

|  |
| --- |
| EVALUATION OF STATISTICAL REPORTS |
| Part 1 Non-sampling errors in polls and surveys. Resource prepared by Dru Rose |
|  |
| This unit is designed to be taught over a one and a half week period. A possible teaching plan is given below and associated resources are provided in the Resource Pack. |
| EVALUATION OF STATISTICAL REPORTS |

Part 1 Non-sampling errors in polls & surveys

Unit Overview

# Teaching Notes

• • •

Notes about activity objectives will be made in this panel.

References to resources used such as power-point slides will be referenced in this panel.

UNIT OVERVIEW

|  |  |  |
| --- | --- | --- |
| Lesson | Topic | Description |
| 1 | Introduction | Class poll to review and extend ideas on sampling  Structure of a media article: *driving infringements article + template* |
| 2 | Sampling in the wider NZ population | Extend ideas on sampling to samples of the NZ population-distribution of key characteristics |
| 3 | Non-sampling errors  Literacy Strategies | Power-point presentation with examples and stories  Pre-reading to get the gist, read aloud asking worry Qs, note-taking *(e.g.Road Tolls article + worry Qs sheet + writing frame)* |
| 4 | Practice using worry Qs and writing frame | e.g.DVT article: What is DVT?  get the gist, ask and answer worry Qs, note 2 positive aspects and 2 concerns |
| 5 | Practice with less familiar context | e.g.ANZUS article: What is ANZUS?  Reading to gain meaning from context  Critically evaluate article |
| 6 | Introduce Political polls | Power-point presentation on research companies-compare methodologies  Compare 2 recent political polls from different companies |
| 7 | Practice Exam style question | See **new** NZQA exemplar*- probably no “worry” Q sheet or* *writing scaffold* |

**LESSON 1 – ACTIVITY 1: What is a poll?**

**Quick class poll**: e.g. **“**What type of driver licence do you have?” (Full, Restricted, Learner, None)

# Teaching Notes

• • •

Biased sample:-all Year 13 from one school etc-how might class results differ from other places in NZ?

* Target population is “school children in NZ aged 16 and over”
* Senior needs to be defined as “aged 16 and over”

Poll %s are “categorical data”-sampling variability is much greater than for quantitative data (means, medians). Large sample sizes are needed in polls-eg 500+ -a new idea for students

Key Features are numbered on the writing template. Students draw around them or number them on their copy of the article and then fill in the writing template.-You will need to show them how to do this initially and go over the answers in class discussion.

Studying the structure of media articles helps students to unpack them to find the information they require.

Introduce the term *Poll Percentage*

Use results to probe existing knowledge base regarding:

*target population, sampling frame, representativeness of sample, sampling variability, sample size, inference etc*

e.g. Write up an incorrect inference such as : *“ ---% of senior students can legally drive by themselves”*  “Is this a valid conclusion?” why not?

*(You may choose to show slides1 to 3 in the non-sampling error power-point)*

What is likely to happen if I ask the same poll Q in another class?-*revise sampling variability and its relationship to sample size*

You may choose to show Wild’s animations for the sampling variability of categorical data (see Resource Pack).

**LESSON 1 – ACTIVITY 2** (may extend into homework)

**How are Media Articles on Polls and Surveys Written ?**

Journalists are taught to write media articles in a particular style. Learning about this style enables a reader to find specific information easily and quickly.

**Group activity:** Give students the writing template entitled: *How is a typical media report on a poll or survey organised?* and a suitable article such as :**“One driver in two admits sometimes ignoring speed limits” (**template, article and answers are in Resource Pack)

**BIG IDEAS: *Random Sampling in the real world***

# Teaching Notes

• • •

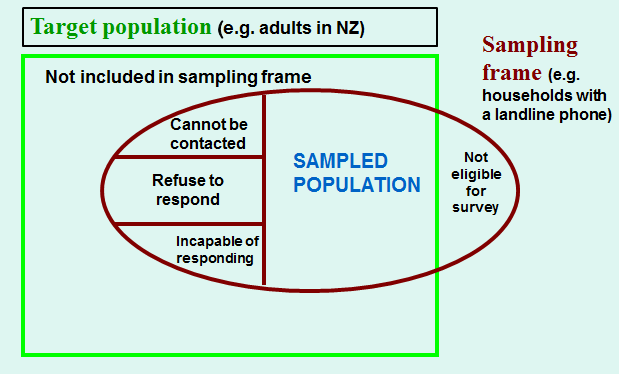
Encourage students to include the **sample size** when describing the sample group to help distinguish it from the target group.

**REAL WORLD:**

Random Digit Dialing (RDD) and Computer Assisted Telephone Interviewing (CATI).-

“CRITICAL” in this course means both POSITIVE and NEGATIVE aspects of a study.

Although the main purpose of this activity is to show students the structure of a typical media article and the type of information usually found, it provides another opportunity for consolidating and extending their statistical knowledge base.



The class discussion should address the distinction between *target population (*adult drivers in NZ)*, sample frame(*households with a landline phone)*, the sample group (*500 NZ drivers)– **This distinction is a difficulty for students.**

You will need to show students that the **method** is in the fine print under the graphic.

The students view of sampling is currently that associated with classroom activities (numbering the class and using the Ran# on a calculator). You could go the DIGIPOLL website <http://www.digipoll.com> and show students how a market research company does it.

**Margin of error:-**Show them where to find it but at this point but simply say it is connected with sampling variability and will be **dealt with in detail later in the course**. (See Workshop 1 material on sampling error).

**“Developing a critical stance”-**informal only at this stage-what seems good and what concerns do they have if any ? (e.g.*GOOD: random sampling, low margin of error, CONCERNS: limited options given for response-e.g. alcohol and not wearing a seat-belt not included so offenders could answer “perfect” driver”-is anyone “perfect”?, age-group results missing in the graphic).*

**LESSON 2: What do the demographics of the New Zealand population look like?**

# Teaching Notes

• • •

Potential resources:

[http://www.stats.govt.nz/tools\_and\_services/services/schools\_corner/SURF%20for%20schools/census.aspx](https://webmail.westlakegirls.school.nz/owa/redir.aspx?C=2862ca6ad67a4e5788f0ce309e036813&URL=http%3a%2f%2fwww.stats.govt.nz%2ftools_and_services%2fservices%2fschools_corner%2fSURF%2520for%2520schools%2fcensus.aspx)

[http://www.censusatschool.org.nz/resources/data-analysis-tools/](https://webmail.westlakegirls.school.nz/owa/redir.aspx?C=2862ca6ad67a4e5788f0ce309e036813&URL=http%3a%2f%2fwww.censusatschool.org.nz%2fresources%2fdata-analysis-tools%2f)

**Get students to predict what a distribution will look like before they go to the software.**

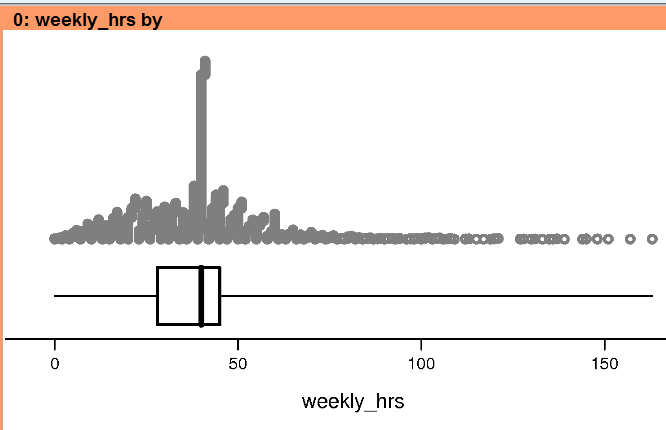
**Describe the SHAPE of each distribution in context-**students should be adopting formal statistical language at this level.

**A common student misconception:** Many students think that key demographic groups in a survey should be sampled in equal numbers e.g. a survey on attitudes to smoking should sample equal numbers of smokers and non-smokers even though smokers now make up only a small proportion of the NZ population.

This idea is key a misunderstanding about the nature and purpose of a random sample.

The demographics of the New Zealand population are very complex. Key characteristics such as age-groups, income groups, ethnic groups etc are not generally distributed in equal proportions (gender being the exception). Research groups use complex weighting techniques to ensure that their samples *accurately reflect the demographics of the New Zealand population*. To understand what this statement means students need to explore the distribution of these key characteristics in the New Zealand population and in their own locality.

The SURFS on the Statistics NZ schools corner are EXCEL files which could also be saved as csv files and imported into iNZight. The income SURF on censusatschool can be used in Data Viewer to show the distribution of any of the variables listed, not just income.-34000 records in this one with the added advantage that sample size can be varied to show the effect of this.



**LESSON 3** – ACTIVITY 1: **Formal introduction of non-sampling errors and survey concerns**

Go through the non-sampling errors power-point with students-there is a contextual example to illustrate each bias or survey concern but get students to add to it by contributing stories from their own experience.

Telling stories (describing situations in context) is a powerful tool for teaching this topic. –when students are trying to make sense of a particular situation ( e.g. in a media article), they compare and contrast it with those they are already familiar with.

**LESSON 3** – ACTIVITY 2: **Literacy Support**

**Resources needed**: *“*Writing frame *for critically evaluating a report”,* a suitable article e.g*.* **”Most favour motorway tolls to plug money gap, survey finds”,**worry questions sheet.

**Statistically-based reports are not like other texts students read in school-**all students seem to need some help making sense of the texts and tables.

The following strategies seem to help.

**Pre-reading:** GETTING THE GIST

The first read-through should be a skim-read to establish what the article is about and unpack-the structure. –They write a 3 sentence summary: e.g. one each on purpose, method, and key finding.

# Teaching Notes

• • •

Re-phrase student contributions using **formal statistical language** e.g. when a *respondent* may not tell the truth in a survey on a sensitive issue survey it is called *behavioural bias*

Choose articles for their potential to highlight specific non-sampling errors and survey concerns

**Teacher Modelling** of the strategies you want students to adopt is essential.

Encourage students to read right through BEFORE attempting to find information.

**While-reading strategies:** ASKING “WORRY” QUESTIONS and NOTE-TAKING.

# Teaching Notes

• • •

.Students seem to make the most progress in this topic when collaborating in groups.

Evaluation involves looking at both the poll itself and how it was portrayed in the article eg anything missing?-were other questions asked?

More detailed information on this survey is available from <http://www.horizonpoll.co.nz/page/241/majority-bac>

**Model** how to use the “worry” questions sheet and the writing frame by **reading the text aloud, asking the questions,** and getting the students to contribute the answers.

**Model** how to fill in a writing template.

Let students discuss in groups the positive and negative aspects of the report and then have a report back session with the whole class.

If using the road tolls survey:

****

**Positive aspects**: Random sample (Horizon Research), sample frame matched to the demographics of Auckland population for age, gender, income, employment and education levels, sufficiently large sample size (see animation) with a small margin of error, anonymous so no interviewer effects

**Concerns:** Selection bias**:** Online-so respondents have to have access to the internet, research panel sample frame chosen by Horizon-may not totally reflect Auckland population eg where people live and ethnicity not mentioned in the demographic list, Hidden agenda: Council want support for their transport policy, Question concerns: support for the 4 big projects not questioned, options geared to produce support for tolling in some form. Anything missing?- see additional information about this survey on Horizon’s website.

**LESSON 4** – STUDENT PRACTICE –moving to less familiar contexts.

# Teaching Notes

• • •

**CONTEXTUAL SUPPORT**

Provide an opportunity for an internet search or supply a background information sheet on DVT-link to “economy class syndrome” which is probably familiar to students.

**An important teaching point:** DEFINING MEASURES

There should be no ambiguity when answering a survey. This ComRes survey clearly defines an “office worker” and “a gamer” but offers the ambiguous response options: *very often, fairly often, not very often* in questions 9 to 12.

**Resources needed**: *“*Writing frame *for critically evaluating a report”,* a suitable report e.g***. “*Young professionals and gamers risk thrombosis”***,* worry questions sheet, additional **contextual** **support needed e.g. What is DVT?**

**Note**: Lifeblood is a British charity promoting thrombosis awareness.

Students should work collaboratively to complete the writing template.

This article mentions the results of other studies on the risk factors for DVT. The Life-blood survey is about time spent seated. The students need to be able to clearly distinguish these two aspects.

There is no graphic of the survey results. They are embedded throughout the text. The students could work in a group to compile a list of the main findings.

**Positive aspects of this survey**: large enough sample, online-so anonymous encouraging people to answer truthfully, clear definitions of an office worker and a gamer.

**Concerns:** Selection method not stated-may not be random-potential for a high non-response rate-(workers too busy)-only young people surveyed but findings extrapolated to “9 million office workers” which may include other age-groups - some ambiguous response options-dubious claim made: “eating lunch at a desk could double risk of DVT”-it is staying seated that increases the risk, not place of lunch.

**OTHER RESOURCES**: The Curia Research “Broadcasting Standards Poll” in the **resource pack for Sampling Error**-hidden agenda, has a very high non-response bias, question concern: 3 different aspects together in one question.

# Teaching Notes

• • •

This topic will be assessed externally.

Students will not be able to look up unfamiliar words or obtain additional background information not supplied in the question.

They need to build up appropriate strategies for coping through experience of reading more difficult contextual reports on current events.

Students need to be encouraged to keep abreast of current events in the media and build up their own general knowledge.

.

The student worksheet in the resource pack covers both non-sampling and sampling errors.

**LESSON 5:** USING CONTEXT to MAKE MEANING

**Resources needed***:* a suitable report e.g***.***

**“Opinion divided on NZ-US exercises”**

The purpose of this activity is for teacher modeling of how to use clues from the context to make meaning of unfamiliar vocabulary.

Potential unfamiliar vocabulary in the above report:

* Exercises in the military sense and the two names “Galvanic Kiwi” and “Alam Halfa”
* ANZUS
* rift
* dubbed
* Marine Corps
* Reciprocal platoon exchange
* Resuming, resumption

Potential unfamiliar general knowledge requiring students to do some additional background research:

* Role of Americans in NZ during world war II
* The ANZUS alliance
* The origins of New Zealand’s anti-nuclear policy and the banning of American ship visits.

(Background research could be set as a homework task prior to class)**Suggested strategy**: Read the report aloud with the class, stopping at appropriate moments to ask students what they think a particular phrase or word means in the context of the report-helping them as required.

**The above report is a good lead-in to the related aspect of this standard: a conceptual understanding of margin of error. (see Workshop 1 material on Sampling Error- )**

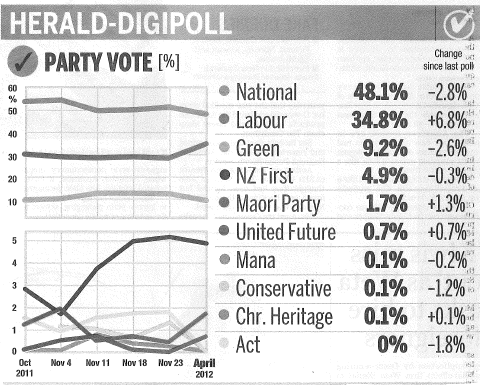
# Teaching Notes

• • •

Political polls are less context dependent than other polls and surveys.

They could come up in a non-sampling error or sampling error question or both.

**LESSON 6 : POLITICAL POLLS**

This topic was unfamiliar to students in the research trial

**Resources:** A power-point presentation on the methodology of the main New Zealand polling companies.

**Note:** Horizon Research uses an on-line panel and deals with the undecided voters very differently to the other main companies. Horizon have put forward an argument in defense of their different methodology: see the Microsoft Word document entitled **“Comparing Polls”** in the Resource Pack.

The issue of CATI in the modern era of cell-phone communication is discussed in the following article by Dr. Andrew Balemi (Auckland University) on statschat:

http://www.statschat.org.nz/2011/10/25/all-about-election-polls/

This item also discussed the margin of error and two of the rules of thumb for estimating it.

**This topic may therefore be better placed after the rules of thumb for estimating margins of error have been introduced (see Workshop 1 materials on Margin of Error).**

However, despite the debate on the pros and cons of the different methodologies, political polling companies in New Zealand generally predict results quite close to those obtained on election day as can be seen in the following resource:

<http://www.scoop.co.nz/stories/PO1111/S00761/analysis-of-political-polls-election-ratings.htm>



|  |
| --- |
| EVALUATION OF STATISTICAL REPORTS |
| Part 2 –Observational Studies and Experiments. Resource prepared by Rachel Passmore |
| **EVALUATION OF STATISTICAL REPORTS** |

Part 2 Observational studies and Experiments

# Teaching Notes

• • •

Notes about activity objectives will be made in this panel.

References to resources used such as powerpoint slides will be referenced in this panel.

Unit Overview

UNIT OVERVIEW

|  |  |  |
| --- | --- | --- |
| Lesson | Topic | Description |
| 1 | Introduction | Review & discuss some media articles |
| 2 | Experiments | What are they? Causal claims. Points to Ponder + identification in media articles. Potential Problems |
| 3 | Experiments – design issues | Design Issues - Control Group, Placebo effect. Blinding. Review articles that use these terms |
| 4 | Observational Studies | Look forward/back. Why OS conducted in preference to an Experiment? Discussion of potential problems |
| 5 | Experiment vs OS | Summarise differences between the two types of studies. Review articles and classify study type with justification. |
| 6 | Evaluation of Report | Evaluation steps, Worry questions, Work-sheet.  Begin first in depth evaluation. |
| 7 | Attack of killer frying pans | In depth evaluation of media report. Examples of possible student responses provided |
| 8 | Exam Style Questions | Examples of shorter exam style |
| Other Resources | Teen drinking may cause irreversible brain damage | Second in depth analysis of media report. |
| Other Resources | More exam style questions & review | Examples of shorter exam style questions & review of Unit as a whole. |

**LESSON 1 – Introduction** Ask students to read both articles and then have a class vote about which one they believe.

# Teaching Notes

• • •

Media Articles –

Select some articles that demonstrate a conflict. Cell phones cause cancer & Cell phones do not cause cancer.

Important to evaluate the study not how well the article was written

SLIDE 3

Focus on importance of defining terms precisely, eg what does ‘drugs’ include? Marijuana, Heroin, Nicotine, Alcohol, Caffeine

How will the effect be measured?

How did they arrive at their decision?

Was it based on evidence presented in the article or just gut feeling?

Having read both articles are they likely to modify their behavior in regard to cell phone use?

**LESSON 1 – ACTIVITY 1**

Do drugs have an adverse effect on teenage brains?

Does listening to music help students to study?

What acne treatments are the most effective?

Define terms in each statement.

DRUGS –

ADVERSE EFFECT-

MUSIC –

HELP FOR STUDY-

ACNE –

MOST EFFECTIVE-

What other information might you need to answer the questions?

**LESSON 1 – ACTIVITY 2**

**How are Media Articles Written ?**

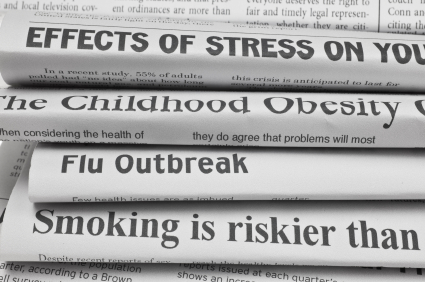
# Teaching Notes

• • •

Slide 4

Journalists are taught to write media articles in a particular style. Learning about this style enables a reader to find the relevant sections of an article easily and quickly.

Many readers skim read articles so it is important that journalist highlight the most important point of the article early on – usually either in the heading or in the first few sentences.



By using this structure their article must answer the following questions about the topic :-

Who, What, When, Where, Why and How?

Using a highlighter pen or underlining, get students to read through a newspaper or magazine article and highlight the

Who, What, When, Where, Why and How?

Having identified these ask students to write a 2 – 3 sentence summary of the article.

Share summaries with the class.

**LESSON 2**

# Teaching Notes

• • •

Slide 5

There are many different types of experimental design. Try to select media articles that allocate one treatment to each participant. Students are not required to cope with more complex designs.

Students are often confused about the difference between experiments and observational studies so spend some time making sure they have this correct before moving on.

Slide 6 Select some from slide for class or group to discuss. FOCUS - For it to be an EXPERIMENT , the treatment received should be the only difference between groups

**What is an Experiment ?**

Many studies are called experiments in media reports but in Statistics the term EXPERIMENT has a very precise meaning. It is important to recognize when research can be regarded as an EXPERIMENT in the statistical sense as only then can CAUSE and EFFECT be claimed. However, it is not always possible to conduct a true experiment for ethical reasons. Designing a perfect experiment is very difficult and often time-consuming and expensive.

Features of an EXPERIMENT

1. Participants are randomly allocated to treatments.
2. Variable of interest is measured and results compared between treatment groups.



LESSON 2 – ACTIVITY 1

Identify which of the following studies might be regarded as an EXPERIMENT using the statistical definition. Discuss as a class or with a neighbour.

The answer for most of them depends on exactly how the research was carried out. Some examples –

Study that compares IQ test results with & without caffeinated drinks. – If students were randomly allocated to a group with caffeinated drink and without caffeinated drink, then this could be an experiment. If students were asked after test who had or had not had a caffeinated drink then it would be an observational study.

Study of impact of smoking during pregnancy on infant performance ability – This is unlikely to be an experiment as researchers would not ethically be able to allocate pregnant women to a smoking or non-smoking group.

# Teaching Notes

• • •

Use media articles from resource pack or find some more current ones.

Slide 7

Effect of birth weight on adult academic performance – As researchers cannot control birth weight again this cannot be an experiment.

LESSON 2 – ACTIVITY 2

Can I identify an experiment from information in an article?

Use actual media articles. Students to read articles and discuss in pairs which are EXPERIMENTS in the statistical sense. Ask students to justify their decision to the rest of the class and summarise what their article was about.

LESSON 2 – ACTIVITY 3

Common terms

**Experiment** – An experiment measures the effect of altering the environment in some way. This may involve giving a pill, an exercise plan, a special diet, or a food supplement to those involved in the experiment. All other possible factors that might affect the environment must remain unchanged.

**Response variable** - the response to the pill or diet is measured. This might involve measurement of blood pressure (after being given a pill) or weight (after being on a special diet).

**Explanatory variables** – the variable that attempts to explain or cause (at least in part) the changes observed in the response variable. This might be a new drug or a new diet. All other variables must remain unchanged.

**Treatment –** one or a combination of explanatory variables assigned by the researcher.

**Confounding variables** – are variable(s) connected with the explanatory variable which may be the contributing to the effect on the response variable. The effect of the confounding variable on the response variable cannot be separated from the effect of the explanatory variables on the response variable.

EXAMPLE – An investigation shows that as imports of lemons to the USA have increased the number of highway fatalities has decreased. Obviously there is another confounding variable such as improved safety in cars, better driving skills, speed limits etc that are the confounding variables in this case.

# Teaching Notes

• • •

Slide 8

Slide 9

Identify terms for the studies on Slide 9.

Examples

**Effect of acupuncture & physiotherapy on back pain**

Explanatory Variable – acupuncture/physiotherapy

Response Variable – level of back pain

Treatment – physiotherapy or acupuncture

**Are defensive drivers better drivers?**

Explanatory Variable – completion of defensive driving course

Response Variable – number of accidents in a given time period

Treatment – complete defensive driving course or not

**IQ test performance and caffeinated drink consumption**

Explanatory Variable – caffeinated drink consumption

Response Variable – result of IQ test

Treatment – caffeinated drink or non-caffeinated drink

LESSON 3 - DESIGNING A GOOD EXPERIMENT

LESSON 3 ACTIVITY 1

How would you design an experiment to test whether listening to music helps students to study better?

Music – what type, how long for, with or without headphones, what if type selected is not your taste, volume controls,etc

Study Better – what measure or measures could be used

Designing a good experiment is not a straightforward task, designing a perfect experiment is nigh on impossible ! However, the following guidelines should help students to recognize when good experimental design principles have been applied in a reported research study.

# Teaching Notes

• • •

Prompt class discussion about clarity over definitions, conditions for experiment.

Remember to emphasise that the aim of an experiment is to keep all factors constant EXCEPT for the explanatory variable.

Slide 10

The following procedures may feature in a well-designed experiment. Not all of them will apply to every experimental situation.

1. Random allocation to treatment groups
2. Use of a Control Group
3. Use of a Placebo
4. Use of Blinding

Random Allocation

Random allocation to treatment groups is NOT THE SAME as the idea of random selection that you may have met when selecting samples from a population. However, the ideas are related. Randomly selected samples from a population ensure that everyone in a population has the same probability of selection. Random allocation is used in experiments in order to ensure that all experimental units ( eg. people, animals, plants) have the same probability of receiving each treatment option.

# Teaching Notes

• • •

Slide 10

Students must be clear about the difference between random allocation and random selection used in sampling methods.

Slide 10

A Control Group helps to establish a base line effect from which to compare Treatment Group effects.

Slide 10

Random allocation serves to create groups that are similar in characteristics apart from the treatment received. This allows researchers to attribute a cause to the observed effect.

Use of a Control Group

To determine whether a ‘treatment’ has had an effect or not, researchers need to establish a base line reaction ie. What would have happened to the response variable if NO treatment had been administered. To do this researchers create a CONTROL GROUP – these participants are treated in EXACTLY the same way as all other participants EXCEPT they do not receive the treatment under investigation. In the example above the Control group would be ‘No Alcohol or Drugs’.

Use of a Placebo

Sometimes the Control Group will be given a PLACEBO. For example, when testing the effectiveness of a new drug – one group will receive the new drug and the control group will receive an identical- looking pill but it will have no active ingredient, often it is just sugar.

A well- known problem of experiments is the Placebo Effect – this is when participants who have received a dummy treatment such as a sugar pill still experience beneficial effects.

Alternatively, researchers may be investigating whether a new treatment is superior to an existing treatment. In this case the control group will receive the currently prescribed medication and not just a sugar pill.

Use of Blinding

# Teaching Notes

• • •

. Watch Youtube clip about Placebo effect <http://www.youtube.com/watch?v=yfRVCaA5o18>

Slide 10

Ask students for suggestions of similar SINGLE BLIND experiments. For example, effects of meditation on anorexia, yoga on migraines etc.

Ask students for suggestions of similar DOUBLE BLIND experiments. For example, medical trials of new drugs. Although double blind is the gold standard of experimental design, remember it is not always possible to achieve. So a study should not be condemned if it has not met this standard.

Researchers who are keen to promote their new drug may be unable to eliminate bias in their measurement of the effect of the explanatory variable on the response variable. To avoid such biases, good researchers will use a blinding procedure. There are two types of blinding – single blinding and double blinding.

Single blinding – EITHER the participant knows OR the researcher knows who has received which treatment.

EXAMPLE – A study into the effects of an exercise programme on blood pressure. Participants would be aware whether they are doing exercise or not, but the research would not know who had exercised and who had not.

Double blinding – In a double blind experiment neither the researcher NOR the participant knows who has received which treatment. In this situation a statistician will randomly allocate participants to treatment groups for the researchers. Only after all measurements have been taken will they reveal which treatment group participants belonged to. Whenever possible double-blind experiments are preferable to single-blind experiments.

EXAMPLE – A study into the effects of nicotine patches and success in giving up smoking. All participants would receive patches but they would not know if they had received a dummy patch or a nicotine patch

LESSON 3 – ACTIVITY 2

# Teaching Notes

• • •

This is an adaptation of an actual report. Often carefully selected extracts will be sufficient for students to make an informed evaluation.

Average cholesterol level in NZ is 5.7 mmol/L. Suggested level is less than 4 & dangerous levels are higher than 6.5. A drop from dangerous ( 6.5) to safe ( 3.8) requires a 42% drop, so a 35% drop could mean many people could reduce to the target level of cholesterol.

Evaluate article below and identify which features of experimental design have or have not been used.

*Extract from newspaper*

*“ Rigorous Veggie Diet Reduces Cholesterol”*



*People with high cholesterol may lower their levels by a surprising one-third with a vegetarian diet that includes a variety of heart-healthy foods a study finds.*

*“The reductions are surprising” said University of Auckland specialist in nutrition.*

*In the experiment, 25 volunteers ate one of the following diets*

*Diet One – foods thought to reduce cholesterol were included in a low-fat diet but participants also received a new cholesterol reducing drug in the form of a pill*

*Diet Two – as Diet One but participants also received a placebo pill*

*Participants were randomly allocated to diets by a statistician using a random number generator. The statistician held the code for which participant had received which of the above treatment options.*

*The laboratory staff who were testing cholesterol levels had no knowledge of who had received which pill. Dieticians had no knowledge of who had received which pill.*

*After a month, cholesterol levels fell on average by 12% for those on Diet Two, and by 35% for those on Diet One.*

# Teaching Notes

• • •

1. YES, allocation to treatments was random.
2. Diet Two represents the Control Group
3. Yes a placebo pill was used
4. Yes Double Blinding was used. Neither participants nor researchers knew who had been allocated to which pill

Remember a causal claim, if based on an observational study, is not justified.

Features of Experimental Design to Consider

1. Allocation of treatments, was this done randomly?
2. Was a Control Group used? If so define.
3. Was a Placebo treatment used? If so define.
4. Was blinding used ? If so which type?

Do you think this study follows good experimental design principles? Explain why it does or does not.

Possible answer

This study does follow good experimental design principles. It has randomly allocated participants to treatment groups. It has used a double blinding method and utilised a Control Group to establish a baseline effect. This Control Group received a placebo pill.

LESSON 3 ACTIVITY 3

CAUSAL CLAIMS

Newspaper headings frequently claim a causal link based on study results. UNLESS the study was conducted along the lines of the statistical definition of an EXPERIMENT, a claim of causality cannot be justified.

**ONLY EXPERIMENTS CAN CLAIM A CAUSAL LINK**

# Teaching Notes

• • •

In this task students must question first of all if a causal claim is being made. If a causal claim is made was the research based on an experiment?

Familiarise students with the sorts of emotive words used in headlines .

Causal words –

Cause,effect,result,eliminate,prevent,kills

Non-causal words –

Associated, related,

Reference

Schield, M. & Raymond, R. (2009) “Distinguishing Association from Causation in Media Headlines”, JSM

Examples of typical headline claims

1. “Drinking eight glasses of water a day prolongs life by 10 years”
2. “Marijuana use by teenagers causes brain damage”
3. “Cell phone usage linked to brain tumours”

Which of the above were causal claim headlines ?

1 & 2 are making causal claims, 3 just claims a link

Were findings based on an experiment or not?

1. – could have been an experiment but would need further details to confirm.

2– could not have been an experiment as unethical to allocate teenagers to a marijuana use group

3- most likely to have been an observational study

Could the headline be justified? If not can you suggest an alternative headline?

1. – might be justified as study could have been an experiment
2. – headline unjustified as study was not an experiment.

Alternative headline “ Link found between marijuana use by teenagers and brain damage”

1. – Headline OK as no causal claim made

If time get students to make up their own headlines and the class can decide on whether claims are causal or not.

LESSON 4 – OBSERVATIONAL STUDIES

# Teaching Notes

• • •

Ask students for other circumstances when an experiment might be unethical.

Slide 11

Lots of reports have come out of this study. Students could find some examples as a homework task.

Students must be familiar with these three potential problems & be able to discuss them in context. Each problem is described below in context of an article.

Sometimes it is not possible to conduct an experiment for ethical reasons. For example, an investigation of smoking or alcohol in pregnancy on babies. It would be unethical to allocate a treatment of smoking or a level of alcohol consumption to a pregnant women. In such situations observational studies are used to observe differences in the explanatory variable. Any observed differences are then analysed to see whether they relate to changes in the response variable.

LESSON 4 – ACTIVITY 1 – Think of some research questions that could not be answered by an experiment because of ethical reasons.

Observational Studies sometimes recruit a group of participants and record their details over a period of time. There is a famous NZ Observational Study which has been assessing a group of babies born in Dunedin in 1972/73. There are over a thousand participants who have all been monitored regularly since 1972. It is called the Dunedin Multidisciplinary Health and Development Study (<http://dunedinstudy.otago.ac.nz/>).

Observational Studies may also consider past behaviour of a group of participants either by questioning them directly about past events or by looking a historic records such as medical records.

Potential Problems with Observational Studies

Confounding variables

Extending results inappropriately

Using the past as a source of data

Confounding Variables

Confounding variables and implications for causation – In an observational study it is impossible to separate out all potential confounding factors as random allocation of treatments has not been used. **Thus a causal link cannot usually be claimed from an observational study**.

LESSON 4 – ACTIVITY 2

Identify confounding variables identified in article?

What other confounding variables can you think of?



**“HAPPY PEOPLE LIVE LONGER’**

*Heart patients who are happy are much more likely to be alive 10 years down the track than unhappy heart patients according to a* *study by a prestigious Medical Centre. The experience of joy appears to be a key factor. Unhappy people apart from being in a bad mood are less likely to take their medicine, eat healthily or exercise on a regular basis”*

# Teaching Notes

• • •

Slide 12

Important fact for students to recognize.

Confounding variables

Unhappy patients were less likely to take pills, eat well or exercise. These are all possible confounding variables.

Others could be age, social network, marital status, climate, living conditions, socio-economic status, pets etc.

It has to be an observational study as you cannot tell someone they MUST be happy!

The headline generalises to ALL happy people live longer, whereas the study was restricted to heart patients only.

LESSON 4 – ACTIVITY 3

Extending Results inappropriately

Can you identify and explain why extending results inappropriately might be a problem in the report below?

# Teaching Notes

• • •

Slide 13

Research was based on a sample of American undergraduate students at one University. The headline suggests that findings can be extended to all women but in fact the research findings are only applicable to the particular group who were investigated or other similar groups – such as undergraduates at other Universities, results certainly cannot be extended to ALL WOMEN.

Other problems – no indication is given about how the sample was selected. Was it representative?

“USA Female Undergrads at X Uni suffer more hangovers than men”

Can you think of any other potential problems?

Many observational studies use convenience samples which cannot be regarded as representative of a wider population. Results of a study can ONLY be extended to a larger population if the sample is representative of that larger population.

“ WOMEN HIT HARDER BY HANGOVERS”



“ *New research finds that women not only get drunk on fewer drinks than men but also suffer worse hangovers. Researchers asked 1230 drinking undergraduate students from an American University to describe how often they experienced certain symptoms after drinking. 26% of students reported experiencing hangovers at least once a month and researchers suggest that these students are at a higher risk of failure.*

What would be a more appropriate headline?

LESSON 4 – ACTIVITY 4

# Teaching Notes

• • •

Students do not need to learn the terms retrospective/ prospective observational study but they do need to be aware of the potential problems associated with each type of study.

Using the past as a source of data

Ask students to recall what they had for breakfast/lunch/dinner on each day last week?

Any study that uses the past as a source of data is known as a **retrospective** study. Such studies rely on people’s ability to recall information about past events accurately – which many people find hard to do! One option to overcome this is to use medical records, if appropriate, as a data source. An additional problem is that over time the influence of confounding variables may change. Consequently, **prospective** observational studies are preferred. Another problem with retrospective data is that variables that might have confounded things in the past may no longer be similar to current confounding variables.

The Dunedin longitudinal study records data prospectively and retrospectively.

Prospective studies – examine cardio , dental, respiratory, sexual & mental health by interviewing participants at regular intervals

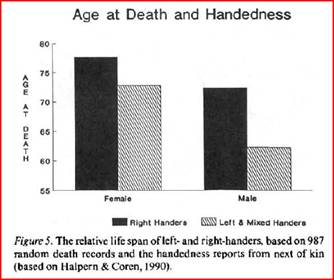
Retrospective studies – have examined teenage involvement in road crashes, link between birth weight, smoking & incidence of fractures.

Over 1100 publications have arisen from this study. To access them

http://dunedinstudy.otago.ac.nz/publications

Can you identify any potential problems with using the past as a source of data in the report below?

“ LEFT- HANDED PEOPLE DIE YOUNGER”



*The researchers, Halpern & Coren, sent letters to the next of kin of recently deceased individuals, and asked them whether the deceased had been right or left-handed. Their results illustrated above appear fairly conclusive that right-handed males and females lived longer than left-handers.*

# Teaching Notes

Slide 14

Shaughnessy stated that “The distinction between Experiments & Observational Studies constitute the single most important contribution that statistics has made to the progress of science”

An Observational Study often precedes an Experiment in a research cycle.

What the researchers failed to remember was that in the early part of the twentieth century, many children were forced to write with their right hand even if naturally they were left-handed. Therefore, people who died in their 70s or 80s were more likely to be right handed. The confounding variable is how long ago participants learned to write.

LESSON 5 – Comparison of experimental and observational studies

# Teaching Notes

• • •

Slide 15

Main features of experimental and observational studies are compared on Slide 15.

|  |  |
| --- | --- |
| **Experiments** | **Observational Studies** |
| Explanatory variable specified | Explanatory variable specified |
| Response variable specified | Response variable specified |
| Participants randomly allocated to treatments | Participants NOT randomly allocated to treatments, may be unethical |
| Treatment is the only factor that varies |  |
| CAUSAL CLAIM CAN BE MADE | CAUSAL CLAIM CANNOT BE MADE |
| Changes in response variable recorded | Changes in response variable recorded |
| Good experiments include random allocation to treatments, control groups, placebos and use blinding | Good observational studies acknowledge & account for all potential problems |

The most important distinction between the two types of study comes from the inferences they support. An Observational Study might reveal that on average women live longer than men, but cannot reveal why. Often the relationship found by an Observational Study precedes an experiment.

LESSON 5 – ACTIVITY 1

# Teaching Notes

• • •

Slide 15

Detecting Experiments and Observational Studies in the Media

Using media reports from resource pack or ones students collect themselves.

1. Read & summarise article findings
2. Decide which type of study has been conducted & why?

Sources of media articles about experiments & observational studies

1. On-line newspaper & magazine websites
2. <http://www.statschat.org.nz/>
3. <http://www.statlit.org/>
4. <http://www.nhs.uk/News/Pages/NewsArticles.aspx>

# Teaching Notes

• • •

Sources of media articles. The key to engaging students is context. So encourage students to source their own articles if possible. Keep a folder of the betters ones and you will soon build up a good resource bank. From online sites a pdf version can usually be saved for future reference as links can disappear.

Stats Chat website will often provide links to original research paper as well as media article link.

Slide 16

1. [www.sciencemediacentre.co.nz](http://www.sciencemediacentre.co.nz)
2. [www.stats.org](http://www.stats.org)
3. [www.wordpress.com](http://www.wordpress.com)

This lesson could be done as an assignment or as a research activity in the library or computer lab.

LESSON 6 – Evaluation of Statistical Reports

Statistical reports include media articles which refer to research findings as well as press releases by organisations promoting publication of research findings. Whenever possible students should be encouraged to find the original research paper on which an article or press release was based before evaluating the report. However, in an assessment a student will not have to read a complete report but will have to evaluate selected extracts from media reports, press releases and research papers.

This teaching unit will demonstrate in-depth evaluations of media reports and will then give examples of assessment style questions.

Assessment style questions will usually contain only an extract from an article and will focus on just a few of the evaluation steps.

IN-DEPTH EVALUATION OF STATISTICAL REPORTS

# Teaching Notes

• • •

Worry question page in Resource Pack. Often useful to print this in a different colour, then can say “Use the yellow page to evaluate your report”

Slide 16 covers steps 1 – 5

Reputable journals include The Lancet, BMJ or any other peer-reviewed journal.

Context is crucial in engaging students. Start by evaluating articles about familiar contexts and move towards more unfamiliar contexts. Unfamiliar contexts will require some research prior to evaluation.

Evaluation of a statistical report will involve students answering a number of questions. These are presented in the Resource Pack under ‘Worry Questions’. Students can use these as a template to get started on evaluation but they will not be available in an assessment.

LESSON 6 – Activity 1 – How to evaluate a report

Evaluation Steps

Step One Is the report based on sound research? Find the original research article to determine this. Was it published in a reputable Journal? Was it conducted by reputable researchers from a University or Research Institute for example?

Step Two Identify the type of study undertaken – experiment or observational study?

Step Three Who funded the research? Did they have a vested interest in the outcome? For example, was it a drug company trying to compare their new drug over an existing treatment?

Step Four Were reliable data collection methods used? If data was collected from the past, how was accuracy checked? Could confounding variables have changed over time?

Step Five Who or what was studied? How were they selected?

Step Six What measurements were taken and how? Are questions defined and responses provided? Are the measurements appropriate for the effect that is being observed? For example, is an IQ test an acceptable measure of intelligence. How were measurements taken – interview, online, questionnaire. Could this have had an impact on responses?

Step Seven Where and when were measurements taken or questions asked? Sometimes if measurements are taken in an unfamiliar setting for the participant, atypical responses can be observed. Some measurements will vary according to the time they are taken – morning/evening, summer/winter, weekend/weekday.

Step Eight Is the size of the observed effect quantified or just described? Is a claim made that Drug A reduces blood pressure OR that Drug A reduces blood pressure by 15%. Beware of unquantified claims.

# Teaching Notes

• • •

Slide 17 covers steps 6 – 12

Slide 18 cover steps 13 & 14

Worry Questions & Report Evaluation template

Step Nine Are there any confounding variables that should be considered?

Step Ten Have the results been extended inappropriately?

Step Eleven Is the information complete? Was non-response or missing data a problem?

Step Twelve Do the results make sense? Is there a scientific or biological reason to support the evidence? Is there any alternative explanation.

Step Thirteen Would the results persuade you or someone you know to modify or change their behaviour?

Step Fourteen Overall conclusion about the research findings. Justify why the study was good or why it was not. Do other studies reveal similar findings? What further information about the study might be useful.

Discuss Worry Questions and Report Evaluation template and demonstrate relationship between the two and the steps identified above.

LESSON 6 – ACTIVITY 2

Read the following NZ Herald article

**Heart disease linked to non-stick cookware: study**

**1:51 PM** Wednesday Sep 5, 2012



**Perfluorooctanoic acid found in non-stick products like lubricants, food packaging and non-stick cookware has been linked to heart disease researchers have found. Photo / Thinkstock**

# Teaching Notes

• • •

Electronic/Hard Copy of NZ Herald article available in Resource Pack

Students read article using their Worry Questions. They can highlight important passages and label them according to which evaluation step they refer to.

The evaluation steps are designed as a scaffold for students beginning to analyse statistical reports. Later on they will need to write reports without a list of worry questions or evaluation template.

US researchers have identified a link between chemicals found in non-stick cookware and heart disease.

But Australian experts have urged caution about the findings, which suggested increasing levels of perfluorooctanoic acid (PFOA) in the blood were associated with cardiovascular disease.

PFOAs are found in products including lubricants, polishes, food packaging and non-stick cookware.

But people can also be exposed to the chemicals in drinking water and the air, a study published in the Archives of Internal Medicine said.

The study reviewed the levels of the chemical in 1216 people with heart problems.

The findings were independent of other factors including age, sex, smoking, weight, and diabetes.

The authors from the West Virginia University School of Public Health said the results contributed to the data emerging on the health effects of perfluoroalkyl chemicals.

But the study could not conclude that the chemicals actually caused heart problems, the authors said.

Chair of Water Quality Research Australia Professor Michael Moore said ubiquitous exposure to PFOA, although at low levels, in non-stick cookware, could cause considerable anxiety.

He said the chemical was also found in coatings on clothing and carpets and in plumbing tape.

Moore said although the authors had tried to remove other lifestyle factors from the equation, there was still a possibility the PFOA levels were caused by other exposures.

"The key point is that there are numerous established contenders for cardiovascular disease risk factors - smoking, lifestyle, exercise, diet - all of which may contribute more to cardiovascular disease than PFOA exposure," Moore said.

# Teaching Notes

• • •

See commentary on this article by Thomas Lumley on Stats Chat website.

http://www.statschat.org.nz/?s=Killer+Frying+pans

Important to cover any vocabulary or context issues after first read through of article and before students attempt formal evaluation.

Monash University Professor Brian Priestly said the results of other studies investigating the link between PFOA and heart disease had been inconsistent.

University of Adelaide pharmacology lecturer Dr Ian Musgrave said single blood measurements were taken in a short period of time and the long-term exposure to PFOA may be different from the results taken from one measurement.

"We know, for example, that blood levels for PFOA are falling over time due to regulation of this chemical," he said.

In the same edition of the journal, Debabrata Mukherjee of Texas Tech University Health Sciences said given the concerns raised by the study, clinicians needed to act now.

"It would make sense to limit or to eliminate the use of PFOA and its congeners in industry through legislation and regulation while improving water purification and treatment techniques to try and remove this potentially toxic chemical from our water supply," Mukherjee said.

- AAP

Copyright ©2012, APN Holdings NZ Limited

Discuss any immediate issues as a class first.

Use worry questions and evaluation template to critically evaluate this report. This could be completed as a homework task.

LESSON 7 – Evaluation of Statistical Report Exemplar

# Teaching Notes

• • •

These responses are not the only possible responses. Encourage discussion of other valid responses students may have made.

Possible responses to evaluation steps are given below. Discuss student responses and compare with those below.

**Step One –** A report on this study appeared in the NZ Herald. The study was conducted by the West Virginia University of Public Health. This organisation can be regarded as reputable. The full research article is published online in the Archives of Internal Medicine, which is a peer-reviewed Journal.

**Step Two** – This study was reported as an observational study.

**Step Three** - No details of specific funding for the study is provided, so it is assumed that the School is an independent body with no vested interests in the results of the study. If the study had been funded by a manufacturer of cookware suspicion might have been warranted.

**Step Four** – Confounding variables are unlikely to have changed. The study used medical records from the NHANES survey of Americans. Accuracy of recall is often a problem with observational studies that use data from the past, but here researchers have used medical records which can be assumed to be reasonably reliable and accurate.

**Step Five** - The article simply states that 1216 people with heart problems were studied, but to discover where these people came from we have to look to the original study. The participants were selected from the National Health and Nutrition Examination Survey (NHANES). This survey collects data from a random sample of the US population. Only some of the survey participants had their blood tested for PFOA levels and probably the participants self-selected themselves as a result. Data from 1999-2000 and 2003-2004 were considered, but no reasons are offered as to why these particular time periods were selected.

The participants were all said to suffer from heart disease but interestingly this was not measured by a doctor but simply as the response to a question “ Has a doctor ever told you that you had a coronary heart disease?” If a participant answers in the affirmative no information is provided about when they suffered from the disease or indeed whether the PFOA measurements taken were taken at a time when they had heart disease or not. Dr Ian Musgrave notes that a single measurement of PFOA levels taken in a relatively short period may not adequately reflect the impact of long-term exposure to PFOA.

# Teaching Notes

• • •

Possible responses provided only – hopefully students will come up with lots of other valid points.

An alternative in-depth evaluation is provided later in the unit. It evaluates a report on teenage drinking and brain damage. Choose whichever context is more likely to engage your students.

**Step Six –** Single blood measurements were taken over a short period of time. The level of PFOA in the blood was assessed and this was a good indicator of the effect researchers were trying to measure.

**Step Seven** – Measurements are a problem with this study. For example, participants were classified as having heart disease if the responded positively to the question “ Has a doctor ever told you that you had coronary heart disease”. No examination by a doctor was performed. The actual measurement on PFOA levels is unlikely to have been affected by where the blood test was actually taken but may have been affected by time of blood test.

**Step Eight –** The newspaper article does not mention the size of the effect, it just states that a link has been found between chemicals present in non-stick cookware and heart disease.

**STEP Nine** - Confounding Variables – the authors of the study did try to control for some confounding variables such as age, sex, smoking, weight and diabetes, but an Australian professor, Michael Moore, notes that even so there was still a possibility that the PFOA levels were caused by other exposures.

**STEP Ten** - Extending results inappropriately – the authors acknowledged that the results of their study do not allow a causal link to be made between heart disease and exposure to chemicals. No inappropriate extension of results was found. The researchers also acknowledge that a prospective longitudinal study is required in order to make definitive claims.

**Step Eleven** - The data was complete for the individuals and time period considered. However, it is odd that since NHANES is an annual survey that the researchers chose such a limited time period for their study.

**Step Twelve** - Many studies have identified numerous factors that may contribute towards heart disease. Exposure to PFOA may be another such factor, but this study fails to determine whether such exposure is a major or minor contributor when other factors, such as lifestyle, are taken into account.

**Step Thirteen** – I don’t think this report will stop people buying non-stick cookware.

**Step Fourteen**– The article starts by saying that US researchers have found a link between exposure to PFOA and heart disease. A ‘link’ however, could be a positive or negative link in terms of increased or decreased risk of heart disease. The commentary which follows this statement, though, suggests that the link showed that higher levels of PFOA were associated with increased risk of heart disease. The article does correctly point out that no causal link between the chemicals and heart disease can be claimed and commentary on the research article suggests that more research is required.

I do not think that people will stop buying non-stick cookware as a result of this article. The article does present several points of view but I think it needed to devote more attention to the actual results of the observational study.

LESSON 8 – ASSESSMENT STYLE QUESTIONS

# Teaching Notes

• • •

Use Worry Questions and evaluation template if required.

Evaluation of statistical reports on experiments or observational studies will only be ONE Question in the assessment. An in-depth analysis will not be possible in the time available. Students will need to read an EXTRACT of an article and will have to answer questions about the article concerning some but not all of evaluation steps used for an in-depth analysis.

ASSESSMENT QUESTION ONE

*Extract from Newspaper article*

“ HEAVY DRINKING DURING PREGNANCY LEADS TO HYPERACTIVE KIDS”

A recent study of 11,500 UK children by the Department of Epidemiology and Public Health, showed that the children of mothers who drank heavily during pregnancy were more likely to suffer from hyperactivity and behavioural or emotional problems.



*Extract from Journal Article*

*This study examines the relationship between light drinking during pregnancy and the risk of emotional or cognitive problems at age 5 years. 11513 UK children were selected from the UK Millennium* Cohort study. *Participants were grouped according to the mothers’ reported alcohol consumption during pregnancy. Children of these mothers were tested at age 5 years. Children born to mothers who drank 1 – 2 drinks per week or per occasion during pregnancy were not at increased risk of behavioural or cognitive difficulties compared with mothers who had not drunk during pregnancy. However, children born to mothers in the heavy/binge drinking category were more likely to have hyperactivity, behavioural or emotional problems compared to those born to mothers who did not drink during pregnancy.*

# Teaching Notes

• • •

Copy of article and questions in resource pack.

1. This research is based on an observational study. Discuss the main differences between observational studies and experiments, and why an experiment was not appropriate in this case.

Possible Answer

In an experiment participants must be randomly allocated to treatments. In this study it was not ethical to allocate pregnant women to an alcohol consumption group so an observational study had to be used.

1. A potential problem with observational studies is using the past as a source of data. Explain why using the past as a source of data might be a problem in this study.

Possible Answer

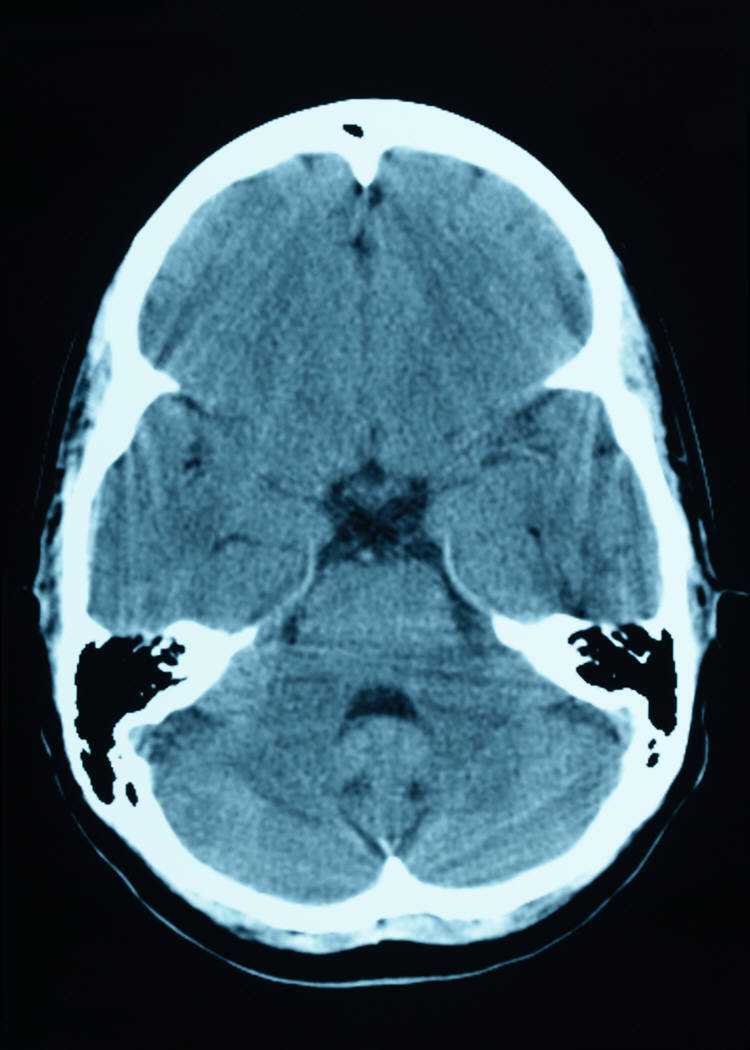
Using the past as a source of data could cause problems in this study since the method relies on mothers recalling accurately their alcohol consumption from some time ago ( several years). It is highly likely that there may be problems with their recall ability – some are likely to under-estimate how much they drank, some are likely to over-estimate how much they drank. Mothers may also not have drunk alcohol in a consistent pattern throughout their whole pregnancy – some may have been heavy drinkers in early pregnancy because they did not realise they were pregnant but cut down to no alcohol or just light drinking once they realised they were pregnant.

Write an alternative headline .

“ POSSIBLE LINK BETWEEN DRINKING DURING PREGNANCY and HYPERACTIVITY FOUND”

OTHER RESOURCES

1. Another example of an in-depth evaluation
2. Further examples of assessment style questions.
3. INDEPTH EVALUATION



TEEN DRINKING MAY CAUSE IRREVERSIBLE BRAIN DAMAGE

(2010)

Article Link <http://www.npr.org/templates/story/story.php?storyId=122765890>

**Step One –** The media report was posted on the National Public Radio website – the US equivalent of NewstalkZB. The research article was published in a reputable peer-reviewed journal Psychology of Addictive Behaviors.

**Step Two** - this is an observational study as researchers could not ethically require teenagers to drink certain amounts of alcohol.

**Step Three** – the study was carried out by neuroscientist Susan Tapert from the University of California. The article does not specifically say who funded the research but I think it can be assumed that Susan Tapert had no vested interest in the outcome.

**Step Four –** Data was not collected from the past. Participants were recruited before they began to drink and then monitored over time about their drinking habits and their performance in thinking and memory tests. The research paper provides a lot of detail about the measures taken so I believe they were taken appropriately.

**Step Five –** The media article does not give details about how participants were recruited. The research article describes how flyers were sent out to households with children attending local middle schools. Participants were excluded if they or their parents had histories of prenatal alcohol, illicit drug use or mental illness. Surprisingly only 13% of respondents met eligibility criteria.

**Step Six** – The measurements taken were brain scans plus scores from thinking and memory tests. The brain scans revealed physical differences between the drinkers’ brains and the non-drinkers brains so were a good indicator of brain damage. The tests administered revealed different consequences of that brain damage for males and females. The tests are not provided but perhaps they contained some gender bias that might have contributed towards these different consequences.

**Step Seven** – The brain scans were most likely conducted at a hospital and as physical attributes of the brain were examined the place of measurement was unlikely to affect scan outcomes. The article does not say where the tests were administered but if they were administered at the hospital this could have affected outcomes as participants would not be used to sitting tests in that environment.

**STEP Eight** -The size of the difference is reported as being 10 per cent between performance of those that do and do not drink heavily in tests of attention span. However it is not clear exactly how this 10 per cent difference is derived or to what it refers specifically.

**Step Nine**  – Researchers did attempt to account for possible confounding variables. For example, genders were examined separately and measures of economic status, home stability, home environment and behaviour level were collected.

**Step Ten** – The headline suggests that all teen drinking my cause irreversible brain damage. In fact this was an observational study so no such claim can be made. In fact the use of ‘irreversible’ is also misleading as “ What remains unknown, says Tapert, is if the cognitive downward slide in teenage binge drinkers is reversible”. The claim is only valid for the particular group of teenage students involved in the study and cannot be generalised to all teenagers without further research.

**Step Eleven** – the research article states that only 13% of respondents to their invitation to join the research programme satisfied their eligibility criteria. However, no information is provided about non-response vs non-eligibility. 13% seems quite a small proportion of respondents so there is a possibility that they possessed some characteristics that were different to those of the non-respondents.

**Step Twelve** – there has been considerable research into this area and but I am not aware of a specific biological reason why too much alcohol affects the white matter in teenage brains.

**Step Thirteen** – the article states that even teenagers who only participated in binge drinking once or twice a month abnormal brain function was observed. Insufficient white matter in the brain apparently indicates poor, inefficient communication between brain cells, resulting in lower levels of intellectual ability. I think teenagers will probably not give up drinking completely after reading this article but it might make some think more carefully about how much they are drinking and how often.

**Step Fourteen** – I think this article explained the effects of teenage binge drinking quite well, but would have liked to see more details about the methodology of the experiment in the media article. For example who were the participants (other than teenagers) and exactly how were they selected. The most telling statement for me was that the researcher was surprised by the size of the difference given that the binge-drinking episodes were not particularly frequent, perhaps only once or twice a month.

FURTHER EXAMPLES OF ASSESSMENT STYLE QUESTIONS

QUESTION TWO

*Extract from Newspaper article*

“GRADUATING IS GOOD FOR YOUR HEALTH”



College graduates feel better emotionally and physically than their high-school drop-out counterparts, a study of over 400,000 US adults reveals. Participants were asked a series of questions including level of education and on how many of the past 30 days they felt physically and emotionally healthy.

1. Define the explanatory and response variables

Explanatory Variable – level of education, Response variable – physical & emotional health

1. Explain whether any of the potential problems that observational studies can suffer from such as using the past as a source of data, extending results inappropriately or confounding variables, might apply to this study.

Using the past as a source of data – participants were asked to recall on how many of the days in the past month had they felt physically and emotionally healthy. Not many people would be accurately be able to recall how they felt on specific days even a week ago.

Extending results inappropriately – the study observed a number of US adults yet the headline extends these results to all graduates.

Confounding variables – the reason for feeling physically and emotionally healthy may be more to do with the economic status of the graduates rather than their level of education. There are lots of other possible confounding variables.

1. Evaluation steps 5, 6 & 7 are concerned with who was studied and how measurements were taken. Discuss these steps in relation to this article.

No information is provided on who these participants were other than that they were American adults. It is not clear whether they were a randomly selected sample of adults or not. If they were not randomly selected then this may have affected results. The physical and emotional health of the adults was assessed by asking participants a series of questions about how they felt over the past 30 days. There are several issues with this. First participants recall of information over the past 30 days is unlikely to be particularly accurate. Secondly, as the questions asked are not provided we cannot assess whether they provided a good indicator of physical & emotional health or not. Thirdly, it is unclear how and when these questions were administered, variation of both of these factors could have impacted on results. For example, if questions were asked in the morning or the evening, or when someone was tired or hungry. If the responses were obtained from an interview – interviewer bias may have been an issue.

QUESTION THREE

Suppose you want to know whether male or female students spend more money on clothes. You consider two different plans for carrying out an observational study:

Plan 1 – Ask the participants how much they spent on clothes during the last 3 months; then compare male and female spending.

Plan 2 – Ask the participants to keep a diary in which they record their clothing expenditure for the next three months; then compare male and female spending.

1. Discuss one disadvantage of each of these plans.
2. Discuss reasons why such a study is more likely to be an observational study rather than an experiment.
3. Plan 1 relies on participants accurately recalling information about spending during the last three months. Such data is unlikely to be very reliable.

Plan 2 relies on participants keeping a diary of expenditure. This should provide more accurate data than Plan 1 but participants could forget to make diary entries.

However, in both plans participants who spend a lot on clothes may be embarrassed about the level of their expenditure and omit some items. Others might be embarrassed about how little they spend on clothes and add some fictitious items of expenditure.

Examination of credit card and bank statements would be a more accurate source of information, but cash sales might still be omitted.

1. The response variable in this study is amount of money spent on clothes, the explanatory variable is gender. This cannot be an experiment therefore because gender is not a treatment option that can be randomly allocated. Other good experiment design principles such as use of a Control Group are also not possible because humans are either male or female , there is no neutral gender.

QUESTION FOUR

An observational study finds that people who travel to work by bus or train have a better knowledge of current affairs than those who drive to work. The link is weaker for those who have a lower level of education. What term applies to each of the following variables:

1. Method of getting to work – Answer Explanatory Variable
2. Knowledge of current affairs – Answer Response Variable
3. Level of education – Answer Confounding variable
4. Whether participant reads a newspaper – Answer Confounding Variable

Definitions of the variables given above is not necessarily straightforward. Discuss some of the possible problems and propose your own definition of each term.

‘Method of getting to work’ –Participants are categorised as using bus/train or driving. However, to be thus categorized would participants have to use that mode of transport every day or just more often than not. I would define method of getting to work – as method used for majority of travel to work.

‘Knowledge of current affairs’ is not straightforward variable to define. For example, what determines ‘current’ – events that occurred in the last week, month, year ? And ‘affairs’ would this cover all topics –national/international politics, law, education, military, sport, arts or just some of these? I would define knowledge of current affairs as performance in a test of events occurring in the last month in participants’ home country. Any topic that appeared in local/national papers or on TV/radio news could be considered for inclusion in the test.

‘Level of Education’ – this is probably more straightforward to define and could be categorised in terms of number of years at school or highest level of academic qualification. Either definition could be suitable.

‘Reads a newspaper’ – again we need to be sure of the definition of ‘newspaper’ and the frequency associated with ‘reads’. Newspapers are obviously large papers such as NZ Herald, Dominion Post, Otago Daily Times, The Press etc but there are many local papers and some people would classify magazines as newspapers too. Should we count online news sites as equivalent to newspapers or should they be excluded ? To classify someone as a person who reads a newspaper we need to define a level of frequency. Do we mean someone who reads the paper every day , every week or just occasionally. I would define this as someone who reads a newspaper ( on-line or paper) at least five times a week.

‘

QUESTION FIVE

*Extract from Newspaper article*

“ DAILY DRINK IS GOOD FOR YOU”



One drink a day can be good for health, scientists report in a new study into the drinking habits of adults. The study recorded the drinking habits of 490,000 adults and noted death rates over a period of 9 years. The article notes that most participants were “ white, middle-class, married and tertiary educated”.

1. Explain why this study could not have been an experiment.

This must be an observational study as it would not be ethical to insist participants have a certain level of alcohol consumption.

1. Do you think the headline is justified by the study? Discuss reasons

The headline does not mention WHO is likely to lead a longer life through drinking nor does it specify the level of alcohol that must be drunk to achieve this effect. The headline is claiming a Cause and Effect but this cannot be justified as an experiment has not been conducted.

1. Explain whether any other the potential problems that observational studies can suffer from might apply to this study.

The participants in this study were white, middle-class, married and tertiary educated. These factors may have more significance on their average life span rather than the level of alcohol consumed. These factors can be considered as possible confounding variables.

1. Describe in detail how you would conduct an observational study in NZ into level of drinking alcohol and life expectancy.

I would randomly select participants from the electoral roll ensuring that I have a sample that reflects the demographics of NZ adults.

I would interview participants at regular intervals over time – probably at least ten years. At the end of ten years I would examine the drinking levels of participants and compare the life expectancy of those that drank and those that did not.

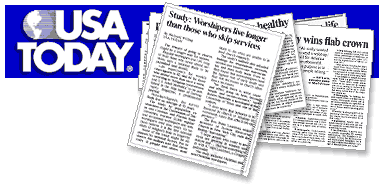
HINT The link below gives a method of calculating the life expectancy of an individual.

http://gosset.wharton.upenn.edu/mortality/perl/CalcForm.html

QUESTION SIX

**Study: Worshipers live longer than those   
who skip services**

**By Richard Willing, USA TODAY** ( 1999)

The reward of going to church might be a longer wait for heaven.

Regular worshipers live 10% longer than those who never attend services, says a national study to be published next month.

Life expectancy for weekly churchgoers is 82, and 83 for those who attend more than once a week, according to "Religious Involvement and U.S. Adult Mortality," to be published in the May issue of Demography magazine, a professional journal.

Nonchurchgoers, the survey finds, live an average of 75 years.

The study, by population researchers at the Universities of Texas and Colorado and Florida State University, tracked 22,000 regular and occasional worshipers and non-worshipers over nine years. Its authors believe it is the first national study of religion's links to longevity.

The study's authors also adjusted findings for respondents' incomes, alcohol and tobacco use; marital status; and even body mass. Even so, they found that non-worshipers have the highest risk of death and that risk decreases as church attendance increases.

"We think there is some cause and effect going on here, beyond health and socioeconomics," says Robert Hummer, a sociologist at the University of Texas and one of the authors. "It does seem that behaviour is influenced by church or religious involvement, (and) that affects life expectancy."

The study included Christian and non-Christian worshipers.

1. Define the explanatory and response variables in this study.

Explanatory variable –frequency of church attendance, Response variable – life expectancy

1. Do you think the headline is justified by the study? Discuss reasons.

I do not think that the headline is justified. This was an observational study as researchers could not force participants to go to church so a causal claim such as worshipping prolongs life cannot be made. I think there are other confounding variables, other than those accounted for by the researchers, that are contributing to this effect. Possible confounding variables might include life-style factors such as occupation, location of residence, social network, and general physical & emotional well-being.

1. How do you think the response variable was measured?

The response variable was life expectancy not age at death. Age at death was probably not used because depending on the age group of the participants there may not have been many deaths within a nine year period. Life expectancy can be calculated using a formula similar to that used by life insurance agencies to determine their premium levels.

1. Evaluation Step One asks if the findings are based on sound published research. Is this the case for the article above ?

The research was conducted by Robert Hummer a sociologist at the University of Texas, so can be considered to have been carried out by a reputable organisation. The findings were published in a peer reviewed journal Demography.

1. Evaluation Step Eight asks if the size of the observed effect is quantified? Is the effect size commented on in this article and is it quantified?

The article quotes three life expectancies 82 years for weekly churchgoers, 83 years for those who attend church more than once a week compared to only 75 years for non- churchgoers. The difference between once a week and more than once a week is very small but the difference between churchgoers and non- churchgoers is quite substantial – 7 or 8 years longer life expectancy.

1. Evaluation Step Thirteen asks whether the results would persuade you or someone else to change their behavior. Do these findings encourage you to change your church attendance habits ?

These results show that there is little difference in life expectancy between those that attend church weekly and those that attend church more than once a week. The findings definitely do not persuade me to attend church more than once a week. The difference in life expectancy between attending church and not attending church is greater but I believe that the confounding variables listed above are probably more important in extending life expectancy so I will not be changing my church attendance habits in the light of these findings.