Chapter 12

Experiments and Observational Studies

Homework

Read Chpt 12

Do p325

1, 7, 9, 10, 18, 22, 24, 26, 29, 33, 40, 41

Objectives

Identify observational studies and experiments, understanding the importance of controlled randomized experiments in establishing cause-and-effect relationships.

Observational Studies

Objective : Identify observational studies and experiments and the significance of the difference.

- In an **observational study**, researchers do not **assign** choices; they simply observe data as it is found.
  - The text's example looked at the relationship between music education and grades.
  - The researchers did not assign students to music education but simply **observed** students in the natural environment, thus it was an observational study.
- Researchers selected subjects that studied music and collected grade data. That data was on past grades making the study **retrospective.**
- Had the researchers selected subjects and collected data as events occured, the study would have been **prospective**.

Observational Studies

Objective : Identify observational studies and experiments and the significance of the difference.

- Observational studies are useful for seeing trends and possible relationships.
  - Keep in mind, it is **not possible** for observational studies, whether prospective or retrospective, to **suggest** a **causal relationship**.
  - When discussing observational studies be very careful not to use causal language such as; because, since, due to, as a result, etc.
  - That does not mean that there is no causal relationship. It simply means that we cannot confirm a causal relationship.

Randomized, Comparative Experiments

Objective : Identify observational studies and experiments and the significance of the difference.

- An experiment is a study design that **does** allow us to suggest a cause-and-effect relationship.
  - In an experiment, the experimenter must identify at least one **Independent (explanatory, predictor) variable**, called a **factor**, to manipulate; and at least one **Dependent (response) variable** to measure.
    - An **experiment**:
      - **Manipulates factor levels** to create treatments.
        - Experimenter determines the factors and levels the subjects will experience.
      - **Randomly** assigns subjects to treatment levels.
        - This is critical, without **random** assignment, there is no experiment.
      - **Compares** the responses of the subject groups across treatment levels.

Randomized, Comparative Experiments

Objective : Identify observational studies and experiments and the significance of the difference.

- In general, the individuals on whom or which we experiment are called **experimental units**.
  - When humans are involved, they are commonly called **subjects** or **participants**.
    - A **factor** is a variable that has an effect on the dependent (response) variable.
    - The specific values that the experimenter chooses for a factor are called the **levels** of the factor.
    - A **treatment** is a combination of specific levels from all the factors that an experimental unit receives.

Four Principles of Experimental Design

Objective : Identify observational studies and experiments and the significance of the difference.

1. Control:

Experimenters control sources of variation other than the factors we are testing by **making conditions as similar as possible** for all treatment groups. That is the purpose and goal of **random** assignment.

- The experimenter manipulates the **independent** variable by determining the different levels of the variable to which subjects will be assigned.
- The primary goal of control is to minimize the influence (variability in observations) of factors that are not part of the study.

Four Principles of Experimental Design

Objective : Identify observational studies and experiments and the significance of the difference.

2. Randomize:

Randomization allows experimenters to balance the effects of unknown or uncontrollable sources of variation (reduces variation due to factors other than those being studied).

- **Randomization** does not eliminate the effects of these sources, but it spreads them out (hopefully, evenly) across the treatment levels so that their effects are equalized within the factors of the independent variable.

- Without **randomization**, you do not have a valid experiment and cannot use the methods of statistics to draw conclusions from your study.

Four Principles of Experimental Design

Objective : Identify observational studies and experiments and the significance of the difference.

3. Replicate:

- **Repeat** the experiment by applying the treatments to several subjects. The more data collected, the more reliable the results.
  - The outcome of an experiment on a **single subject** is an anecdote, not a study.
  - When the experimental group is not a representative sample of the population, we might want to replicate an entire experiment for different groups.

Four Principles of Experimental Design

Objective : Identify observational studies and experiments and the significance of the difference.

4. Block:

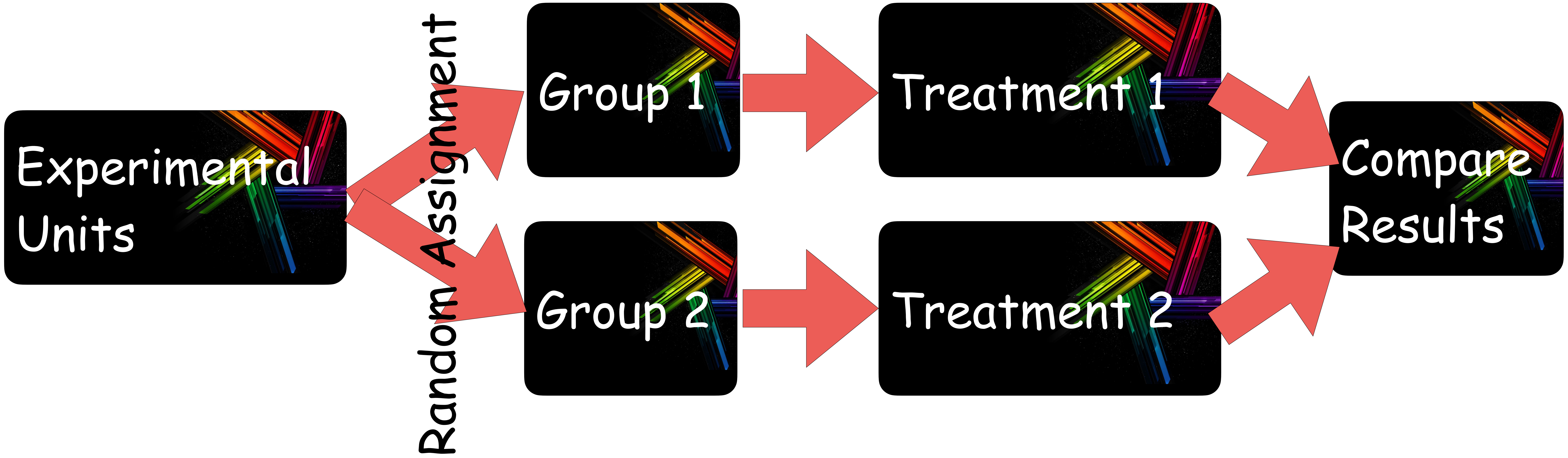
- Sometimes, attributes of the experimental units that are of no interest to the study, but that cannot be controlled or eliminated may have an effect on the outcomes of an experiment (i.e. gender).
- If we group together (**block**) individuals based on some similar characteristic and then randomize within each of these **blocks**, we can reduce much of the variability due to the difference between the blocks.
  - In a sense, we run simultaneous, parallel experiments.

- Note: Blocking is an important compromise between randomization and control, but is not necessary in an experimental design.

Diagrams of Experiments

Objective : Identify observational studies and experiments and the significance of the difference.

- It's often helpful to diagram the procedure of an experiment. (And AP wants to see it.)
  - The following diagram emphasizes the randomization of subjects to treatment groups, the separate treatments applied to these groups, and the comparison of results:



- I want to emphasize that the fancy treatment is solely for getting your attention. Do not do this in a study, or on anything you turn in to me or AP.

Significance

Objective : Identify observational studies and experiments and the significance of the difference.

- In an experiment (or observational study) we are looking at how different levels of the independent variable (treatments) result in different measures of the dependent variable. I should be obvious that there will always be differences.
  - That begs the question, How large do the differences need to be to say that any difference is due to the different treatments?
  - Differences that are larger than we would get just from chance (the randomization alone) are called:

statistically significant

- For now, the important point is that a difference is **statistically significant** if we do **not** believe that it is likely to have occurred simply **by chance**.

Significance

Objective : Identify observational studies and experiments and the significance of the difference.

- The meaning of statistical significance is so important in this course that I must repeat…
  - **Statistical significance** means the differences we observe are **greater than we would get just from chance**.
  - Results not **statistically significant** are considered to be the result of **chance**.

Experiments and Samples

Objective : Identify observational studies and experiments and the significance of the difference.

- Both experiments and sampling use randomization to get unbiased data, but they do so differently and for different goals:
  - Sampling is an attempt to estimate population parameters, so the sample needs to be as **representative of the population** as possible. To effect that representation, **random sampling** is critical.
  - Experiments try to assess the effects of treatments. Experimental units are not always drawn randomly from a population, but they are always **randomly assigned** to treatment levels **to make the experimental levels as similar as possible prior to any treatment**.

Two Randoms

Objective : Identify observational studies and experiments and the significance of the difference.

- **Random selection** is sampling to ensure your sample is representative of the population. Thus **random sampling.**

- **Random assignment** is designed to ensure treatment groups are as alike as possible **prior** to receiving treatment. Thus differences in response between the groups can be attributed to the differing treatments. Thus **random assignment.**

Control Treatments

Objective : Identify observational studies and experiments and the significance of the difference.

- Often, we want to compare a condition involving a specific treatment to a **no treatment** (status quo) condition.
  - A baseline ("SOP") measurement is called a **control** treatment, and the experimental units to whom the baseline is applied is called the **control group**.
- If an experimenter knows what treatment was assigned, that knowledge may very well, unintentionally or intentionally, influence the assessment of the response.
  - In order to avoid the bias that might result from knowing what treatment was assigned, researchers use **blinding**.

Blinding

Objective : Identify observational studies and experiments and the significance of the difference.

- There are two main classes of individuals who can affect the outcome of the experiment:
  - 1. Those participants who could **influence** the results. The subjects themselves and anyone administering the treatments. In other words, those actually participating in the study in some way.
  - 2. Those who **evaluate** the results (statisticians, research doctors, etc.)
- When all individuals in **either** one of these classes are blinded, an experiment is said to be **single-blind**.
- When everyone in both classes is blinded, the experiment is called **double-blind**.
  - Someone knows who received what, but as long as the subjects, participants, and evaluators do not know, the research is double-blind.

Placebos

Objective : Identify observational studies and experiments and the significance of the difference.

- It is human nature to respond to the applying of any treatment and that can induce an improvement.
  - To separate out the effects of "treatment", we use a control treatment that appears, for all intents and purposes, just exactly like the actual treatment.
    - The control (faux) treatment that looks just exactly like the treatment being tested is called a **placebo**. Placebos are the best way to keep participants (subjects and experimenters) from knowing whether or not subjects are receiving the actual experimental treatment.
    - A **placebo effect** occurs when taking the faux treatment actually results in a change in the dependent variable being measured.

Randomized, Comparative Experiments

Objective : Identify observational studies and experiments and the significance of the difference.

- To summarize; we have a good, valid, randomized, comparative experiment when the study is:
    - randomized.
    - comparative.
    - placebo-controlled.
    - double-blind.

Blocking

Objective : Identify observational studies and experiments and the significance of the difference.

- When groups of experimental units are similar, it's a good idea to gather them together into **blocks**.
  - **Blocking** is for experiments what **stratifying** is for sampling.
    - Both methods group together subjects that are similar and randomize within those groups as a way to reduce unwanted variation.

- The study of statistics is the study of variation. Probably the most important purpose of control in any study is the attempt to reduce variation resulting from any variable other than the variable of interest that is being observed and measured. All the methods we have discussed are attempts to reduce that unwelcome variability.

Blocking

Objective : Identify observational studies and experiments and the significance of the difference.

- Blocking isolates the **variability due to the differences between the blocks** to help insure measured differences between groups are, in fact, due to the treatments.
  - Blocking **reduces variability** by separating any effects of the blocking factor. Researchers are not often interested in the effects of the blocking factor.
  - When researchers randomly assigns subjects within blocks, the study design is a **randomized block design**.
  - For example; if we were concerned about the effect of grade level on an experiment, we could block by grade level, then randomly assign subjects within each grade level to the various treatment levels. Then, in effect, run parallel experiments to compare results **within each block**.

Blocking

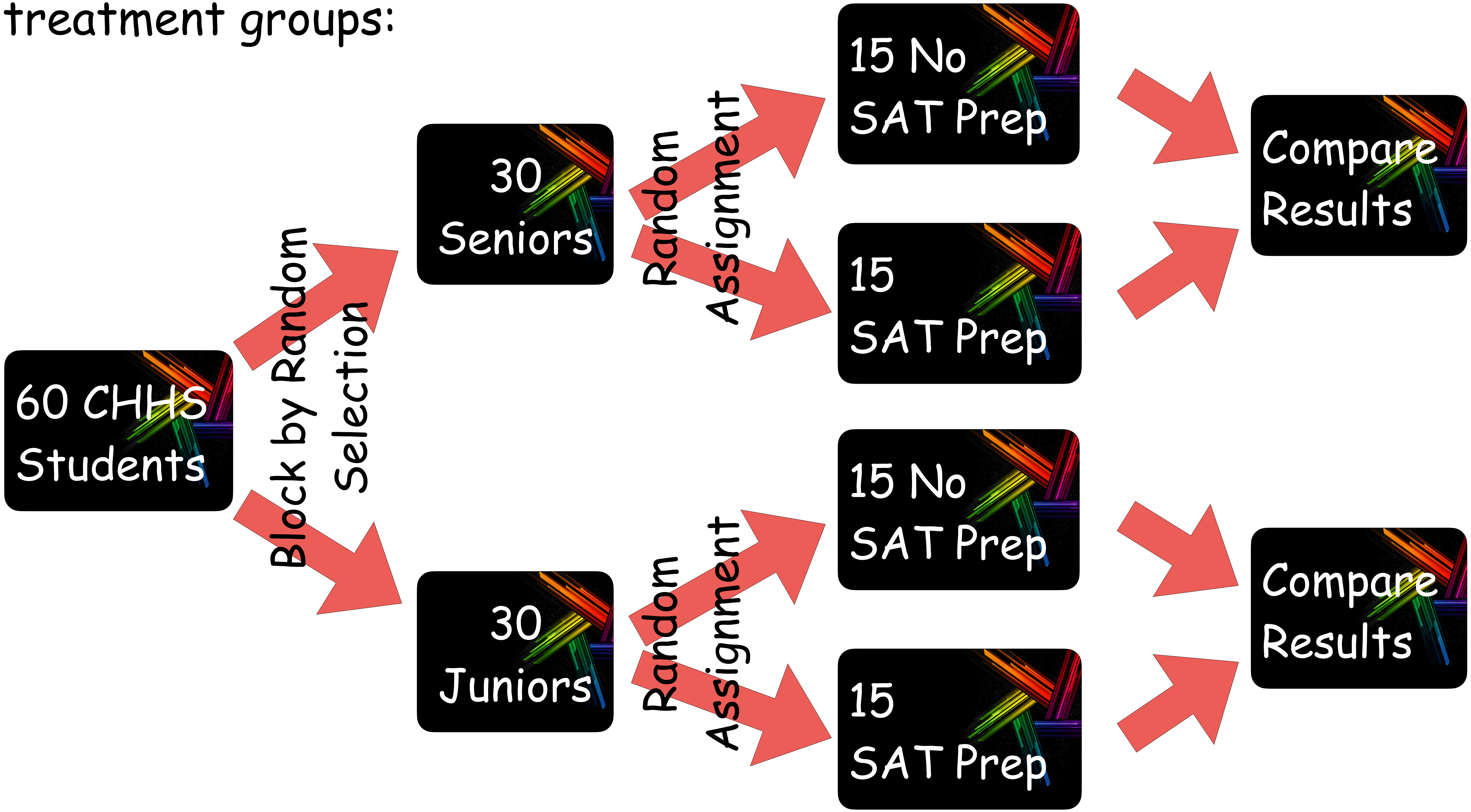
Objective : Identify observational studies and experiments and the significance of the difference.

- Suppose we are interested in the effect of SAT Prep classes on SAT scores. Further, suppose we are concerned that the grade level at the time the student takes the course might affect SAT results. To control the undesirable effect, we might block our subjects by Junior and Senior to control the variability due to year.
  - Our experimental units (participants) are students at CHHS.
  - Our treatments are SAT prep course and no SAT prep course.
  - Factors affecting SAT scores are myriad; IQ, motivation, courses taken in high school, study habits throughout high school, etc. To control all those factors we will randomly assign students to two treatments. We will block by year, so seniors will get the two treatments (course/no course) and juniors will get the same treatments.
- The response variable will be SAT test scores. We will compare results for seniors and we will compare results for juniors. We are unlikely to compare juniors with seniors.

Blocking

Objective : Identify observational studies and experiments and the significance of the difference.

- Here is a diagram of a blocked experiment with two treatment groups:



Matching

Objective : Identify observational studies and experiments and the significance of the difference.

- In a retrospective or prospective study, subjects are sometimes paired because they are similar in ways **not** under study. Twin studies are a good example.
  - **Matched pairs design** creates blocks by matching pairs of similar experimental units. Chance (random assignment) decides which member of a pair gets the first treatment. The other subject in that pair gets the other treatment.
  - Matching is a form of blocking by creating blocks of matched characteristics.
- The AP rubrics require that blocking descriptions focus on the use of blocking to **reduce variability** within the treatment groups. Not a discussion of comparing differences in the blocked groups.

Matching

Objective : Identify observational studies and experiments and the significance of the difference.

- Probably the most common type of matching is the "before and after" (pretest/ posttest). Each "pair" in the matched pairs design is actually one subject that gets both treatments, one following the other.
  - In that case, each participant serves as its own control. Since the order of the treatments may influence the response variable, we must randomize the order of the treatments for each subject. We randomly assign which treatment is given first among the subjects.
  - For example we may want to study the effect of different music styles (classical vs country) on ability to concentrate. Subjects will be required to complete a task requiring concentration. Some subjects will be randomly assigned to hear classical music first and country second while attempting the task. Other subjects will be randomly assigned to hear country first and classical second.

Make it More Complicated

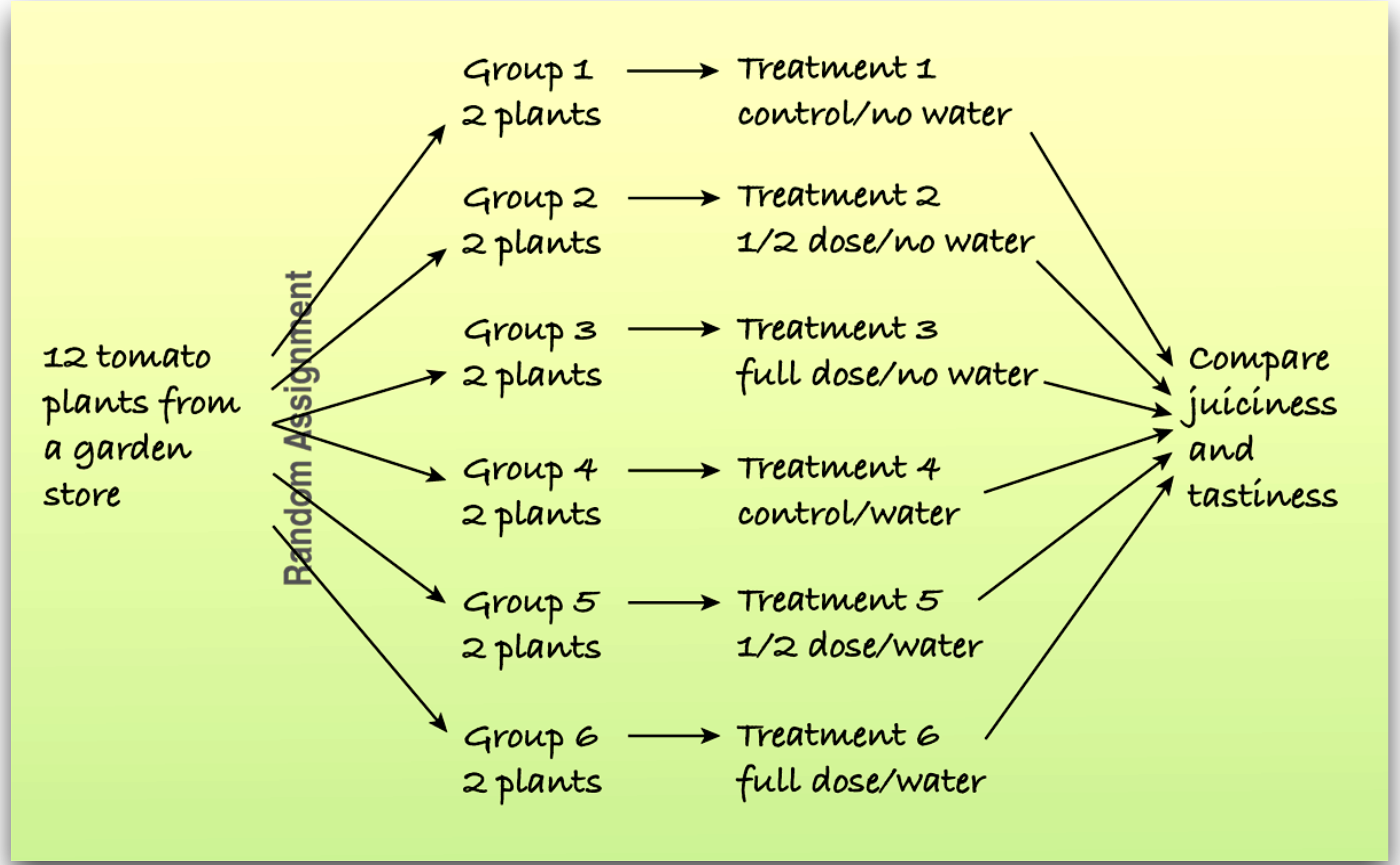
Objective : Identify observational studies and experiments and the significance of the difference.

- It is often important to include multiple factors in the same experiment in order to examine what happens when the factor levels are applied in different **combinations**.
  - For example, the following diagram from your text shows a study of the effects of different fertilizer/water combinations on the quality of tomatoes:
  - In this example from your text combinations of watering and fertilizing are compared for effect on the juiciness and "tastiness". Since tomatoes do not taste good, I cannot imagine how the "tastiness" is determined.
  - This examines the effects of different combinations of factor levels.

Make it More Complicated

Objective : Identify observational studies and experiments and the significance of the difference.

- This is what your diagrams should look like. Nothing fancy. Simple and straightforward.



Lurking or Confounding

Objective : Identify observational studies and experiments and the significance of the difference.

- When the changes in the levels of one factor are associated with changes in the levels of another factor, we say that these two factors are **confounded**.
  - When we have confounded factors, we cannot separate out the effects of one factor from the effects of the other factor.
  - A lurking variable (variable 1) creates an association between two other variables (variables 2 and 3) that leads us to believe that variable 2 may affect variable 3 because we are unaware of the effect of the sneaky, lurking variable 1.
  - Communities with greater number of churches tend to have higher crime rates. Think there may be a lurking variable there?

Lurking or Confounding

Objective : Identify observational studies and experiments and the significance of the difference.

- **Confounding** can arise in experiments when some other variables **associated with a factor** has an effect on the response variable.
  - Since the experimenter assigns subjects (at random) to treatments (not simply observing the treatments) and cannot separate out the associated confounding variable, the **confounding** variable has nothing to do with the assignment.
  - Because of the **confounding**, we cannot determine whether the results found are caused by our factor or by the **confounding** factor (or some combination of both working together).
  - Our goal is to measure how the factor affects the response variable. But when another variable is intertwined with the factor and we cannot tell how much effect each of these variables is contributing, we describe this problem as **confounding**.

Lurking or Confounding

Objective : Identify observational studies and experiments and the significance of the difference.

- I try to think of the difference between lurking and confounding like this:
  - Confounding is what happens when two variables are working together to affect changes in a third variable. We may know about the connection between the two variables but we cannot separate them.
  - Lurking variables are usually unknown to the researcher. Thus the effects observed may not be due the the variable the researcher believes, but unknowingly due to the variable hidden behind the experimental factors.
  - Suppose we are measuring your ability to drive a car. We measure your ability to go around a curve, come to a stop, park, etc. I test one of you in a Mustang, and another of you in a Lincoln. Do you think the model of vehicle would have an effect on the results? That is confounding.

Summary

Objective : Identify observational studies and experiments and the significance of the difference.

- Terms with which you should be familiar going forward
  - Experimental Units (humans = subjects)
    - Factors
    - Treatment
    - Independent Variable
      - Explanatory Variable
      - Predictor Variable
    - Dependent Variable
      - Response Variable
  - Observational Study
    - Retrospective
    - Prospective
  - Completely Randomized Experiment
  - Randomized Block Experiment
  - Matched Pairs Experiment
  - Placebo
  - Blinding

Example

Objective : Identify observational studies and experiments and the significance of the difference.

- **Newsweek** - "Of all pre-college curricula, the highest level of mathematics one studies in secondary school has the strongest continuing influence on bachelor's degree completion. Finishing a course beyond the level of Algebra 2 (for example, trigonometry or pre-calculus) more than doubles the odds that a student who enters postsecondary education will complete a bachelor's degree."
    - Anything wrong with this statement?
    - How might we design a study to test this theory? What problems do you foresee?

Experiment

Objective : Identify observational studies and experiments and the significance of the difference.

- Question: Do people better remember what they see or what they hear?
  - Materials: 2 number cards (1 each A & B), 1 stopwatch, 1 coin (Working in pairs)
  - 1. Collect your materials, but **DO NOT** look at the numbers on the index cards unless and until you are instructed to do so.
  - 2. Flip the coin to decide who will be student A and who will be student B, and each of you take the appropriate card. (still DO NOT look at the other side!)
  - 3. Flip the coin again to decide who will go first.
    - 3a. **Student A:** (1) **Look at the numbers on your card** while Student B times you for **30 seconds**, then give the card to Student B. (2) Next recite the alphabet aloud to Student B. (3) Finally, recite all of the numbers you remember on the card. Student B will record the number correctly recalled.
    - 3b. **Student B:** (1) **Student A will read the numbers on your index card to you three times.** (2) Student B will then recite the alphabet aloud to Student A. (3) Finally, recite all of the numbers you remember being read to you. Student A will record the number correctly recalled.

Experiment

Objective : Identify observational studies and experiments and the significance of the difference.

- Consider the experiment just completed.
  - What might be confounding factors?
  - How might we block this experiment?
  - How might we run this experiment as a matched pair?