08/29/2013

PART I – EXPERIMENTAL DESIGN
CHAPTER 1 — CONTROLLED EXPERIMENTS

Experiments are done to see if a treatment has an effect. The treatment could be a drug, an advertising campaign, an educational method, etc.

experiment is an orderly procedure carried out to verify refute or establish an hypothesis

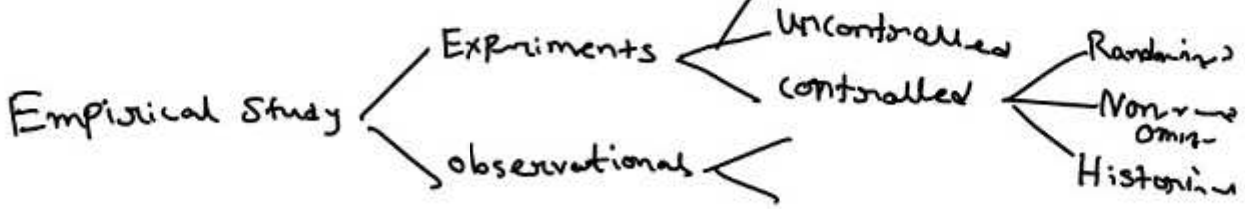
If the researcher ***decides*** who gets the treatment, the investigation is called an ***experiment.*** If the researcher just studies the effects of the treatment, but plays no part in choosing who gets the treatment then the investigation is called an observational study (see next chapter).

Example: CITES has asked me to do an experiment to see if using tablets are better than blackboard in terms of student engagement.

If the experiment ~~*compares*~~ the effects of getting the treatment to not getting the treatment, it's called a ***controlled*** experiment. If the researcher just gives the treatment to a group of subjects and records the results without making an explicit comparison, it's called an *uncontrolled* experiment.

To make the experiment controlled I can compare the effects of using tablet in this class with previous classes when I didn't use tablets. OR, I could divide the class in 2 groups. Use tablet in one group and blackboard in the other.

In a **controlled experiment** the researcher decides who gets the treatment (***treatment* group**) and then **compares** their responses to the responses of those who didn't get the treatment (***control* group**).



MAIN IDEA of a controlled experiment:

Controls should be as much like the treatment group as possible-- the only difference being the actual treatment. That way any differences in the responses of the 2 groups can be attributed to the treatment and not to some other difference between the groups.

TYPES OF CONTROLS in Designed Experiments: Randomized and Non-randomized

Randomized— The evaluator uses an objective chance procedure (like flipping a coin) to decide who's in the treatment group and who's in the control group.

Pros: Eliminates Bias/difference between the groups in all respects/variables Because it average outs the random differences between subjects/subgroups of subjects

Cons: 1) Expensive and time consuming
2) If the sample is not large it may not work

Non-Randomized: The treatment and control group are chosen by some other non-random method.

Different Types of Non-randomized Controls:

- Historical—Experimenter chooses people to be in treatment group and compares them to subjects from the past who didn't get the treatment.

Pros: 1) Cheaper. You may have data from Past
Example: If I compare using tablet in this class with not using it in last year

Cons: 1) Past conditions may not be same as present in all aspects.

- Experimenter tries to match treatment and control to make the 2 groups as alike as possible.

  Pros: 1) Its a good design only if you compare between the subjects / subgroups that were matched to be alike.

  Cons: 1) You may match them in variables you consider to be important. What about variables you haven't considered? Randomization eliminates bias on all variables

  Human Bias : 1 2 3 4

  Pick a number between 1 and 4

- Experiment chooses the treatment and control groups in the best interests of the subjects.

  Pros: None

  Cons: It's a bad design, because the experimenter might be biased and give the treatment to healthier or sicker people

- Experimenter decides to let the subjects choose whether to be in treatment or control group.

  Pros: None

  Cons: same as above - The subjects who are sicker might want to be in treatment group of a drug trial

Random assignment is ideal. As long as there are enough subjects, flipping a coin to decide who's in treatment and who's in control is ideal for 2 basic reasons:

- **Random assignment is most likely to make treatment and control groups as**

alike as possible because it eliminates human bias. With enough subjects, **random differences average out**, not only on the characteristics that the researcher has identified as relevant, but on *all* characteristics, including hidden ones that the researcher might not realize are important

- Random controls are based on ***chance procedures***. **Without a chance process we can't use statistics to analyze whether the treatment really worked**. This will be explained later in the course but the basic idea is this: You see a difference in the responses of the treatment and control group. There'll always be *some* difference in responses even if the treatment is a dud, just due to chance. So the question is how *big* a difference must you observe before you can determine that it's real and not just due to chance. Without random controls you have no way of answering that.

BLIND and DOUBLE-BLIND controlled experiments

Subject Bias-The Placebo Effect- Whether people *think* they've received a treatment affects their response. To separate the effects of the *actual* treatment from the *idea* of treatment, the assignment to treatment and control should be ***blind***. In other words, the subjects shouldn't know which group they're in. This can be achieved by giving the control group a fake treatment called a ***placebo.*** Since both groups *believe* they've been given the treatment, any differences in their responses can be attributed to the treatment itself and not the idea of treatment.

Evaluator Bias- Knowing which subjects received the treatment and which didn't can bias those evaluating the results. To eliminate such bias, the evaluators should ***blind*** to this knowledge. An experiment is called ***double-blind*** when neither the subjects nor those evaluating their responses know who's in the treatment group and who's in the control group.

IDEAL EXPERIMENTAL DESIGN:
Randomized, controlled, double-blind design

- Basic Method is comparison. Investigator **randomly** assigns subjects into 2 groups, **treatment** and **control.**

  Treatment gets treatment, control gets **placebo**. The responses of the 2 groups are compared to see if the treatment made any difference.

- **Double-blind:** neither the subjects nor the experimenters who measure their responses know who has received the treatment and who has received a placebo.

Example 1: An example from Homework. Skip the example in book
Hypothetical_Study_Comparing_2phrasings_of_test_question

Suppose I have included one particular multiple-choice question on the final for the past 2 years. I'd like to determine whether or not a new phrasing of this question would improve student responses. I decide to include the question with the new phrasing on this semester's final and then compare the percentage of students who answered the question correctly this semester to the percentage of students who answered the question correctly last semester with the old phrasing.

1) What sort of controls have I used in this experiment?

(a) Randomized controls
(b) Non-randomized, historical controls
(c) Non-randomized, non-historical controls
(d) No controls

2) Which of the following represents the best improvement on the design of the experiment in the previous question?

(a) Compare this semester's results (with the new phrasing) to the combination of all past semesters' results with the old phrasing (rather than just last semester's results).

(b) Give half of this semester's students a version of the final with the old phrasing and half of this semester's students a version with the new phrasing. Make sure that both groups have equal numbers of "A" students, "B" students, etc. Compare the results from these two groups.

(c) Give half of this semester's students a version of the final with the old phrasing and half of this semester's students a version with the new phrasing. Use a random procedure (like flipping a fair coin) to decide who gets which version. Compare the results from these two groups.

(d) This design needs no improvement.

Example 2: Portacaval Shunt—surgery to redirect the flow of blood in-patients who are at risk for hemorrhaging due to cirrhosis of the liver.

51 studies were done to see if the operation helped (see p. 7 of recommended text *Statistics*).

	Degree of Enthusiasm		
	Marked	Moderate	None
No Controls	24	7	1
Non-Randomized Controls	10	3	2
Randomized Controls	0	1	3

Why do you think the no-controls study were most enthusiastic about the surgery? (24/32 = 75%)

In the absence of a group to compare to, even small improvement might seem a large in subjective judgement - especially the doctors wanted the surgery to succeed

The non-randomized studies were much more enthusiastic than the randomized studies. (10/15 compared to 0/4) What are some possible explanations for that?

In non randomized trial the doctors may have chosen healthier patients for surgery and sicker patients for control

Compare 3 year survival rates in randomized vs. non-randomized studies.

	Randomized	Non-randomized
Surgery	60%	60%
Controls	60%	45%

Results from surgery are the same. The difference lies in the results from no surgery. The non-randomized controls are probably sicker to begin. Doctors often exclude the sickest patients from surgery.

Could you design these studies to include placebos and be double-blind? If so, how much of a difference do you think it would make?

Example 3: Arthroscopic Knee Surgery
A popular operation for arthritis of the knee was recently found to be no better than a placebo procedure in which patients were sedated while surgeons pretended to operate.

180 patients were randomly assigned to have the operation or to have placebo surgery in which surgeons simply made cuts in the knee so the patients wouldn't know if they had the surgery. Most patients in both treatment and control said their knee pain had improved, but neither group improved on any objective measures.

(You can view the full article on Lon Capa under Assignment #1.)

Example 4: Hormone Replacement Therapy
At least 6 million women in this country had been taking HRT for years on the basis of numerous studies which showed that women on the hormone therapy had a lower incidence of heart disease, stroke, colon cancer and bone fracture. But none of the studies used **randomized controls**.

Then a study sponsored by the National Institutes of Health, enrolled 16,608 healthy women from ages 50 to 79 and **randomly** assigned them to take hormone replacement therapy or a placebo. Much to everyone's surprise, the study was prematurely stopped in the summer of 2002 because it showed that while hormone treatment did lower the risk of hip fractures, it actually **increased** the risk of heart attack, stroke and breast cancer.

What are some possible reasons why the randomized controlled study showed an increase in heart disease and stroke while the non-randomized studies showed a decrease?

<table>
<tr><td colspan="2">Non-Random Studies</td><td>Randomized Study</td></tr>
<tr><td>Treatment- Healthy</td><td>Control-Sicker</td><td rowspan="2">Treatment and Control group are choses from the SAME population—Healthier women who are also eating well, exercising etc. , then half are randomly assigned to HRT and half to placebo.</td></tr>
<tr><td>Healthier Women who are more likely to be eating well, exercising, Keeping their cholesterol and blood pressure under control</td><td>Includes more sick women who can't take HRT, because they smoke, are at risk for various diseases, too poor to see a doctor, etc.</td></tr>
</table>

CHAPTER 2--OBSERVATIONAL STUDIES

OBSERVATIONAL *STUDIES*	***EXPERIMENTS***	**CONTROLLED *EXPERIMENTS***
The researcher has NO power over assignment into treatment and control groups. • Subjects themselves or simply fate determines who gets the treatment and who doesn't. • Researcher just **observes** what happens.	**The researcher decides who gets the treatment. A comparison/control group may or may not be present**	**Researcher decides how to divide the subjects into treatment and control groups.** • Subjects may be randomly assigned to treatment and control. • Subjects may be non-randomly assigned to treatment and control.

Note that both observational studies and controlled experiments compare treatment and control groups. The key difference is how the subjects got into the treatment or control groups. If the researcher assigned them, it's a controlled experiment. If not, it's an observational study.

Some studies have to be observational.
Examples:

- Effects of malnutrition--researcher cannot starve people for a study.
- Effects of alcohol on fetal development.

Effect of Smoking on lung Cancer

Observational studies are done out of necessity. Whenever possible it is better to do a controlled experiment.

Main Problem of Observational Studies: **They can show association, but difficult to show causation.**

Since the treatment and control groups just "happened" they are often very different from each other. These differences **confound** (mix up) the results.

Possible **confounders** make it difficult to prove **causation** by **association**. Did the treatment **cause** the response or is the treatment simply **associated** with the response?

Maybe both treatment and response were caused by a third **confounding factor**.

TREATMENT TRULY *CAUSES* RESPONSE.	**TREATMENT IS JUST *ASSOCIATED* WITH RESPONSE. A *CONFOUNDING* FACTOR IS CAUSING BOTH THE TREATMENT AND RESPONSE**
Treatment →→ Response	**Confounder** →→ Treatment →→ Response

Statisticians adjust for these confounding variables by breaking the control and treatment groups into **smaller more homogeneous sub-groups, where the confounding factor is the same.**

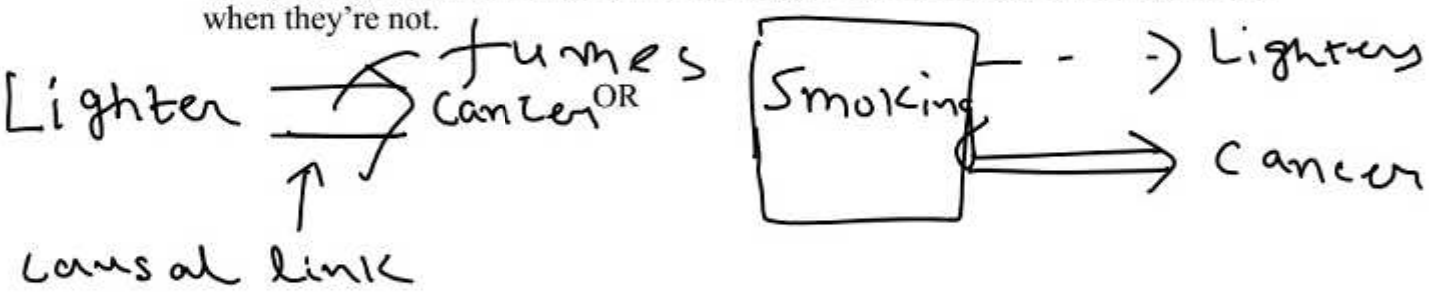
Examples of Observational Studies:
The following studies show <u>association</u>. Do they also show <u>causation</u>? What are the likely <u>confounders</u> and how would you adjust for them?

Let's start with an extreme example that will help you remember what a confounder is:

Lighters and Lung Cancer--Suppose ***observational*** studies showed that people who carry lighters have a higher rate of lung cancer. There are 2 possibilities:

- Lighters cause cancer. If so, how? The arrow in the diagram on the left represents a possible **causal (causal, as in cause, not casual!) link** that could connect lighters to cancer to support the argument that lighters ***cause*** cancer. But remember in observational studies, this causal link is pure speculation, observational studies do NOT provide evidence for causation. We first have to eliminate the possibility of major confounders before investigating causation, which brings us to the second possibility…

- Lighters don't cause cancer. If not, then why would people who carry lighters get more lung cancer? The mystery box in the diagram on the right represents a ***confounder*** or lurking variable-- something else about people who carry lighters that is the true cause of lung cancer making it look like lighters are responsible when they're not.



Below are either possible **confounders** (that would fit inside the mystery box on the right) mixing up the results making it look like lighters are to blame when they're not, possible **causal links** (that would fit inside the arrow) explaining *how* lighters may cause cancer, or neither (other causes of cancer but ones that don't help explain why people with lighters are getting more cancer).

- Genetics—Some people are more genetically prone to lung cancer than others
  i) Confounder ii) Causal Link iii) Neither

- Age—Lung Cancer rates rise with age
  i) Confounder ii) Causal Link iii) Neither

- Smoking Cigarettes—People who smoke cigarettes have a higher rate of lung cancer and are also more like to carry lighters.
  i) Confounder ii) Causal Link iii) Neither

- Lighter fluid—Perhaps inhaling the fumes from lighters can cause lung damage leading to cancer.
  i) Confounder ii) Causal Link iii) Neither

- Radon—Radon exposure raises one's risk for lung cancer
  i) Confounder ii) Causal Link iii) Neither

Now how can we determine whether it's the lighters or the confounders or (maybe some combination of both) that is the true cause of the lung cancer?

Break the population into sub groups where the confounding factor is the same.

Non Smokers	Compare lung cancer rates of those who carry lighters to those who don't.
Moderate Smokers	Compare lung cancer rates of those who carry lighters to those who don't.
Heavy Smokers	Compare lung cancer rates of those who carry lighters to those who don't.

In this case there's no difference in cancer rates between those who carry lighters and those who don't **within** each group. Of course the heavy smoker group has the highest cancer rates but rates between those who carry lighters and those who don't are the same in that group.

In other words, the lighters are just a marker for people who smoke—smoking causes cancer, not lighters.

Examples of Observational Studies:
The following studies show <u>association</u>. Do they also show <u>causation</u>? What are the likely <u>confounders</u> and how would you adjust for them?

- **Smoking and Liver Cancer** --Studies show that smokers have a higher rate of liver cancer. Does that show that smoking causes liver cancer?