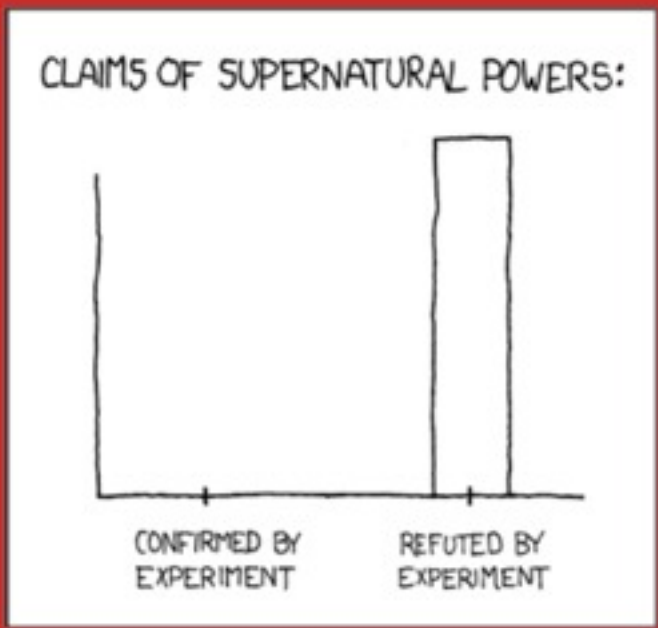
Tenei pukapuka no ki te ____
Please return if found.

CONDUCTING EXPERIMENTS

Statistical Experiments AS91583



Written and complied by Mr A Rambhai who is currently
the Teacher in Charge of Level 3 Statistics at Ormiston Senior College.
Inspiration, images and ideas provided by Ms A Martin of Avondale College.

For more information, please contact:

Mr A Rambhai
rambhai.inc@gmail.com

Before you begin....

I created this book to compliment the teaching and learning for *Achievement Standard 91583: Conduct an experiment to investigate a situation using experimental design principals.*

This book should not excuse any student from attending classes; there will be crucial things taught by the teacher and discovered together as a class. This also includes suggested answers to parts of this book!

You should work through this book from start to finish, reading all the explanations and attempting all the questions. Show your teacher what you have done regularly!

Work hard, and enjoy this year.

Mr A Rambhai

ex•per•i•ment

noun |ikˈsperəmənt|

a scientific procedure undertaken to make a discovery, test a hypothesis, or demonstrate a known fact : *laboratory* ***experiments on*** *guinea pigs | I have tested this by experiment.*

• a course of action tentatively adopted without being sure of the eventual outcome : *the previous* ***experiment in*** *liberal democracy had ended in disaster.*

Glossary of vocabulary for this topic:

Version 2.0

Bias	Something that causes favourititism.
Cause	This is usually the *treatment*.
Context	The real world story or facts behind an experiment.
Control group	The group who does not receive the treatment.
Effect	The outcome of applying a treatment, measured by the *response variable*.
Experiment	Process of planning, running, and looking at the results of a test.
Experimental Group	Group of experimental units.
Experimental Unit	Single person who is being tested upon in an experiment.
Experimenter	Person or group of people in charge of running an experiment.
Hypothesis	Predication, or expectation. Usually made before an experiment.
Independent variable	Usually takes only two values, *placebo* and *treatment.*
Median	The central or middle value of an ordered dataset
Paired Comparison	An experiment on a single experimental group, taking a before and after measurement.
Placebo	Simply put, a fake *treatment.*
Purpose	The throughly-developed-line-of-reasons for running an experiment.
Random Allocation	Process of randomly assigning *experimental units* to groups using, for example a deck of cards.
Randomisation test	Process of testing if chance alone is influencing the results from an experiment.
Response variable	The measurement that is the main focus of an experiment.
Spread	The spread of the data around the median, measured by the interquartile range (IQR) or standard deviation.
Treatment	An applied change or influence that should result in a change in the *response variable.*
Treatment group	The group who receives the treatment

Chapter One

INFORMED CONTEXTUAL KNOWLEDGE

... setting up a purposeful experiment.


Maybe time to get a new haircut?
Does hairstyle affect age estimates?


Swearing helps you withstand pain better!

Context is all about making an experiment real. Without **context** or real world application you are left with nothing more than a theoretical argument.

The examples you saw on the previous page (hopefully) introduce some important ideas or features of experiments you will need to think about.

Remember this lady? Take a moment and study the picture carefully... what is actually going on?

The picture is from an **experiment** which is testing the **hypothesis:**

Swearing helps you withstand pain better!

The **experimental unit** is the female. She has her hand in a bowl of ice water. In the bowl is a probe, which monitors the temperature of the the ice water... this is important because each **experimental unit** should experience the same coldness of the ice water. There is also a timer in the background.

The experiment was run like this: 2 groups of people are selected randomly from a larger group, one is the **control** group who are not allowed to swear, and the other is the **treatment** group who are allowed to swear; in fact, they are encouraged to swear! Being able to swear is the the **treatment**.

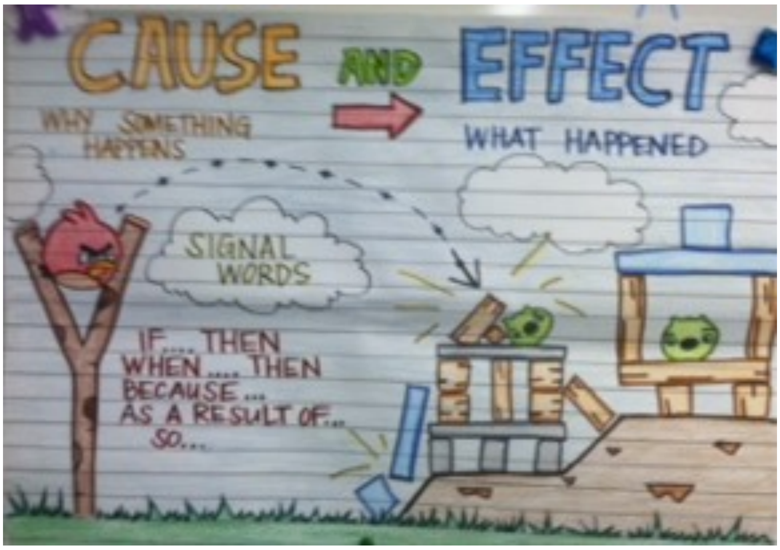
The **control** group each has a turn at submerging their right hand in the ice - water. Once a person (**experimental unit**) puts his / her hand into the bowl, the timer starts. When the person cannot bear the pain any more, he / she will remove their hand from the bowl and the timer will stop. The **experimenter** will then record the time. Each person in the group will have a go, and have a time recorded. Remember, the **control** group will be told from the start of the experiment they are not allowed to swear.

The **treatment** group is next, and the experiment will be repeated and their times recorded. Remember they will be encouraged to swear.

(continued on next page...)

The 2 groups are the **independent variables**, and the time measured is the **response variable**. Analysing the “time” variable will help the **experimenter** determine if allowing the participant to swear translates to being able to withstand pain for longer (or is it just extreme cold?).

A summary of the results of the experiment will help justify the claim that “Swearing helps withstand pain better”. This is an example of **cause** and **effect**.



Purpose_I : the context...

In a statistical experiment, there must be a well developed set of reasons that justify the running of an experiment (and therefore also the use of the results from the experiment!).

You will be provided with a context or scenario.
You will then need to find out more about the context or scenario (research!!).
You will use what you have found out to pose an investigation question (see Purpose_II).

Of course, one of the best places to get information from is via www.google.co.nz

I will illustrate the point using an example:

Context or scenario: People who own smart phones are losing brain cells.

#One - Start by using looking up smart phones, and it might be a good idea to define or explain what a smart phone actually is..

A cellular telephone with built-in applications and Internet access. In addition to digital voice service, modern smart phones provide text messaging, e-mail, Web browsing, still and video cameras, MP3 player and video playback and calling

Source: www.pcmag.com

#Two - Find out what causes people to lose brain cells..

4mind4life.com (a mental health prevention website) lists 50 causes of brain cell loss such as concussions, illicit drug taking and even lack of quality sleep. Out of the 50 causes listed, there is no section on cellular phones (or radio waves).

There is some evidence that exposure to electromagnetic radiation from cellular phones, causes brain tissue to heat up. This may have implications for brain cell loss but conclusive evidence was not stated.

Source: news.menshealth.com

#Three - Find evidence of experiments that have already been conducted..

The University of Pittsburgh Cancer Institute ran an *observational* study of newly diagnosed glioma (general term of brain cancer) cases over 4 years, and compared it with their hours of cell phone use. The study found that there was no association (so no relationship) between hours of cell phone use and glioma cases. The study is not an experiment, but the results do contradict the claim that cellphones cause brain cell loss.

Purpose_II : the variables...

Before you write an investigation question, you should have a clear idea of the type or kind of experiment you will be conducting. I would advise a discussion with your teacher once you have your idea.

A good **investigation question** will have:

1) A well-defined **response variable** (something that is measurable or quantifiable).
2) A well-defined **explanatory variable** (2 values only, *control* and *treatment*).
3) A well-defined **experimental group**.

... you should try to write a reason or reasons why you are choosing each of these!! Ideally, you should have ideas from your research to back up your choice.

I will illustrate using the example from before:

Context or scenario: People who own smart phones are losing brain cells.

Memory retention is dependent on the number of connections between cells. This implies that if a person has a good memory retention they are likely to have a high number of connections between brain cells; which implies the person will have a higher number of brain cells. I can't practically count the number of brain cells a person has, so I will be using the 'ability of remember' as a proxy for the number of brain cells.

The more numerous and healthy the brain-cell connections and the faster the signals can go back and forth, the better the mind and the memory will work.

Source: heatlh.howstuffworks.com

In order to test this claim, I will be running a "memory test" experiment. In this experiment I will have a group of people in a room and they will be shown an image on the screen for a set amount of time. After the image disappears I will ask each individual to write down all the items from the images they could remember.

The **response variable** will be the number of correct items each individual can remember.

The **explanatory variable** will be the individuals having their cellular / smart phone on in their pocket for 10mins prior to the memory test. I cannot study prolonged exposure to electromagnetic radiation (like in the University of Pittsburgh study), so this will have to be sufficient.

The **experimental group** will be a Year 11 Mathematics class at Ormiston Senior College. Any participants selected must be from Ormiston Senior College, and for the sake of simplicity and ease of data collection a class is the best option.

Purpose_III : the question...

A good **investigation question** will have:

1) A well-defined **response variable** (something that is measurable or quantifiable).
2) A well-defined **explanatory variable**.
3) A well-defined **experimental group**.

I am going to investigate the relationship between the memory score of Year 11 Math's students in Mr R's class who have their cell phones OFF and Year 11 Math's students in Mr R's class who have their cell phones ON.

I am going to investigate the difference between the median memory score of Year 11 Math's students in Mr R's class who have their cell phones OFF and the median memory score of Year 11 Math's students in Mr R's class who have their cell phones ON.

I am going to investigate if their is a difference between the median memory score of Year 11 Math's students in Mr R's class who have their cellphones OFF (control group) and the median memory score of Year 11 Math's students in Mr R's class who have their cell phones ON (treatment group).

Lastly, make a **prediction** or **hypothesis**. It is a good idea to justify your prediction or hypothesis with evidence found in your research.

I will illustrate using the example from before:

Context or scenario: People who own smart phones are losing brain cells.

I predict that there will be no difference between the memory scores of the students who have their cell phones ON, and students who have their cell phones OFF.

This is because the results of the *observational* study by The University of Pittsburgh Cancer Institute found that there was no association (so no relationship) between hours of cell phone use and glioma cases. The study is not an experiment, but the results do contradict the claim that cellphones cause brain cell loss.

The article of the following page claims that the “Secret of happiness is money”.

You are required to go through the *Informed Contextual Knowledge* process outlined in this chapter using the article as the context or scenario.

There is space available for you to write on the following pages. Use either well structured paragraphs, or simply write notes.

Use the tick-boxes below to help track your progress

Purpose_I: the context..

❑ Research money, define it and find evidence about how money contributes to society.
❑ Research happiness, define it, and find out how you could measure it.
❑ Research any experiments / studies done about money and happiness, record how these studies where done, and what the important results are.

Purpose_II: the variables..

❑ Define your response variable, and justify using evidence from research on your choice.
❑ Define your explanatory variable.
❑ Define your experimental group, and justify your choice of the group.

Purpose_III: the question..

❑ Write a clear investigation question; make sure it has all the important features!
❑ Write a clear prediction / hypothesis, and justify your opinion with evidence from research.

stuff.co.nz

The secret to happiness is money, research shows

KATIE CHAPMAN Last updated 05:00 17/01/2013

The Beatles sang "can't buy me love" - but apparently you can buy happiness.

While som people may cling to the adage that "money can't buy happiness", research suggests the opposite is true, and the more money you have, the happier you'll be.

In fact, those sayings were simply a "delusion", the Auckland University study found.

"It's an axiom that money can't buy love. Our results, however, show that, to a certain extent, it can buy happiness and good health," said the report published last year in the New Zealand Journal of Psychology.

The findings will be made public later this year.

The report's lead author, Nikhil Sengupta, said the aim was to test the sayings that those with less money used to comfort themselves - such as "more money, more problems".

They discovered those sayings were "comforting illusions", and as people's bank balances grew, so did their levels of happiness, while stress melted away, he said.

"People who are wealthier are less stressed because they have a stronger ability - a perceived ability - to meet their needs."

Those on lower incomes worried about meeting basic needs, which increased stress, the report said.

"Poorer people tended to report more stress, less happiness, lower levels of satisfaction with their lives, and a lower ability to meet basic needs and life necessities.

"The comforting belief that poorer people are compensated by increased happiness and less stress is a delusion in the New Zealand context, and a dangerous one at that."

But when it came to the link between income and quality of life, the dollar amount wasn't as notable as a percentage jump in pay packets, Mr Sengupta said.

People whose income doubled from $10,000 to $20,000 reported the same boost in happiness as $100,000-earners whose income rose to $200,000.

"Income does predict how happy people are going to be with their lives . . . regardless of how much they earn."

Decision-makers now had to take note of the findings and stop holding on to the "false belief" that poorer people were at least happier, and look at "how we are going to lift them out of poverty to make them happier", he said.

The research was based on statistical analysis of phone surveys of 2746 women and 251 men.

The median household income of those surveyed was $65,000.

While the survey was done in 2008 the results were unlikely to vary much today, because international research showed similar results at different times, Mr Sengupta said.

Once you think you have given this your best shot, ask your teacher to look over it and give you feedback. This is a very important step in the learning process!

Chapter Two

EXPERIMENTAL DESIGN PRINCIPLES

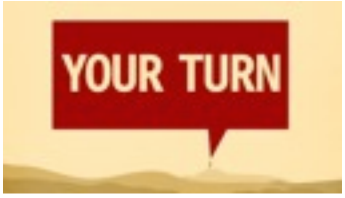
... how to design an experiment

Chapter one covered researching a context and creating a well-defined **purpose** and **hypothesis**. If you are still not sure about all of this, read chapter one carefully!

Soon it is time to *test* the **hypothesis**. But to do this with some degree of reliability, we need to *designed an experiment* that includes explanation of some key ideas, which are detailed in this chapter (hint: the red font from page 17 onwards).

An experiment needs to be well-planned-and-executed, taking into account all things that are practically possible (e.g. student noise, lighting, confounding variables), otherwise, any **cause** and **effect** conclusion you make from the collected data will not be meaningful!! ...and we don't want that.

On the page that follows is a list of important terms you need to know, that will help you design a well-planned-and-executed experiment.



Find definitions for these important features of experimental design. The terms with a * next to them are optional and may not always be relevant.

When designing an experiment, include:

Two independent groups	
Random allocation	
Treatment	
Treatment level	
Response variable	
Explanatory variable	
Experimental group	
Experimental Units	
Control group	
Treatment group	
Other sources of variation	
Placebo effect*	
Hawthorne effect*	
Confounding variables*	
Single blinding*	
Double blinding*	

Design feature explained: Experimental Group / Unit

What is an experimental group?

The group or collection of people that take part in your experiment. For example, Teachers at Ormiston Senior College or, Male Year 13 Statistics students at Ormiston Senior College.

When conducting an experiment, you must split your **experimental group** into two **independent groups**, by **randomly allocating** people to groups (this is covered in the next page).

**NOTE: you are not taking a sample from a population, as the experiment is about establishing a *cause and effect,* not using a sample to make an inference about the population.

What is an experimental unit?

A single person who is part of your **experimental group.**

How do you know whom should be part of an experimental group?

This comes back to the research and context given to you. You will also need to be realistic about who you can get to participate.

Why should I care?

The data from your experiment will tell you if the **treatment** causes the change in what you are measuring, for example, playing video games makes Year 13 males studying calculus more violent. You will also comment on how strong the **effect** is... that is, how certain is your conclusion? 20% 52% 96%?

Creating an **unbiased experimental group** will help balance out the **variation** in the results caused by differences in humans! People are different; these differences could be as simple as eye colour, and complicated as how funny someone thinks a joke is.

Design feature explained: Random allocation

What does random allocation mean?

Your experiment will conducted on two **independent groups**; that is a **control** group and **treatment** group.

Before and after measurements on one group only (**paired** comparison) are not part of this achievement standard.

For an experiment based on two **independent groups,** once you have selected an **experimental group**, you will sort each **experimental unit** (or person) into one of two groups; the **control** group or the **treatment** group.

You will need to use some sort of *random* method to sort each **experimental unit** (or person) into the **control** group or **treatment** group.

How do I randomly allocate my experimental group?

Below are some suggested tools to sort or allocate your **experimental group** into two independent groups.

Why should I care?

If **random allocation** does not occur it introduces *experimenter bias* into the experiment, which means any conclusion you make using the results of your experiment are likely to be **bias** or unfair.

It is important to control <u>all the different things you can</u> when running your experiment. All the things you can't control, for example - how funny a person thinks a joke is, differences in hair colour, differences in opinion; can be said to be evenly spread over the **experimental group** by **random allocation**.

Design feature explained: Blinding

What does Blinding mean?

Blinding (or **single blinding)** refers to giving the people in the **experimental group** a cover story.

For example, I run an experiment testing the hypothesis that wearing makeup causes a person to look younger. I use **random allocation** to sort my Year 11 mathematics class into a **control** group and **treatment** group. Each person in the **control** group will estimate the age of the person in the picture without makeup, each person in the **treatment** group will estimate the age of the person in the picture with makeup.

At no point do I tell anyone in the **control** group they are viewing a picture of a person without makeup.

At no point do I tell anyone in the **treatment** group they are viewing a picture of a person with makeup.

How do I do this?

Avoid telling the people you are experimenting on, what the **treatment** is, and what you are collecting or measuring.

Why should I care?

Cover story's are important, because if you tell the experimental group what the experiment is about or what the **treatment** is, there is a chance that the knowledge you have communicated can *bias* your data and thus your conclusion about **cause** and **effect**.

In some cases, you may wish to **double blind**. Experiments where both the **experimenter** and the **experimental unit** do not know what or when the **treatment** has been applied or not. For example, a glass of coke vs. a glass of pepsi.

Usually **single blinding** is enough.

Design feature explained: Other sources of variation

What are 'other sources of variation'?

For an experiment, you will try to control as many factors as possible when designing your experiment, for example:

- background noise
- other people walking into the classroom (distractions)
- students trying to sabotage the experiment
- levels of lights (e.g. is it dark enough to see the projected image on the screen?)
- temperature of the bucket of ice-water

What other examples can you think of? ...

It is important to conduct a fair or **unbiased** experiment so your results and conclusion are not **biased** in any way. The **treatment** is the only thing you want to **effect** the experiment.

"Control the things your can, those you cannot control should be balanced out during **random allocation** of **experimental units**"

How do I control these factors?

Careful planning of how the experiment will run, where it will happen, what you will say etc.

Why should I care?

Your experiment is ultimately about that **treatment** having an **effect** on something, for example, makeup on age estimates. In order to make a firm conclusion that the **treatment** is the cause, you must eliminate or 'hold-constant' all the factors you can.

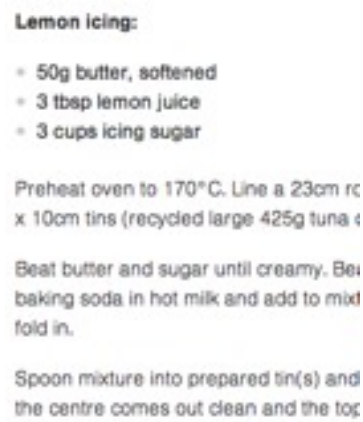
Design feature explained: Instructions...

What 'instructions'?

In your experimental design you must include <u>a clearly written sequence of steps</u>. These steps will help you run your experiment smoothly, and with minimal errors.

Can you give me an example?

Consider a recipe for a banana cake. You firstly need the ingredients (e.g. experimental units, treatment) and then you need the *method* or *order of mixing the ingredients.*


Lemon icing:

- 50g butter, softened
- 3 tbsp lemon juice
- 3 cups icing sugar

Preheat oven to 170°C. Line a 23cm rc
x 10cm tins (recycled large 425g tuna c

Beat butter and sugar until creamy. Bea
baking soda in hot milk and add to mixt
fold in.

Spoon mixture into prepared tin(s) and
the centre comes out clean and the top

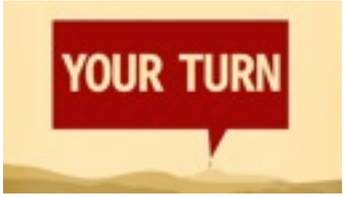
Why should I care?

If you don't have a well written experimental design, your actual experiment may not run as planned, and therefore the data from your experiment may <u>not be reliable</u> as the results of the experiment may be influenced by the differing methods of running the experiment.

Well written instructions show that you are taking into account **other sources of variation.**

Ideally include:

a. Instructions for the experimenter.
b. Instructions for for the experimental group.



Using one of the claims below, write out the full design of an experiment. Remember to include the important features!

“People who own smart phones are losing brain cells (see Chapter 1)
“Secret to happiness is money” (see Chapter 1)

Once you think you have given this your best shot, ask your teacher to look over it and give you feedback. This is a very important step in the learning process!

Chapter Three

EXPLORATORY DATA ANALYSIS

... analysing your data

In chapter one we discussed the importance of investigating the context and deciding on a clear **purpose** and **hypothesis** for an experiment. Chapter two is about the important aspects of designing an experiment, remembering that you must explain both your thinking and your instructions regarding 'how-to-conduct-your-experiment'.

This chapter is all about what to do with the data collected from an experiment. In other words, you will be looking at plots of the data and making statements *using your knowledge of statistics and the context (scenario) you are investigating.*

These statements should describe the picture that the data is painting. "Context" or making links with the experiment, is extremely important when making statements, because if you remove the context from the data ... the data is no longer able to tell the *real world story.*

HOW TO conduct data analysis of an experiment.

<u>What will I be analysing?</u>
You will be producing the familiar dot plot, and box & whisker plot; this time you will comparing the data of the *control* group with the data of the *treatment* group. Your job is to describe what you are seeing in the data from the experiment. The 3 sections are:

1. the Median,
2. the symmetry and/or spread around the Median, and
3. describing any unusual features.

The next page will explain each of these parts in more detail, and at the end of this chapter is your opportunity to practice writing analysis statements and get them checked by your teacher!

<u>How will I write these analysis statements?</u>
The 3 main steps for writing good statements for your experimental data are:

WHAT do you SEE?

S = Statement; using the correct (but obvious) language to describe the feature.

N = Numbers; where possible quantify your statement, e.g. IQR = 56.

C = Context; <u>always</u> write your statement with the scenario in mind.

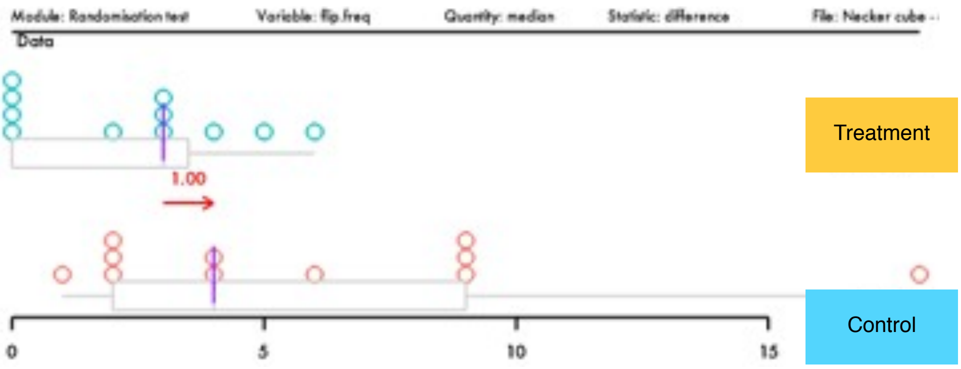
WHAT does it MEAN?

This part will be tricky, and your answers will vary from experiment to experiment. As a guideline, what could the feature you are writing about (e.g. the median), be telling you about the overall / final results of the experiment?

WHY should I believe you?

Using evidence from research (remember to reference!), or even events in your experiment or facts about your experimental group, to *justify* or ***back up*** *"what you see".*
******it may not always make sense to do this part, but if you can, please do.*

A really quick example.
See the exemplars later in this book, but for now have look at this.



Median

The median number of flips for the control group is 4 flips *which is greater than* the median number of flips for the treatment group of 3 flips, by 1 flip.

This suggests that once chance is eliminated as an influencing factor in the results (that is, it’s not a fluke that this difference of 1 flip exists), I should be able to conclude that the *treatment* image causes the number of flips observed to decrease.

A possible reason for this is, in a similar test run by the University of Michigan they concluded that it was easier for students to concentrate on the the *control* image rather than the *treatment* image.

Symmetry - Skewness

<u>In the treatment group</u>, there are 8 out of 11 students who saw between 0 and 3 flips only. A possible reason is that the image used for the *treatment* group requires more concentration than the image used from the *control* group... and because the experimental group were from a low level math’s class, it is likely they don’t have good concentration skills to begin with!

<u>In the control group</u>, the number of flips observed are spread out from 1 flip all the way to 15 flips. There is no symmetry.

Unusual features

One student in the control group saw 17 flips. I don’t think this student was being honest with his answers, as he did not follow instructions in the experiment.

You must NOT bring in any 'sampling / population' ideas from statistical inference.

Statistical inference is linking *what you see in the sample data* to *what you expect to see in the population data*. An inferencial investigation question looks like:

"I will investigate the difference between the median income ($000) of males and the median income ($000) of females, who are between 18 - 65 who live and work in New Zealand".

The difference between the sample median incomes (sample statistic) will help us make a prediction or judgement about population median incomes (population statistic).

Statistical experiments is restricted only to the control and treatment group of your experiment. The idea is to see if the *treatment* causes some change in the *response variable* of the experimental group only. An experimental investigation question looks like:

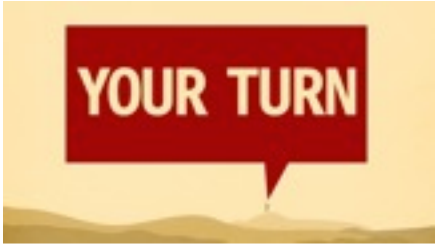
"I wonder if the median number of flips observed by the control group* will be higher than the median number of flips observed by the treatment group*?".

***Of course, you will need to define the control and treatment group in your report.*

Once chance is eliminated as a factor in the experiment (see the Randomisation Testing section later in the workbook) we will be able to write a **cause** and **effect** statement.

The below table summarises what you need to 'see' in the analysis part of your investigation.

Statistical Inference	Statistical Experiment
Centre: median	Centre: median
Centre: shape	
Spread: IQR / Range	Symmetry - Skewness only
Spread: Shift	
Spread: Overlap	
Unusual features	Unusual features



What claim is this experiment testing?

What reasons would someone have for running this experiment?

Background information on the experimental design

32 Year 12 students from Avondale College were **randomly allocated** into two **experimental groups**, the **control** group (Group A) and **treatment** group (Group B). The **control** group was placed in a classroom, and each person was given a blank piece of paper and pen. The **experimenter** then projected the image of the female without makeup onto the screen, and each person was asked to write down an estimate of the females' age in years, with a 30 sec time limit. Each piece of paper was collected up by the **experimenter.**

The experiment was repeated for the **treatment** group, except an image of the same female with makeup was projected on the screen

Here is the data set from the experiment:

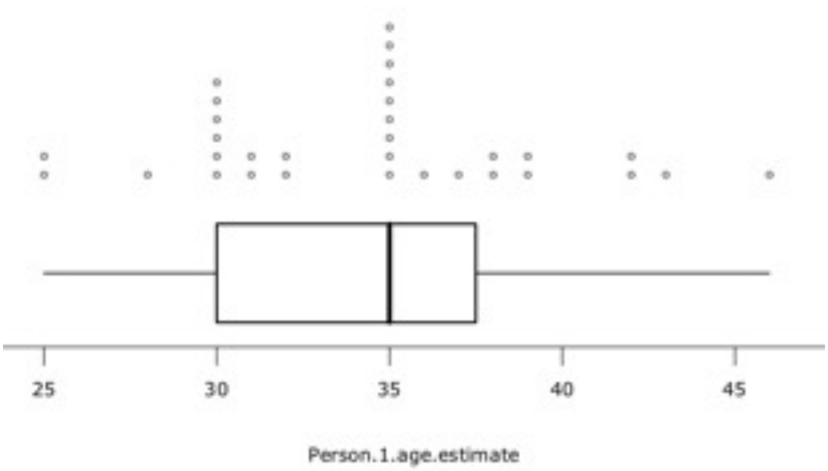
Face	**Age Estimate**	**Face**	**Age Estimate**
No makeup	30	Makeup	38
No makeup	35	Makeup	25
No makeup	29	Makeup	28
No makeup	29	Makeup	34
No makeup	25	Makeup	26
No makeup	32	Makeup	35
No makeup	28	Makeup	35
No makeup	26	Makeup	25
No makeup	28	Makeup	28
No makeup	26	Makeup	38
No makeup	30	Makeup	29
No makeup	48	Makeup	27
No makeup	43	Makeup	43
No makeup	32	Makeup	28
No makeup	33	Makeup	23
No makeup	25	Makeup	25

Analysis Box and Whisker plots: the overall age estimate.

This is a plot of <u>all</u> the estimates of the females' age. You will notice the graph is not divided by group, nor is it obvious which age estimate belongs to which group.

This singular graph is only to help you practice writing the required analysis statements, before you start *comparing the control and treatment groups.*

The point of looking at this plot is to see if there are any particular age's that people tend to estimate. In our analysis we will focus on these three features <u>only</u>; the **median**, **symmetry or skewness**, and **unusual features.**



Have a go at writing some **statements**. The teacher will give you suggested answers in class.

My statement about the **median** of the overall age estimate...

Teachers statement about the **median** of the overall age estimate...

My statement about the **symmetry or skewness** of the overall age estimate...

Teachers statement about the **symmetry or skewness** of the overall age estimate...

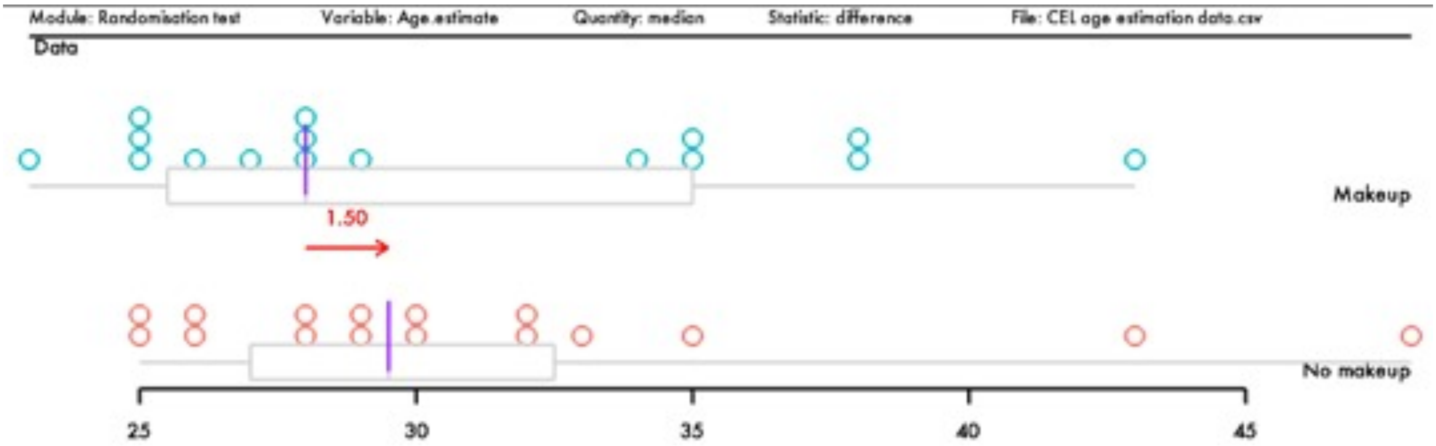
My statement about any **unusual features** of age estimates...

Teachers statement about any **unusual features** of age estimates...

Analysis Box and Whisker plots: comparing the two groups.

Below is a plot comparing the age estimates made by the **control** group verses the age estimates made by the **treatment** group.

The point of looking at this plot is to compare the age estimates of the **control** group with the **treatment group**. In our analysis we will focus on comparing these three features only; the **median**, **symmetry or skewness**, and **interquartile range** (**IQR).**



Hint: see page 37 & 38 to find out how to produce this graph.

My statement *comparing* the **median** of the age estimates...

Teachers statement *comparing* the **median** of the age estimates...

My statement *comparing* the **symmetry or skewness** of the overall age estimate...

Teachers statement *comparing* the **symmetry or skewness** of the overall age estimate...

My statement *comparing* the **unusual features** of age estimates...

Teachers statement *comparing* the **unusual features** of age estimates...

Chapter Four

RANDOMISATION TESTING

... and concluding remarks.

Analysis: Randomisation testing.

We assume an experiment has been designed taking into account all factors possible (**variation from other sources)** and the **control** and **treatment** groups were determined by **random allocation**.

Therefore, all efforts have been made to eliminate **bias** or other influencing factors / variables from the experiment, and our conclusion should produce a fair and valid **cause** and **effect** conclusion for the experimental group.

The only *other* source of variation we have NOT yet taken in account, is “Chance”.

Recall, that a **control** and **treatment** group will be formed using **random allocation**. Then, what if it’s a *coincidence* that one group has more males that the other? Or what if it’s a coincidence that one group contains all the ‘smarter’ people?

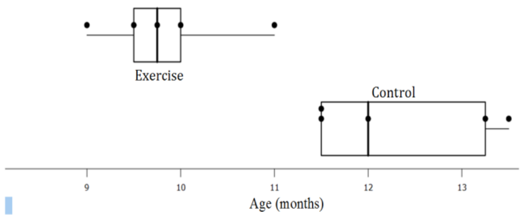
We will use what is called a randomisation test to deal with “Chance”.

Randomisation Tests

The Big Ideas:

Can we rule out the 'chance alone' explanation because it's implausible?

1. Looking at the world using data is like looking through a window with ripples in the glass: *'What we see is not quite the way it really is'*
2. The **possibility** that **chance acting alone** has produced the observed difference pattern
3. **Assessing** the plausibility of the 'chance alone' explanation
4. **Forming a conclusion** about what the data are telling us



What do we mean by *chance alone*?

Under chance alone the difference pattern we observe is purely and simply the result of the luck-of-the-draw as to which units (e.g., babies) *just happened by chance* to be assigned to which group and nothing else.

Under chance alone, the value of the measure made on a unit (e.g. a baby's walking age) has nothing to do with the group to which that unit (baby) had been assigned, i.e., **it doesn't matter to which group the units** (babies) **were randomly assigned, each unit** (baby) **would still have the same actual value of the measure made** (walking age)**.**

How do we assess the plausibility of the chance alone explanation?

Randomisation Test (3Rs: ***R****e-randomise,* ***R****epeat,* ***R****eview*)

Re-randomise:

Ignore each unit's (baby's) assigned group, randomly re-assign each unit to a group – always keeping its actual measure, that is, **re-assign as if chance is acting alone**. Determine and record the value of the statistic of interest (e.g., $\bar{x}_{Control} - \bar{x}_{Exercise}$).

Repeat:

Repeat the re-randomisation a large number of times. Plot the recorded values of the statistic of interest to generate the re-randomisation distribution under chance alone. See what are likely and what are unlikely (hard to get) values for the statistic of interest (e.g., $\bar{x}_{Control} - \bar{x}_{Exercise}$) when chance is acting alone.

Review:

Review the 'chance-is-acting-alone' explanation by locating the observed difference (e.g., the actual observed data value of $\bar{x}_{Control} - \bar{x}_{Exercise}$) on the re-randomisation distribution.

- An observed difference located far out in the tails discredits the chance-alone explanation. The observed difference would be highly unlikely if chance were acting alone, therefore chance probably isn't acting alone. It gives us: evidence that chance is not acting alone; evidence that something other than just chance is acting.

- An observed difference not located far out in the tails does not discredit the 'chance-alone' explanation. The observed difference would be not be surprising if chance were acting alone, therefore chance <u>*could*</u> be acting alone. But equally, something as well as 'chance' <u>*could*</u> also be acting.

Guide Lines for Assessing 'Chance alone'

When the **tail proportion** in the re-randomisation distribution is **less than 10%** then:

- the observed difference would be unlikely when chance is acting alone, therefore it's a fairly safe bet chance isn't acting alone.
- we have evidence **against** 'chance is acting alone'
- we have evidence that **chance is not acting alone**

When the **tail proportion** in the re-randomisation distribution is bigger than **about 10%** then:

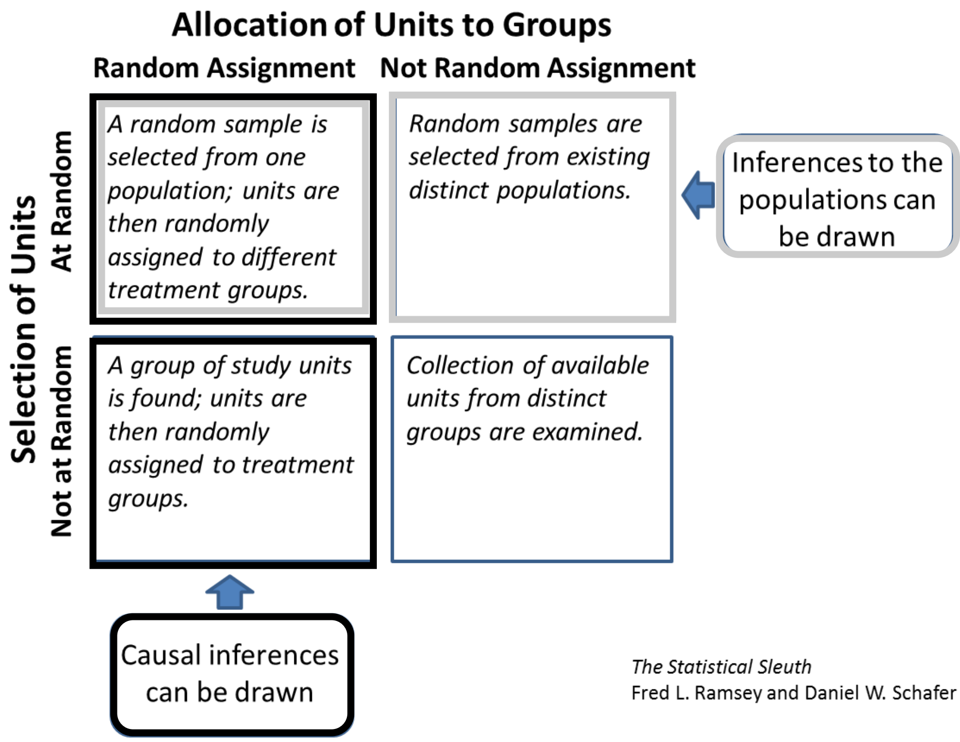
- the observed difference is not unusual when chance is acting alone, therefore chance **COULD** be acting alone
- we have **NO** evidence **against** 'chance is acting alone'
- 'chance' COULD be acting alone ***OR*** something else as well as 'chance' COULD also be acting (we don't have enough information to determine which one of these two possibilities applies).

The Link between the Randomisation (data production) and the Conclusion

Randomisation (data production)

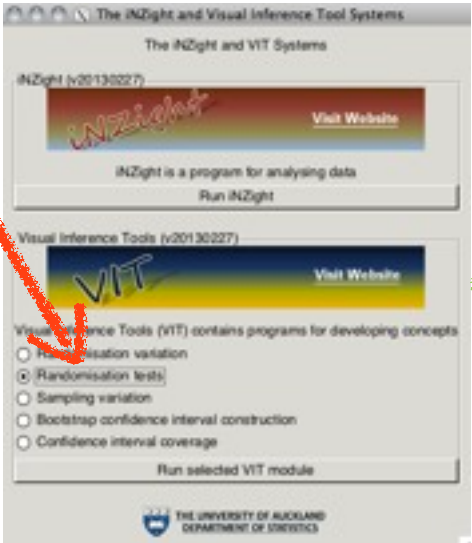
1. Random assignment of units to treatment groups (an experiment)
2. Random sampling of units from a population or populations

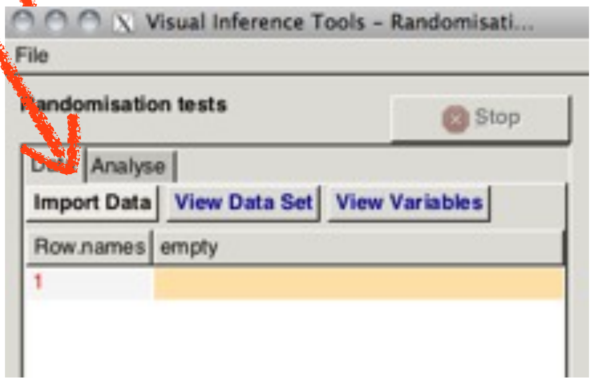
Where is the randomisation and what inference does it support?



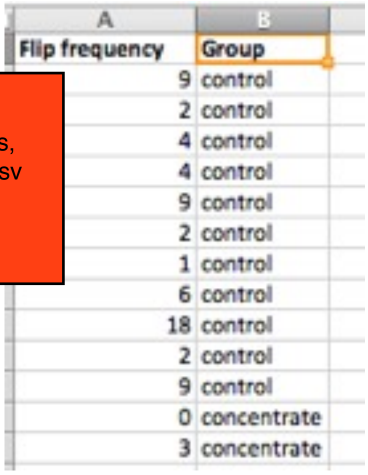
The Statistical Sleuth
Fred L. Ramsey and Daniel W. Schafer

HOW TO: Randomisation testing.

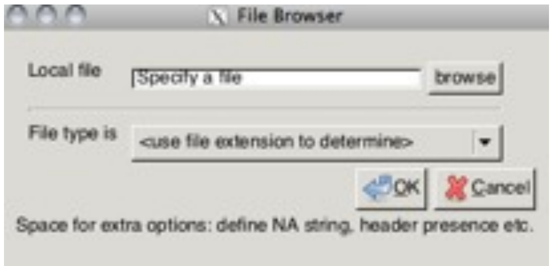


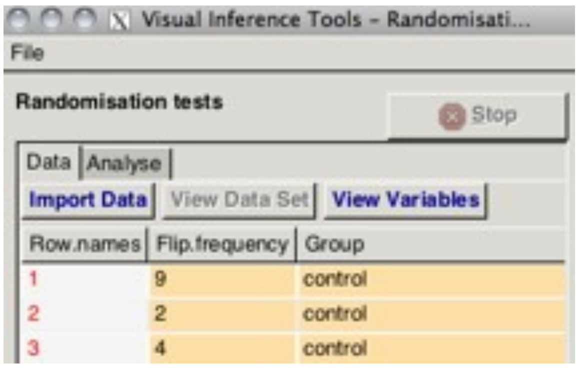


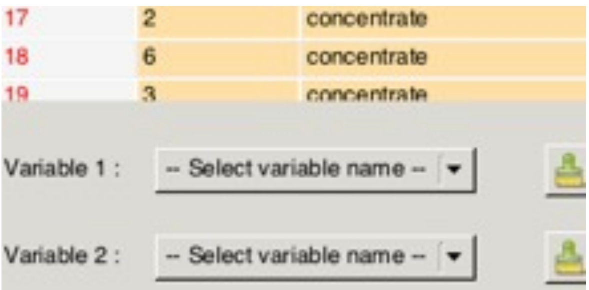
Check data is formatted like this, and saved as *.csv (OSX) or *.xls (Windows).

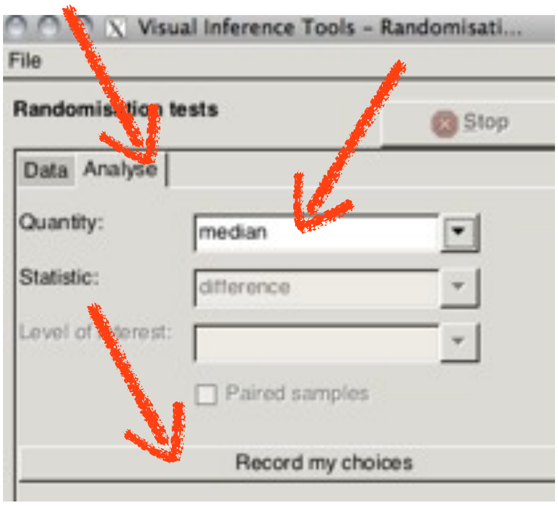


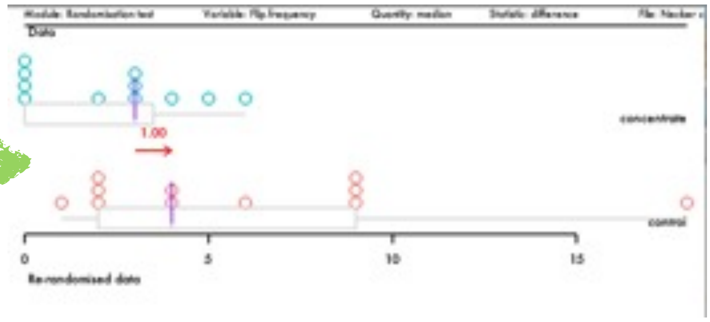
A	B
Flip frequency	Group
9	control
2	control
4	control
4	control
9	control
2	control
1	control
6	control
18	control
2	control
9	control
0	concentrate
3	concentrate

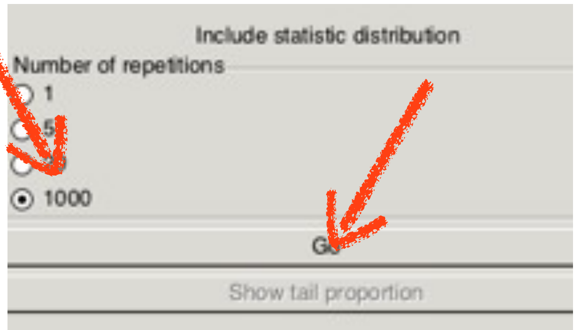


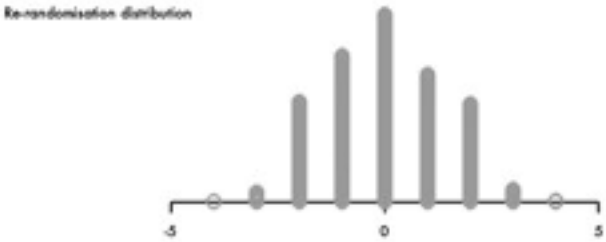


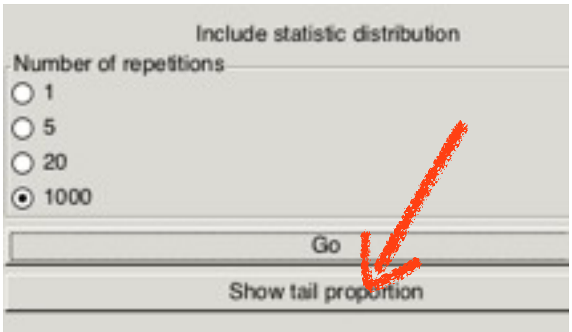


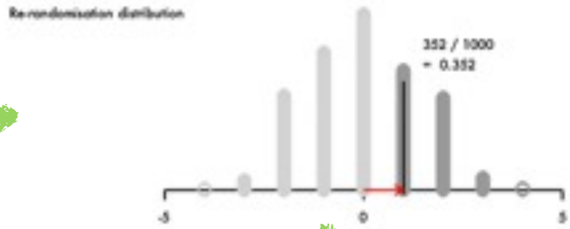

Visual Inference Tools - Randomisati...
File
Randomisation tests
Stop
Data
Analyse
Quantity:
median
Statistic:
difference
Level of interest:
Paired samples
Record my choices


Data
1.00
concentrate
control
0
5
10
15
Re-randomised data


Include statistic distribution
Number of repetitions
1
1000
Go
Show tail proportion


Re-randomisation distribution
-5
0
5


Include statistic distribution
Number of repetitions
1
5
20
1000
Go
Show tail proportion

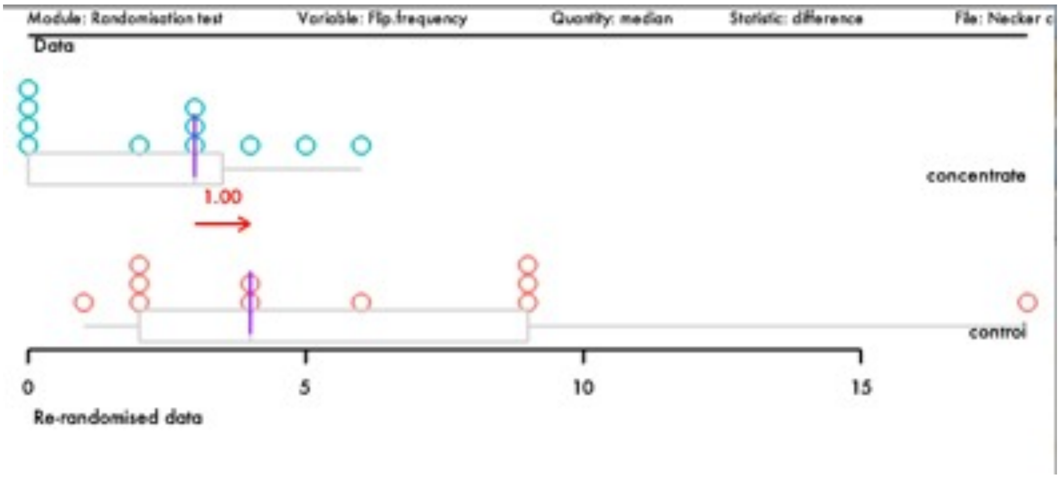

Re-randomisation distribution
352 / 1000
= 0.352
-5
0
5


HOORAY!

...explaining what iNZight has done.

Again, here is a quick example:

After importing the data and 'recording your choice', the results of your experiment are displayed on a dot plot and box & whisker graph.

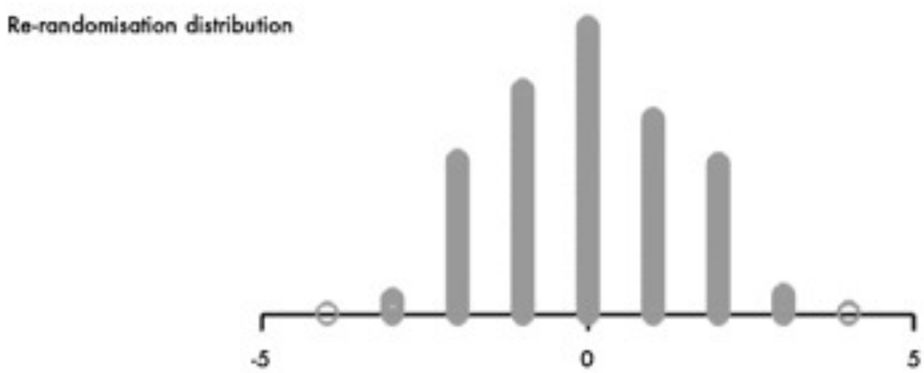


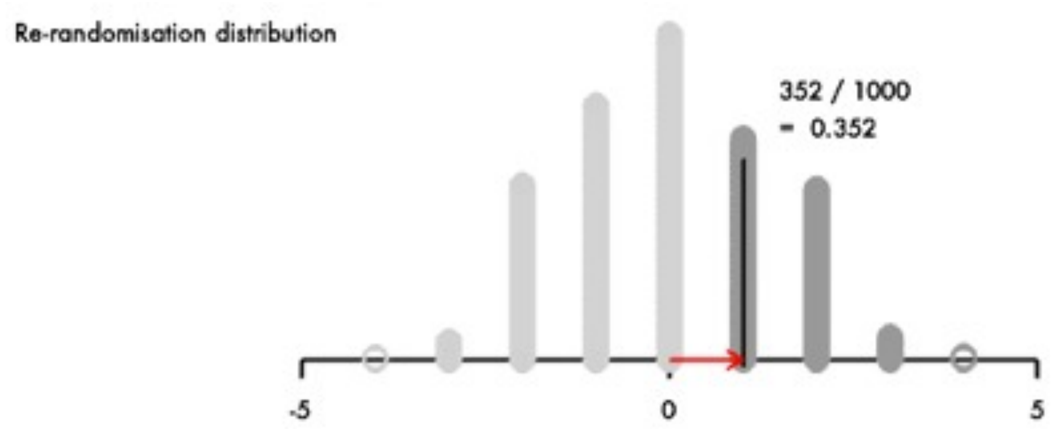
From the experimental data, we can see a difference is 1 flip between the medians for the **control** and **treatment** group (red arrow).

Is this difference of 1 flip by chance alone (that is, simply because of random variation through **random allocation**)? Is there enough of a difference that is caused by the **treatment,** that we can make a **cause** and **effect** link between the **treatment** and the difference between the number of flips?

The difference that we are seeing must be large or *significant* enough so that it's not a fluke; we try to minimise the effect of 'chance' (or random variation) by using **random allocation**. The data collected from the experiment will be mixed up and then re-allocated into groups (using **random allocation**), the medians for the new control and new treatment group will be recorded, and a difference in the re-randomised medians will be recorded, and compared to the original difference in the medians of 1 flip.

We will need to calculate the difference in re-randomised medians 1000 times.





The red arrow is the original difference in the medians of 1 flip. Looking at the distribution of re-randomised differences between the medians we can already tell that our result of 1 flip is a fairly typical result (can you explain why?).

Interpreting the tail proportion.

The tail proportion tells us the at least 352 times out of 1000, we will get a difference in median number of flips which is 1 or more.

This means that the difference of 1 flip between the control and treatment group is NOT unusual (rather, it is fairly typical!).

Therefore we have no evidence that anything else other than chance is giving us our result of 1 flip.

Therefore we cannot conclude that the **treatment (cause)** applied will **effect** the number of flips a person in the experimental group sees.

When CAN we conclude that the treatment (cause) applied will effect the number of flips a person in the experimental group sees ???

When the TAIL PROPORTION is LESS THAN OR EQUAL TO 10%.

The conclusion: the importance of cause and effect.

What is and **cause** and **effect?**

Simply put, the **cause** is the **treatment** that is applied in your experiment. The *change to your data* that happens when you apply the **treatment**, is the **effect**. For example, using makeup (the **cause**) might change how old a person looks (the **effect**), as measured by age.

Why should I care?

The purpose of your experiment is always testing if the **treatment** changes a **variable** you choose to measure.

Writing a conclusion: the steps required

Numbers 1 - 4 below give you a quick idea of what you should cover when making the conclusion. The rest of this chapter will give you further details.

1. **Write a pre-conclusion based on your experimental data.**

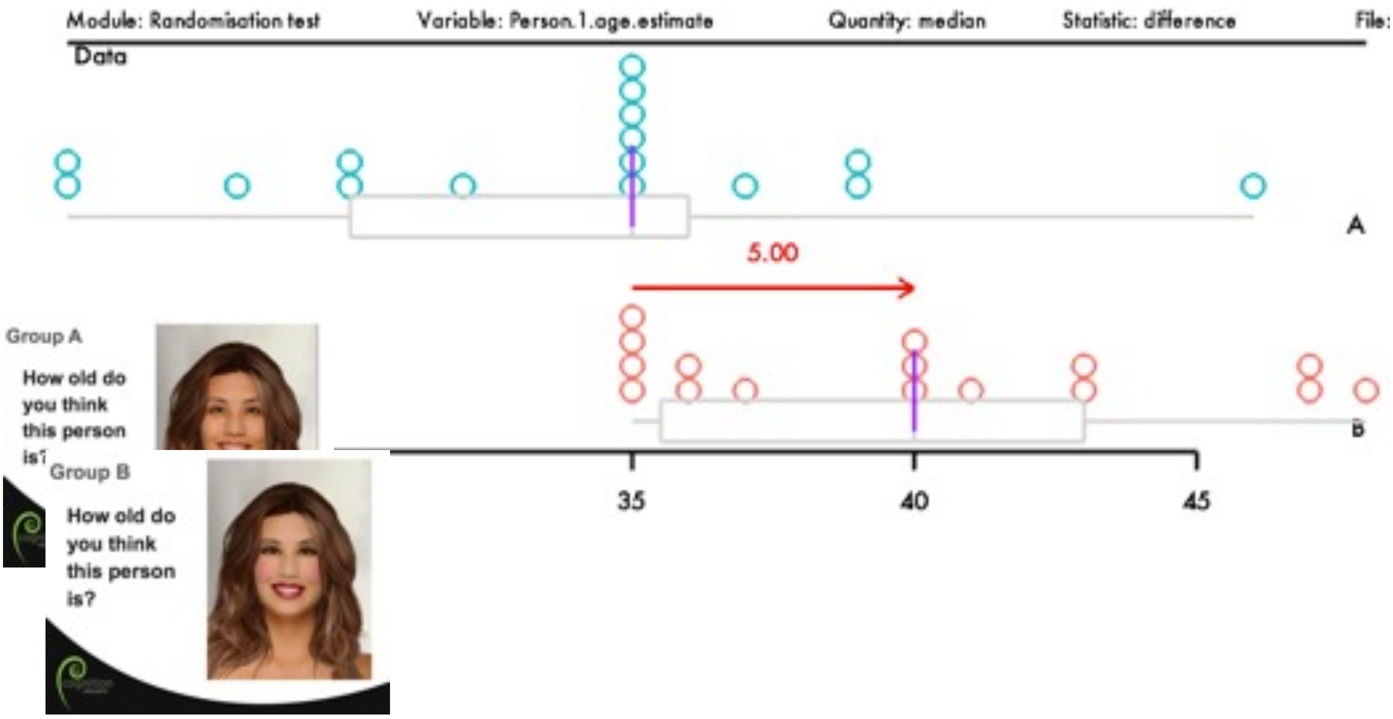
2. **Write a statement about unbiased experimental design, and the need to perform a randomisation test.**

3. **Using the results from your randomisation test, write a conclusion that details cause and effect**

4. **Write down closing arguments, supported by contextual evidence; detail how your experiment could be improved and why.**

Write a pre-conclusion based on your experimental data

Remember the age estimate plot from chapter two about age estimates between the **treatment** and **control** group? We will now make a concluding statement about what the plot shows us about the **median**(s).



My pre-conclusion...

Teachers pre-conclusion...

2

Write a statement about <u>unbiased experimental design</u>, and the need to perform a randomisation test.

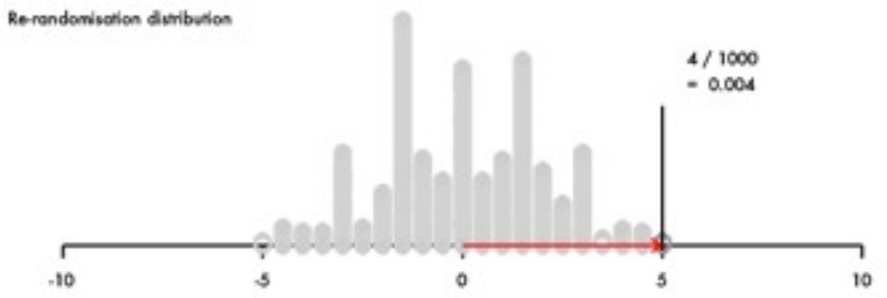
Recall, if you design an experiment that controls all the factors you can realistically control (for example, noise, light, distractions, cheating) then your data and thus any conclusions you make are likely to be **unbiased**.

In addition, we need to perform a **randomisation test**. This is so that any difference in the medians we observe is not happening by fluke or **chance** alone.

My statement about unbiased experiment design, and the need to perform a randomisation test...
Teachers statement about unbiased experiment design, and the need to perform a randomisation test...

3

Using the results from your randomisation test, write a conclusion that details <u>cause</u> and <u>effect</u>. Compare this result with the hypothesis.



My cause the effect conclusion...
Teachers cause and effect conclusion...

4

Write down closing arguments, supported by contextual evidence; detail how your experiment could be improved and why.

Compare your **conclusion** with experiments that have been done testing the effect of makeup on age. What result do these experiments have?

Compare your **conclusion** with newspaper articles or research papers (theories) that already exist. Are the articles or research drawing a similar conclusion to your experiment?

If you could, name two or three things could you change about your **experimental design** that would make your data and therefore the **conclusion** more accurate.

Once you think you have given this your best shot, ask your teacher to look over it and give you feedback. This is a very important step in the learning process!

Chapter Five

Practice writing a report.

The Final Report

A report is a document that is written in the *past tense.* In basic terms, it is a storey of your journey along the 'experiment' road.

The strength of your grade will based on your **Final Report**. Once you have:

a. conducted your experiment,
b. performed your data analysis and,
c. of course taken notes along the way...

.... you will then be ready to write your **Final Report.**

WRITE sentences appropriate to answer each box. This is a practice report, using a familiar context (if you have read the earlier chapters!!)

**For background information, see page 29. Assume the data is from a different class.

Introduction

Background
Problem
Expectation

Why do people (in particular, females) use makeup?

What **treatment** was used in the experiment?

What **variable** was chosen to measure the **effect** on?

Write down the clear and detailed **investigation question** for the experiment.

What do you expect the conclusion to be? This is the **hypothesis**.

Why select a **control** group and **treatment** group selected? Why not a **paired** comparison?

Who will be part of the experiment?

Where and when will the experiment take place?

What **random allocation** tool will be used to form the **control** group and **treatment** group?

Why is **random allocation** necessary in this experiment?

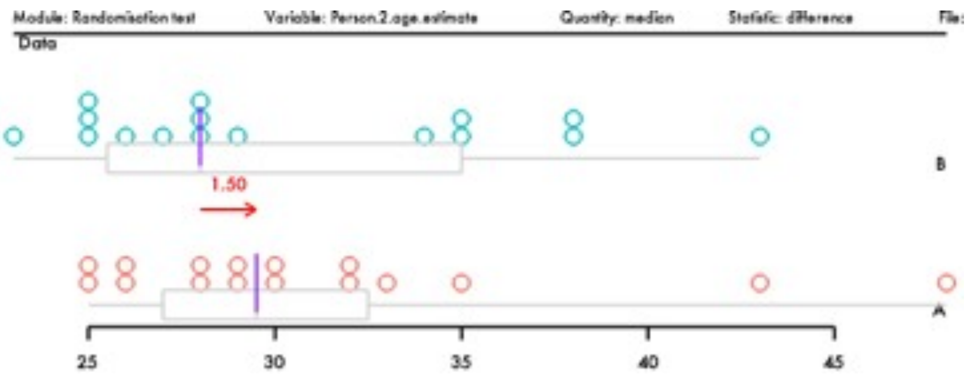
What other **factors** will you be able to control? (be realistic)

What **factors** will you be not be able to control?

Write a set of instructions, about "how to run the experiment".

Results

Dot and box plots and summary stats
Descriptive statements
Randomisation test



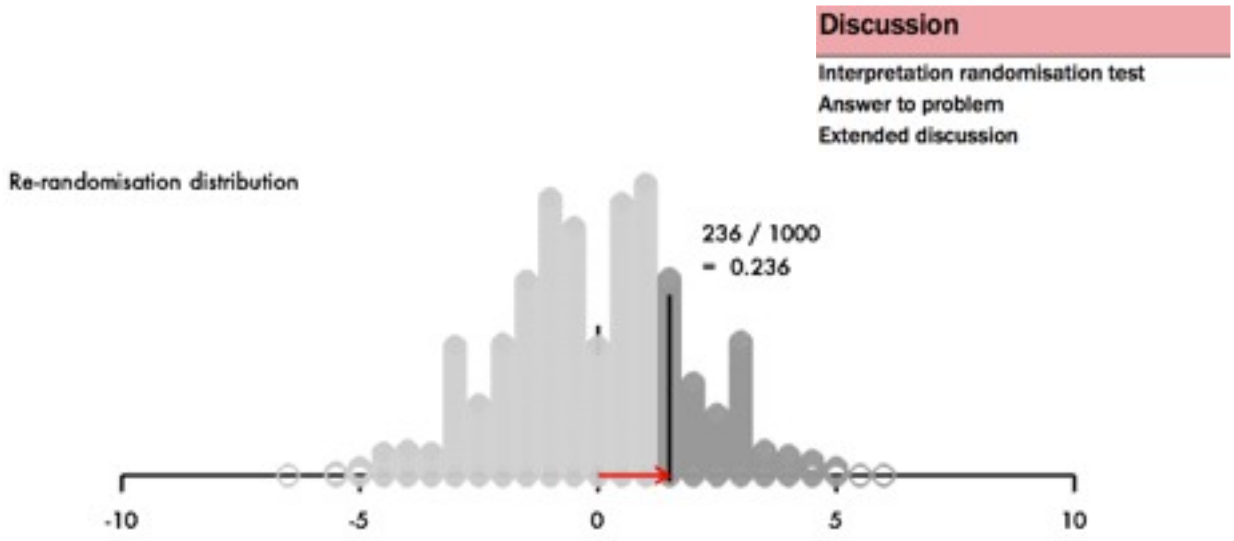
Provide an analysis of the centre (medians) for the **treatment** and **control** groups.

Provide an analysis of the **symmetry and skewness** for the **treatment** and **control** groups.

Provide an analysis of anything **unusual** for the **treatment** and **control** groups.

1 Write a pre-conclusion.

2 What makes the **experimental design**, **unbiased**? Why do we care?



2 Why do we need a randomisation test?

3 What does this tail proportion of 236 / 1000 tell us?

3 Is the observed difference in the medians unusual?

3 We have no evidence to suggest......

3

Conclude the experiment. Remember to mention **cause** and **effect**.

4

Does the conclusion match the hypothesis?
Does your conclusion match existing research?
What could you have realistically done to improve your experiment?
Always explain <u>why</u> and include any other comments.

Once you think you have given this your best shot, ask your teacher to look over it and give you feedback. This is a very important step in the learning process!

Chapter Six

Exemplars

NOT YET PUBLISHED.

Available late May 2014.