Quick Check

You are interested in estimating the number of cat owners in Tamaqua. Explain how to gather a sample of 500 residents using:

1. a SRS

2. a Stratified Sample

3. a Cluster Sample

4.2A

Observational Studies vs. Experiments, The Language of Experiments, How to Experiment Badly

Observational study – Observes individuals and measures variables of interest but does not attempt to influence the responses (sample survey).

- Poor way to gauge the effect that changes in one variable have on another variable.

Ex. I sit in the back of the room and record the number of times Johnny taps his pencil on his desk.

Experiment – Deliberately imposes some treatment on individuals to measure their responses.

- When our goal is to understand cause and effect, experiments are the only source of fully convincing data.

Ex. I ring a bell every time Johnny taps his pencil on his desk to see how it affects his behavior.

Lurking variable – Variable that is not among the explanatory or response variables in a study that may influence the response variable.

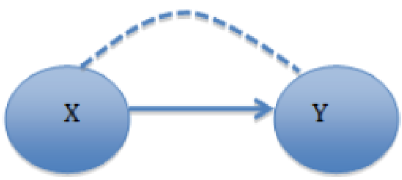
Create two problems:

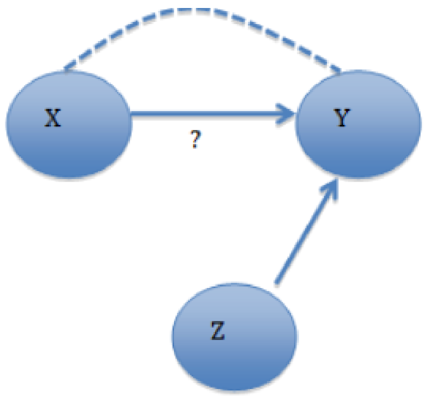
1. They can create extra variability in the response variable.

2. Have the potential to become confounded with the explanatory variable if not controlled.

Confounding – When two variables are associated in such a way that their effects on a response variable cannot be distinguished from each other.

- Each lurking variable is a potential confounding variable.





Example

For the study, researchers from Cincinnati Children's Hospital Medical Center surveyed over 4,700 children between the ages of 4 and 15 and their parents.Over 4 percent of the children included had ADHD. The researchers found that those children whose mother smoked during pregnancy were over twice as likely to develop ADHD than a child whose mother had not smoked. In addition, a child who had been exposed to lead, giving them high lead blood levels, were four times as likely to have ADHD, as compared to a child with low lead levels in his blood.

Based on this study, should we conclude that smoking during pregnancy *causes* an increase in the likelihood that a child develops ADHD? Explain.

Example

Many students regularly consume caffeine to help them stay alert. Thus, it seems plausible that taking caffeine might increase an individual's pulse rate. Is this true? One way to investigate this is to have volunteers measure their pulse rates, drink some cola with caffeine, measure their pulses again after 10 minutes, and calculate the increase in pulse rate. What is wrong with this experiment?

Even if every student's pulse rate went up, we couldn't attribute the increase to caffeine. Perhaps the excitement of being in an experiment or the sugar in the cola made their pulse rates increase.
In other words, there are many variables that are potentially confounded with caffeine.

The Language of Experiments

Treatment – A specific condition applied to the individuals. It is often applied at different levels.

There can be one or many treatments.

The Language of Experiments

Experimental Units – Smallest collection of individuals to which treatments are applied. When the units are human beings, they are often called **subjects.**

Example

A study published in the *New England Journal of Medicine* compared two medicines to treat head lice: an oral medication called ivermectin and a topical solution containing malathion. Researchers studied 812 people in 376 households in seven areas around the world. Of the 185 households randomly assigned to ivermectin, 171 were free from head lice after two weeks compared with only 151 of the 191 households randomly assigned to malathion.

Units - Households

Response - Lice free?

Exp - Type of Med.

Treatment - Drug A

Drug B

a) Identify the experimental units, explanatory and response variables, and the treatments in this experiment.

The experimental units are the households, not the individual people, since the treatments were assigned to entire households, not separately to individuals within the household. The explanatory variable is type of medication and the response variable is whether the household was lice-free. The treatments were ivermectin and malathion.

Example

Does adding fertilizer affect the productivity of tomato plants? How about the amount of water given to the plants? To answer these questions, a gardener plants 24 similar tomato plants in identical pots in his greenhouse. He will add fertilizer to the soil in half of the pots. Also, he will water 8 of the plants with 0.5 gallons of water per day, 8 of the plants with 1 gallon of water per day and the remaining 8 plants with 1.5 gallons of water per day. At the end of three months, he will record the total weight of tomatoes produced on each plant.

E. U. - Tomato plants

Exp Var - H_2O and Fertilizer

Resp. Var - Weight of T. P.

Treatments -

① Fert + 0.5 gal H_2O

② No Fert + 0.5 gal H_2O

③ Fert + 1 gal H_2O

④ No Fert. + 1 gal H_2O

" 1.5 gal H_2O

" 1.5 gal H_2O

a) Identify the experimental units, explanatory and response variables, and the treatments in this experiment.

The two explanatory variables are fertilizer and water. The response variable is the weight of tomatoes produced. The experimental units are the tomato plants. There are 6 treatments:

(1) fertilizer, 0.5 gallon

(2) fertilizer, 1 gallon

(3) fertilizer, 1.5 gallons

(4) no fertilizer, 0.5 gallons

(5) no fertilizer, 1 gallon

(6) no fertilizer, 1.5 gallons

Designing Experiments

Suppose we wanted to design an experiment to see if caffeine affects pulse rate.

Here is an initial plan:

- measure initial pulse rate
- give each student some caffeine
- wait for a specified time
- measure final pulse rate
- compare final and initial rates

What are some problems with this plan? What other variables are most likely to be sources of variability in pulse rates?

How to Experiment Well: The Randomized Comparative Experiment

- Experiments without random assignment are never a good idea

○ A good experiment needs comparison to prevent some lurking variables from becoming confounded with the explanatory variable.

○ Ex. Have some students drink cola with caffeine while others drink cola without caffeine

Quick Check

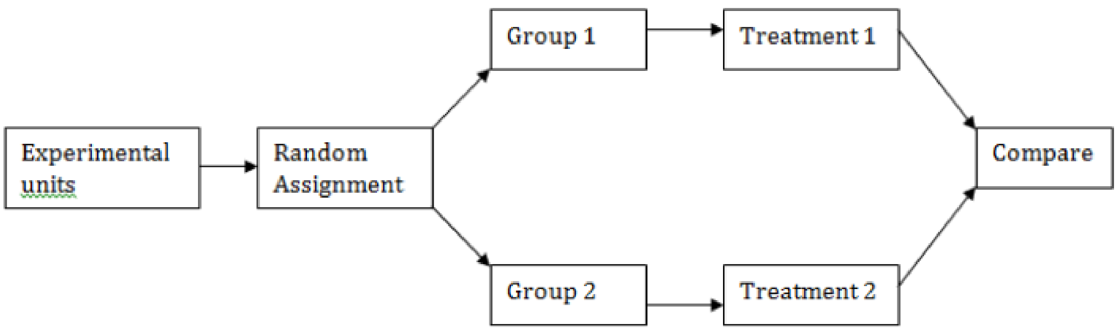
Two brands of toothpastes are being studied for effectiveness in reducing the number of cavities in children. There are 100 children available for the study.

- What are the experimental units?
- What is the response variable?
- What are the treatments?
- What are the factors?
- What are the possible confounding variables?

4.2B

How to Experiment Well, Three Principles of Experimental Design

Completely randomized design – The treatments are assigned to all the experimental units completely by chance.



Completely randomized design

Control group – Provides a baseline for comparing the effects of the other treatments

- Control groups are sometimes given an inactive treatment called a **placebo**

- **Placebo effect** – Response to a dummy treatment

- When a parent kisses a child's "boo-boo", the kiss has no "active treatment" but it makes the child feel better.

Completely randomized design

Single-Blind – The subjects don't know what treatment they received.

Double-Blind – Neither the subjects nor those who interact with them and measure the response variable know which treatment a subject received.

Completely randomized design

Random assignment – Experimental units are assigned to treatments at random using some sort of chance process.

Example

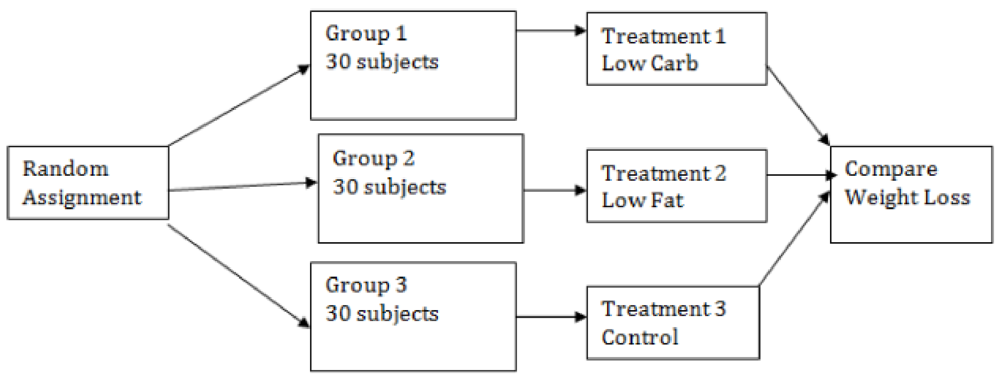
Suppose you have a class of 30 students who volunteer to be subjects in the caffeine experiment. Explain how you would randomly assign 15 students to each of the two treatments.

Using 30 identical slips of paper, write *A* on 15 and *B* on the other 15. Mix them thoroughly in a hat and have each student select one paper. Have each student who received an *A* drink the cola with caffeine and each student who received a *B* drink the cola without caffeine.

Example

A health organization wants to know if a low-carb or a low-fat diet is more effective for long-term weight loss. The organization decides to conduct an experiment to compare these two diet plans with a control group that is only provided with a brochure about healthy eating. Ninety volunteers agree to participate in the study for one year.

a) Outline a completely randomized design for this experiment. Write a few sentences describing how you would implement your design.



To implement the design, use 90 equally sized slips of paper. Label 30 of the slips "1", 30 of the slips "2" and 30 of the slips "3". Then, mix them up in a hat and have each subject draw a number without looking. The number that each subject chooses will be the group to which he or she is assigned. At the end of the year, the amount of weight loss will be recorded for each subject and the mean weight loss will be compared for the three treatments.

Three Principles of Experimental Design

Control lurking variables that might affect the response. Use a comparative design and ensure that the only systematic difference between the groups is the treatment administered.

Control all treatment groups not just the "control group"

Three Principles of Experimental Design

Random Assignment: Use impersonal chance to assign experimental units to treatments. This balances the effects of lurking variables that aren’t controlled on the treatment groups.

Three Principles of Experimental Design

Replication: Use enough experimental units in each group so that any differences in the effects of the treatments can be distinguished from chance differences between the groups.

Means “use enough subjects” not repeatability

Example

Explain how to use all three principles of experimental design in the caffeine experiment.

Control: There should be a control group that receives non-caffeinated cola. Also, the subjects in each group should receive exactly the same amount of cola served at the same temperature. Also, each type of cola should look and taste exactly the same and have the same amount of sugar. Subjects should drink the cola at the same rate and wait the same amount of time before measuring their pulse rates. If all of these lurking variables are controlled, they will not be confounded with caffeine or be an additional source of variability in pulse rates.

Randomization: Subjects should be randomly assigned to one of the two treatments. This should roughly balance out the effects of the lurking variables we cannot control, such as body size, caffeine tolerance, and the amount of food recently eaten.

Replication: We want to use as many subjects as possible to help make the treatment groups as equivalent as possible. This will give us a better chance to see the effects of caffeine, if there are any.

Example

Explain how we can conduct the caffeine experiment in a double-blind manner.

This means that neither the subjects nor the individuals measuring the results know what treatment was administered. In this case, we would need to ensure the subjects and those who came in contact with them were not told what type of cola they were drinking.

What is the advantage of this?

SUMMARY:

With randomization, replication, and control, each treatment group should be nearly identical, and the effects of other variables should be about the same in each group. Now, if changes in the explanatory variable are associated with changes in the response variable, we can attribute the changes to the explanatory variable *or the chance variation in the random assignment*.

4.2D

Blocking

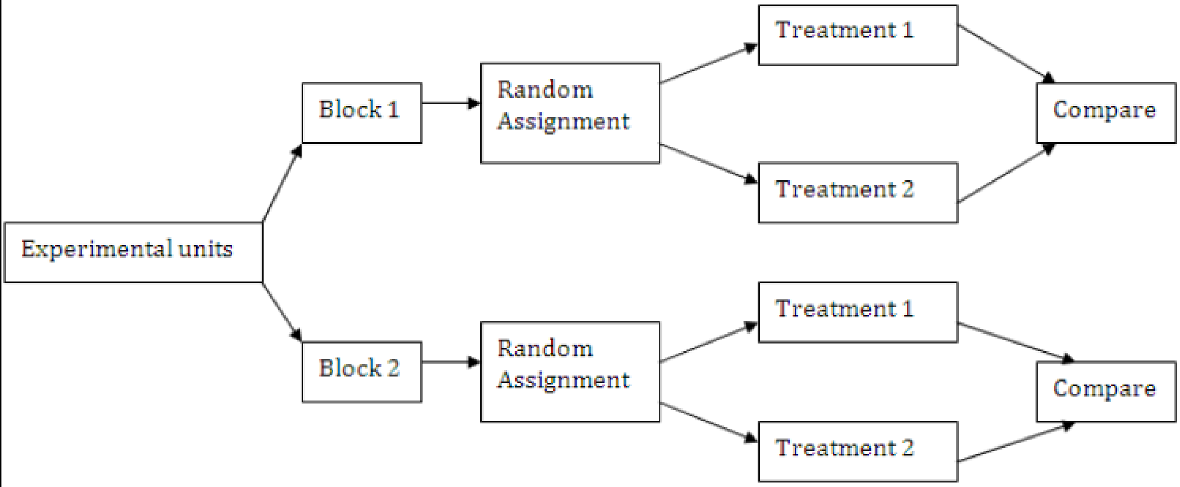
Matched Pairs Design

Block and Randomized Block Design

The random assignment of experimental units to treatments is carried out separately within each block.

Block - Group of experimental units that are known before the experiment to be similar in some way that is expected to affect the response to the treatments.

- Be able to explain why you chose a particular blocking variable. The best explanation is usually that the blocking variable has a strong association with the response variable.



Blocking in experiments is similar to stratification in sampling.

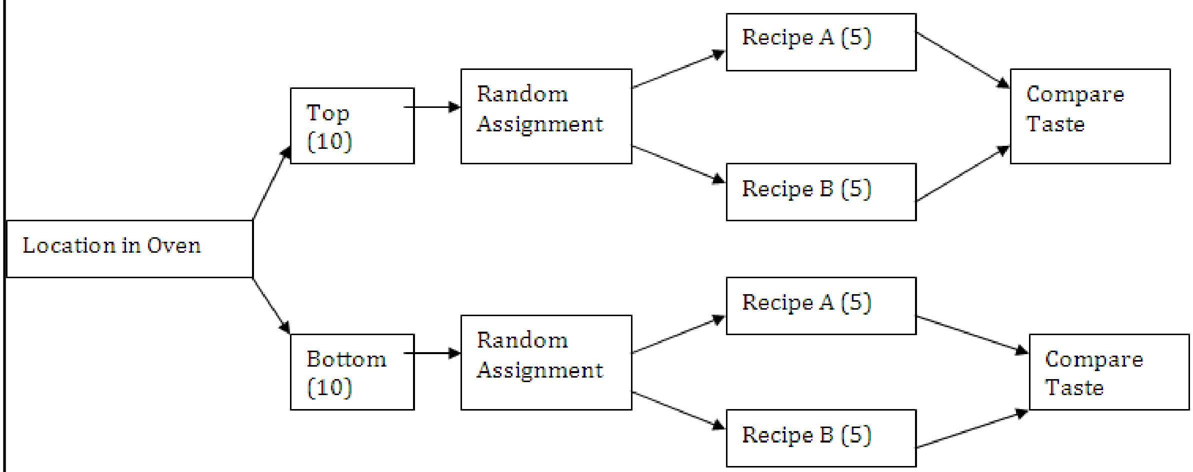
- Blocking accounts for a source of variability, just like stratifying. This means that blocking is another form of control.
- Blocks should be chosen like strata: the units within the block should be similar, but different than the units in the other blocks. You should only block when you expect that the blocking variable is associated with the response variable.
- Blocks, like strata, are not formed at random!

What are some variables that we can block for in the caffeine experiment? In general, how can we determine which variables might be best for blocking?

Example

Anne is an avid baker who would like to compare two different chocolate chip recipes (A and B). So she recruits 10 volunteer taste testers to rate each type of cookie on a scale from 1 (very bad) to 10 (very good). She will make 10 of each type of cookie, for a total of 20. Each cookie tray will hold only 10 cookies, so she will use two trays and bake them at the same time in the same oven, one sheet on the lower rack and one sheet on the upper rack. Because cookies might bake differently depending on which rack they are on, we will use the 10 locations on the lower-rack cookie sheet as one block and the 10 locations on the upper-rack cookie sheet as the other block. On each of the sheets, Anne will randomly place 5 of each type of cookie.

State the experimental units and the treatments imposed on the units. Also, draw a diagram illustrating her experiment.



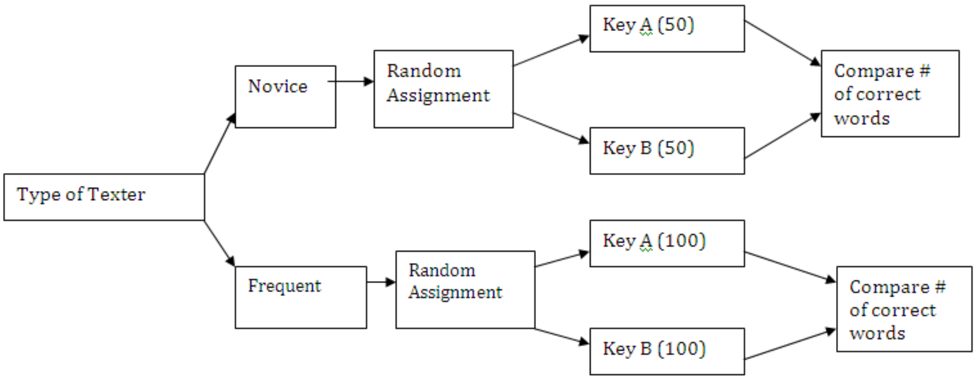
Example

A cell phone company is considering two different keyboard designs (A and B) for its new line of cell phones. Researchers would like to conduct an experiment using subjects who are frequent texters and subjects who are not frequent texters. The subjects will be asked to text several different messages in 5 minutes. The response variable will be the number of correctly typed words.

a) Explain why a randomized block design might be preferable to a completely randomized design for this experiment.

Because the subjects include people of varying texting abilities, a completely randomized design would lead to lots of variability in the response variable. However, in a randomized block design, the experienced texters would be compared to each other and the novice texters would be compared to each other. This will make it easier to see difference in the response variable if there is one.

b) Outline a randomized block experiment using 100 frequent texters and 200 novice texters.



To randomly assign the experienced subjects to the two groups, put 50 slips of paper labeled "A" in a hat and 50 slips of paper labeled "B" in a hat. Mix up the papers and have each subject select one to determine which design he or she will use. Follow a similar process for the 200 novice texters. After the experiment compare the mean number of words correct for the two different designs within each block.

Matched Pairs Design

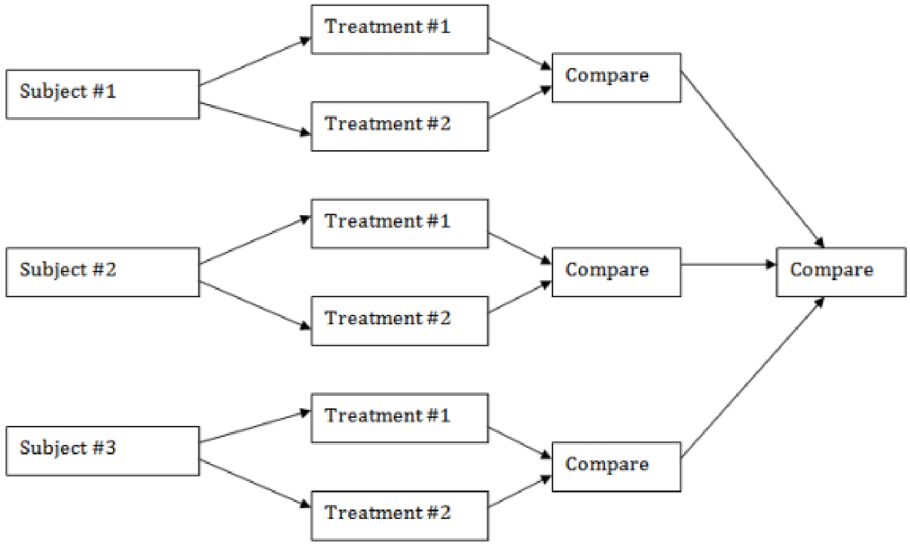
Create blocks by matching pairs of similar experimental units

- Special case of a randomized block design that uses blocks of size 2

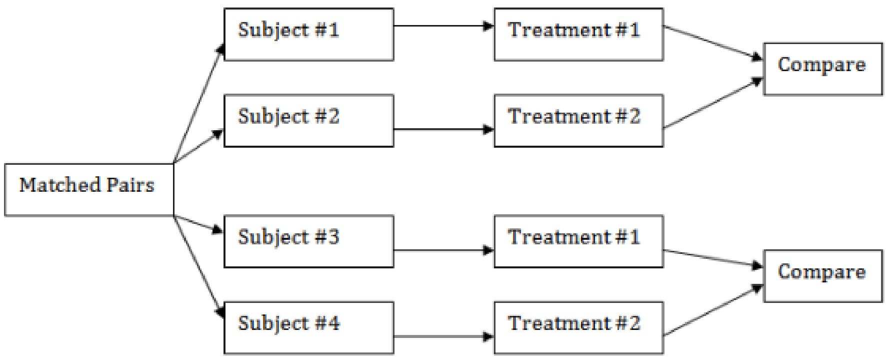
- Use chance to decide which member of a pair gets the first treatment. The other subject in the pair gets the other treatment

- Many times, the subjects in the pair are the same person

Matched Pairs – Single subject



Matched Pairs – paired subjects



Example

Are standing pulse rates generally higher than sitting pulse rates?

This experiment is conducted using a class of size 30

Completely randomized design: You will randomly assign half the students in the class to stand and the other half to sit. You can use the hat method, table D, or technology to carry out the random assignment. Once the two treatment groups have been formed, students should stand or sit as required. They should measure their pulses for one minute. Have the subjects in each group record their data.

Matched pairs design: Each student should receive both treatments in a random order. Since they already stood or sat in the first design, they just need to do the opposite now. As before, everyone should measure their pulses for one minute after completing the treatment. Have all the subjects record their data in a chart.

Example

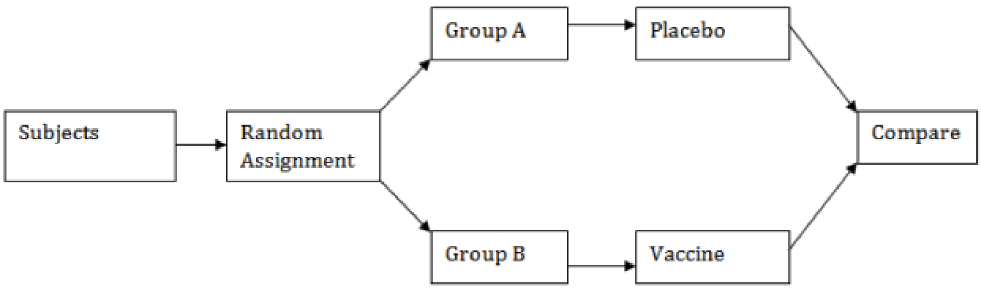
Consider the following hypothetical experiment. Acme Medicine is conducting an experiment to test a new vaccine, developed to immunize people against the common cold. To test the vaccine, Acme has 1000 volunteers - 500 men and 500 women. The participants range in age from 21 to 70.

Show how a completely randomized design, a randomized block design, and a matched pairs design might be applied by Acme Medicine to understand the effect of the vaccine, while ruling out confounding effects of other factors.

Completely Randomized

In this design, we will randomly assigned participants to one of two treatments. The same number of participants (500) will be given a placebo (A) and the real vaccine (B). To implement the design, use 1,000 equally sized slips of paper. Label 500 of the slips “A” and 500 of the slips “B”. Then, mix them up in a hat and have each subject draw a slip without looking. The number of colds reported for each treatment will be recorded and compared.

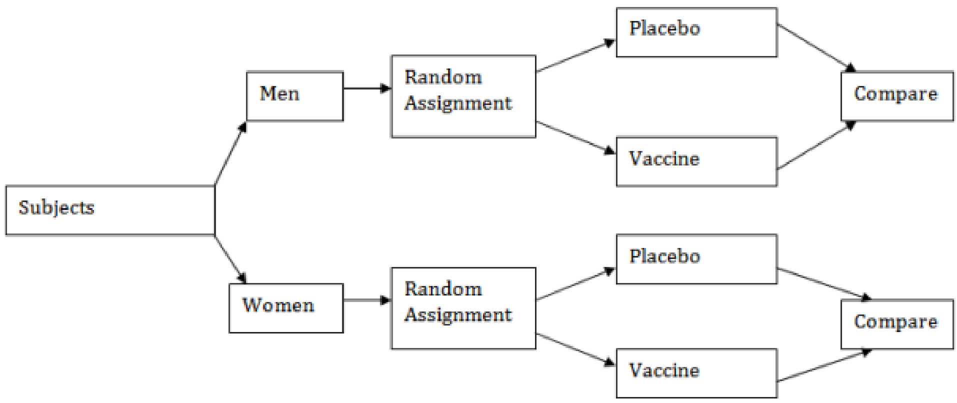
Completely Randomized



Randomized Block

It is known that men and women are physiologically different and react differently to medication. Participants are assigned to blocks, based on gender. Then, within each block, participants are randomly assigned to treatments. For this design, 250 men get the placebo, 250 men get the vaccine, 250 women get the placebo, and 250 women get the vaccine.

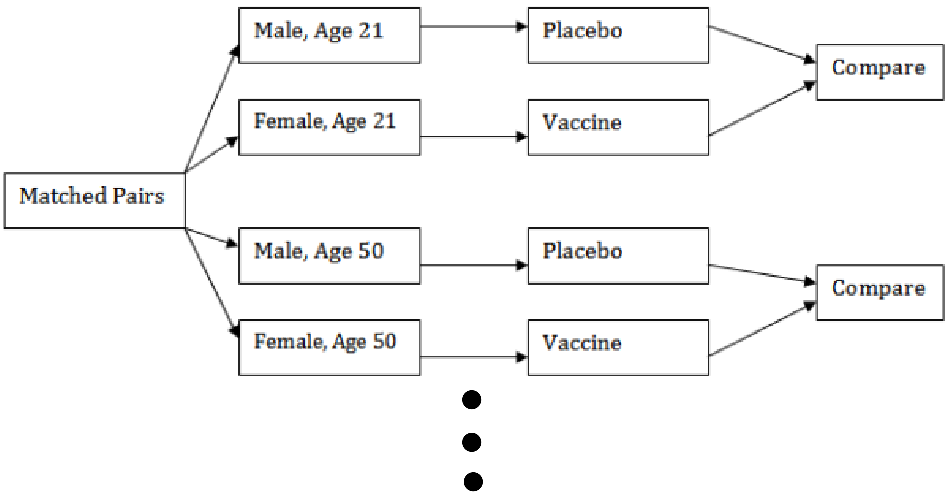
Randomized Block



Matched Pairs

The 1000 participants are grouped into 500 matched pairs. Each pair is matched on gender and age. For example, Pair 1 might be two women, both age 21. Pair 2 might be two women, both age 22, and so on. You could also match two subjects of the same age, but different gender.

Matched Pairs



SUMMARY:

Think about all possible sources of variability in the response variable. Control everything you can, block for the things you can measure but can't control, and randomly assign treatments within the blocks to balance out the effects of any remaining variables.

4.2C

Experiments: What Can Go Wrong? Inference for Experiments

Inference for Experiments

- In an experiment, researchers usually hope to see a difference in the responses so large that it is unlikely to happen just because of chance variation.
- We can use the laws of probability, which describe chance behavior, to learn whether the treatment effects are larger than we would expect to see if only chance were operating.
- If they are, we call them **statistically significant**.

Definition:
An observed effect so large that it would rarely occur by chance is called **statistically significant**.

A statistically significant association in data from a well-designed experiment does imply causation.

For example, if caffeine really has no effect on pulse rates, then the average change in pulse rate of the two groups should be exactly the same. However, because the results will vary depending on which subjects are assigned to which group, the average change in the two groups will probably differ slightly. Thus, whenever we do an experiment and find a difference between two groups, we need to determine if this difference could be attributed to the chance variation in random assignment or because there really is a difference in effect of the treatments.