Strength Shoe®

The Strength Shoe® is a modified athletic shoe with a 4-cm platform attached to the front half of the sole. Its manufacturer claims that this shoe can increase a person's jumping ability. A 1993 study published in the *American Journal of Sports Medicine* investigated the Strength Shoe® claim using 12 intercollegiate track and field athletes as study participants[1]. In this activity you will be replicating this investigation using data collected from 12 other subjects. to examine the following question:

How does random assignment make groups probabilistically equivalent?

Remember, when investigating whether or not one variable **causes an effect** on another, researchers seek to exert control by creating a comparison group and then assigning subjects to either the treatment group or the comparison group. An **experiment** is a study in which the experimenter actively imposes the treatment condition on the subjects. Ideally, the groups of subjects are identical in all respects other than the condition, so the researcher can then see the variable's direct effects on the response variable.

Study Design

[1] Cook, S. D., Schultz, G., Omey, M. L., Wolf, M. W., & Brunet, M. F. (1993). Development of lower leg strength and flexibility with the strength shoe. *American Journal of Sports Medicine, 21*, 445–448.

To determine the efficacy of training in Strength Shoes®, you are planning a study design in which 6 of your subjects will train in regular shows, and 6 will train in Strength Shoes®. Then all subjects vertical jump height will be measured.

The following subjects have volunteered to take part in your study.

Andreas	Jasmine	Mary
Antonio	John	Paul
Davieon	Ka Nong	Ringo
George	Keyaina	Tong

1. Assign the 12 subjects to two groups (ordinary shoe group and Strength Shoe® group) by assigning every other subject to the Strength Shoe® group, alphabetically. (Andreas = Ordinary Shoe group; Antonio = Strength Shoe® group, etc.) Write the names of the participants for the two groups below.

Strength Shoe® Group		**Ordinary Shoe Group**
Name		**Name**
Antonio		Andreas

Now imagine that you carried out the study and found that the participants you assigned to the Strength Shoe® group jumped 5” higher on average than the participants in the Ordinary Shoe group. Furthermore, a simulation showed that this difference was more than would be expected because of chance.

2. Are you willing to attribute the mean difference in jumping height to training in Strength Shoes®? Explain.

Observed Confounding Variables

In every study, there are potentially many factors (aside from the treatment) that may be related to the response variable and, in turn, affect the results of the study. Statisticians refer to these variables as **confounding (lurking) variables**.

3. Are there confounding factors that might affect the results? Identify a few.

If we can consider what these factors might be before we collect our data, we can measure them. We might refer to these confounding variables as **observed**

confounding variables. For example, two factors that might affect jumping distance are the participant's sex and their height. Luckily you have collected this information from your 12 participants. (Check out the *strength-shoe.tp* file.)

4. Based on your assignment of subjects from Question #1, compute the mean height for the participants in both groups. Based on these means, does it seem that height will bias the results? Or does it seem like the two groups are "equivalent" when it comes to height?

5. Based on your assignment of subjects from Question #1, compute the proportion of males and females in both groups. Based on these proportions, does it seem that sex will bias the results? Or does it seem like the two groups are "equivalent" when it comes to sex?

6. Describe how you might re-assign the participants to group so that sex and height do not bias the results?

7. Try to carry out a new assignment of subjects to make the two groups "equivalent" on sex and height.

Strength Shoe® Group				Ordinary Shoe Group		
Name	Sex	Height		Name	Sex	Height

Unobserved Confounding Variables

While some confounding variables may be identified and controlled in a study, others may not be identified initially by the researcher. These unidentified confounding variables may also mislead researchers into thinking that a treatment is effective (or not effective), when in reality, all of the difference in the response variable is a function of differences in the confounding variable, not differences in the treatment. In practice, erroneous results because of unobserved confounding variables are prevalent in every field. Even the smartest and most experienced researchers will probably not identify all of the confounding factors related to differences in the response variable need to be controlled prior to the study.

One unobserved confounding factor that might affect jumping distance are the participant's genetics. It turns out that there is a jumping gene, called Gene X.

- Open the s*trength-shoe.tp* TinkerPlots™ file.
- To see whether the participant has Gene X, right click anywhere in the case table and select Show Hidden Attribute.

8. Based on your re-assignment of subjects from Question #7, compute the proportion of participants with Gene X in both groups. Based on these proportions, does it seem that Gene X will bias the results? Or does it seem like the two groups are "equivalent" when it comes to Gene X?

9. Without having identified Gene X as a confounding variable prior to the study, how would we know if our results were biased because of Gene X?

Random Assignment

Bias could result because the groups are not "equivalent" on any confounding variable that is related to the outcome, in this case, jumping height. It turns out that the key to making groups "equivalent" on *all* confounding variables (both observed and unobserved) is to use random assignment in forming experimental groups.

For the remainder of this course activity, you will examine how random assignment "equalizes" not only the observed confounding variables (e.g., sex), but also unobserved confounding variables (factors we haven't yet thought of).

Remember that statisticians consider groups to be "equivalent" if the difference in means between the experimental groups have an expected value of zero (i.e., the average difference in means across all possible randomizations of the participants is zero). We will use TinkerPlots™ to simulate the randomization of subjects to experimental conditions many times to explore the expected difference in means for many confounding variables.

Confounding Variable: Sex

To explore whether random assignment makes the groups "equivalent" on sex, we are going to simulate the randomization of subjects to experimental conditions many times, and examine the expected difference in means for the dummy coded sex variable. (Note: We are thinking about the confounding variable of sex as the "outcome" in this simulation.)

Set up a sampler to randomly assign participants to the experimental conditions. To do this:

- Set up a sampling device that includes the 12 fixed values for sex. Dummy code these values so that 0 = male and 1 = female. Name this device Female.
- Set up a sampling device that includes the 12 fixed values for condition; 6 ordinary shoe labels, and 6 Strength Shoe labels. Name this device Group.
- Click Run to randomly assign the participants to groups. (Note that since we are only interested in the participant's sex, their name is just superfluous information and is not needed for the simulation.)
- Plot the attributes Female (x-axis) and Group (y-axis) in a single plot;

- Compute and display the mean for each group.

10. Report the **proportion of females** in each group. Also subtract these two proportions (taking the Strength Shoe® group's proportion minus the ordinary shoe group's proportion).

    Proportion of females in Strength Shoe® Group:

    Proportion of females in Ordinary Shoe Group:

    Difference in proportions (Strength Shoe®– Ordinary Shoe):

This is just a single random assignment (trial). Whether the groups are statistically equivalent asks: Is the expected difference zero across many, many trials?

- Collect the proportion of females in both the Ordinary Shoe group and the Strength Shoe® group.
- Use the Formula Editor to compute the difference in the proportion of females between the two groups. (Note: Subtract the Ordinary Shoe group from the Strength Shoe® group.)
- Collect 499 more trials.
- Plot the 500 differences.
- Organize and fully separate the results (no bin lines) for the plot.
- Compute and display the average for the 500 randomized differences.

11. Sketch the plot below.

12. Where is this plot centered? What does this imply about the expected difference in the proportion of females in the two groups? Explain.

Confounding Variable: Height

To explore whether random assignment makes the groups “equivalent” on height, we are going to simulate the randomization of subjects to experimental conditions many times, and examine the expected difference in mean height. (Note: We are thinking about the confounding variable of height as the “outcome” in this simulation.)

Set up a sampler to randomly assign participants to the experimental conditions. To do this:

- Set up a sampling device that includes the 12 fixed values for height. Name this device Height.
- Set up a sampling device that includes the 12 fixed values for condition; 6 ordinary shoe labels, and 6 Strength Shoe labels. Name this device Group.
- Click Run to randomly assign the participants to groups.

- Plot the attributes Height (x-axis) and Group (y-axis) in a single plot;
- Organize and separate the cases based on both attributes.
- Compute and display the average height for both groups.

13. Report the average height for each group. Also find the difference in these two averages (taking the Strength Shoe® group's average minus the ordinary shoe group’s average).

    Average height in Strength Shoe® Group:

    Average height in Ordinary Shoe Group:

    Difference in average height (Strength Shoe®– Ordinary Shoe):

Again, this is just a single random assignment and we want to get a sense of the difference in the average height across many random assignments.

- Collect the average height in both the Ordinary Shoe group and the Strength Shoe® group.
- Use the Formula Editor to compute the difference in the means between the two groups. (Note: Subtract the Ordinary Shoe group from the Strength Shoe® group.)
- Collect 499 more trials.

- Plot the 500 differences.
- Organize and fully separate the results (no bin lines) for the plot.
- Compute and display the average for the 500 randomized differences.

14. Sketch the plot below.

15. Where is this plot centered? What does this imply about the expected difference in average heights between the two groups? Explain.

Unobserved Confounding Variables

Although it is great that random assignment will tend to make the groups "equivalent" on the observed variables (e.g., proportion of females; average height), what about confounding variables that we did not consider or even think about.

To explore whether random assignment makes the groups "equivalent" on Gene X (an unobserved confounding variable), we are going to simulate the randomization of subjects to experimental conditions many times, and examine the expected difference in the proportion of participants with Gene X. (Note: We are thinking about the confounding variable of Gene X as the "outcome" in this simulation.)

Set up a sampler to randomly assign participants to the experimental conditions. To do this:

- Set up a sampling device that includes the 12 fixed values for Gene X. Dummy code these values so that 0 = Does not have Gene X and 1 = Has Gene X. Name this device GeneX.
- Set up a sampling device that includes the 12 fixed values for condition; 6 ordinary shoe labels, and 6 Strength Shoe labels. Name this device Group.
- Click Run to randomly assign the participants to groups.

- Plot the attributes GeneX (x-axis) and Group (y-axis) in a single plot;
- Organize and separate the cases based on both attributes.
- Compute and display the average for both groups.

Again, this is just a single random assignment and we want to get a sense of the difference in the average height across many random assignments.

- Collect the proportion of participants having Gene X in both the Ordinary Shoe group and the Strength Shoe® group.
- Use the Formula Editor to compute the difference in the proportion of participants with Gene X between the two groups. (Note: Subtract the Ordinary Shoe group from the Strength Shoe® group.)
- Collect 499 more trials.

- Plot the 500 differences.

- Organize and fully separate the results (no bin lines) for the plot.
- Compute and display the average for the 500 randomized differences.

16. Sketch the plot below.

17. Where is this plot centered? What does this imply about the expected value of the difference in proportion of participants with the Gene X in the two groups? Explain.