample and population

- **Population**
 - All the scores, events, etc that we're interested in
 - e.g., heights of the entire PSYCH 109 class
- **Sample**
 - A representative subgroup of the population
 - e.g., heights of a random selected 10 people in the class
- We usually can't measure the entire population we're interested in, so we take a sample and infer general results about the population from the sample

Terminology – sampling error and sampling bias

- Two ways that sampling can go wrong :
- **Sampling error**
 - Results of repeated samples from the same population will always differ
 - Unavoidable
 - But you can minimise it by using bigger samples
- **Sampling bias**
 - Sample misrepresents population in a systematic way
 - Avoid by doing **random** sampling
 - Serious sampling bias **invalidates** the research
 - Gleitman *et al* – this specifically damages **external** validity

Examples of sampling error and sampling bias

- Population = heights of the 109 class
- DV = average height
- How sampling error depends on sample size
- Suppose I did a sample where I chose everyone whose first name ends in a vowel to be in the sample – how does this lead to sampling bias?
- Historical example of sampling bias :
 - 1948 US election – phone polls predicted win to Thomas Dewey (Republican), but Harry Truman (Democrat) won easily. Why?

Two general kinds of research design ...

- **Observational** design
 - Measure two DVs and look for a relationship between them
 - Sometimes called a **correlational** design, because a relationship between variables is a **correlation**
 - There is no IV, because nothing is manipulated
 - Gleitman *et al* say that there is an IV, but they're using IV to mean 'cause of the change in DV' – believe me not them
 - e.g., is self-esteem related to height?
 - Often only choice for ethical or practical reasons
 - e.g., discovery of Broca's area for speech in the brain

The basic weakness of observational designs

- Just because two DVs are correlated, we can't conclude that one affects the other
- e.g., suppose we find that watching violent films and aggressive behaviour are correlated
- Maybe watching violent films makes you aggressive
 - A causes B
- But maybe naturally aggressive people like watching violent films
 - B causes A
- Or maybe they're both caused by another variable
 - C causes both A and B
 - Called the third variable problem

The basic weakness of observational designs

- If there's one thing I'd most like you to remember from these lectures, it's :

Correlation does not imply causation

Two general kinds of research design ...

- **Experimental** design
 - Manipulate IV and look to see effect on DV
 - If it's replicable, **can** imply causation
 - e.g., is self-esteem related to results of a fake IQ test?
 - e.g., Flourens and experimental ablation
- Experimental designs are more powerful than observational, so we should use them when it's ethical and practical to do so
- When you're evaluating research that you read about, always consider this issue – has causation really been proved?

Two general kinds of research design ...

- This explains, I hope, why I'm a bit fussy about what an independent variable really is (see footnote in Laboratory One in your lab manual)
- A lot of people, including the authors of your lab manual, use 'IV' more loosely than I think we should, to mean 'the variable that caused the changes in the DV'
- But if you haven't done an experiment, then how do you know what caused the changes?
- So Lab One isn't really an experiment, because it compares the behaviour of males *vs* females, a pre-existing variable not manipulated by us
- And sex isn't an IV
- And it's not possible to be confident about causation ...

The fundamental principle of research design

- We want to eliminate all explanations of our results except one
- That is, if we conduct an experiment, and see a change in behaviour (our DV), we want to be sure that it was caused by our IV
- Alternative explanations of the results are called **confounds** or **confounding variables**
- A confounding variable is anything, other than our IV, that might have produced the change in DV that we saw
- They are bad, and we need to eliminate them if want our research to be worthwhile

General approaches to eliminating confounds

- Hold the confounding variable constant
 - especially good for environmental or external confounds
 - called "standardization" in your textbook
- Randomize the confound
 - especially good for subject-based or internal confounds
- (Decide that they're more interesting and study them instead)

General approaches to eliminating confounds

- You should always think about research that you design and that you read in these terms :
 - What possible confounds were there ?
 - Have they been controlled ?
- If not, don't trust the research

Example – within-subjects ν between-groups

- **Within-subjects design**
 - Each subject is exposed to all levels of the IV (all experimental conditions)
 - e.g., does alcohol affect short-term memory?
 - Each subject tries to learn a list of words in both “sober” and “drunk” conditions
 - There is an effect if individual subjects behave differently in the two conditions
- Good for controlling subject confounds
- But there could be environmental confounds – such as ?
- Especially note the risk of an **order effect**

Example – within-subjects ν between-groups

- **Between-groups design**
 - Each subject is only exposed to one level of the IV (one experimental conditions)
 - e.g., separate groups of subject in “sober” and “drunk” conditions
 - There is a difference if the average behaviour of the two groups is different
- Good for controlling environmental confounds
- But could be subject confounds – maybe the people in the “sober” condition just had better memories anyway
- Can partly prevent this by **randomly** allocating subjects to groups

Other names for these designs

- Between-groups design also called
 - Between subjects
 - Independent samples
 - Independent groups
- Within-subjects design also called
 - Repeated measures

A way to get the best of both designs

- **Matched-pairs design**
 - Give a pretest on memory ability before the experiment
 - Find pairs of subjects with similar memory ability
 - Randomly allocate one member of each pair to each condition, “sober” and “drunk”
 - There is an effect if the “sober” member of each pair consistently scores better than their “drunk” partner
- Keeps environmental confounds constant, and nearly keeps subject confounds constant too
- This is a very good design, but quite labour-intensive, so it isn't used as often as it should be

Experimental and control groups

- Easily the most common design in Psychology is a between-groups design that compares an experimental group with a control group
- e.g., assessing the effectiveness of a new drug
- The two groups are identically treated in all respects except that :
 - The experimental group receives the treatment
 - The control group doesn't
- One possible confounding variable is **subject expectation**, so it's important that the subjects don't know which group they're in
- This is a **single-blind** design

Experimental and control groups

- Another important confounding variable is **experimenter expectation** – why?
 - The experimenter needs to remain objective
 - The experimenter's expectation might affect how the subjects behave
- So it's best if the experimenter doesn't know who is in which group either
- This is a **double-blind** design
- Usually thought of as the “gold standard” for research