ESS 200 (Stapleton) Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Inquiry/Experimental Design Notes

Suppose your experimental question is “do students get better test grades when the completion of homework is mandatory?”

1. What is the manipulated variable?
2. Provide a synonym for manipulated variable.
3. What is the responding variable?
4. Provide a synonym for responding variable.
5. What types of variables should be controlled in this experiment?
6. What types of variables should not be controlled in this experiment?
7. What would you have to do in order to have a good sample size for this experiment?

8. What is the purpose of having a very clear protocol for collecting data in an experiment?

Suppose your hypothesis is “the choice to make homework mandatory (or not) does have an effect on students’ grades.”

6. Is this a one-tailed or two-tailed hypothesis? Explain why.

7. In addition to the quote above, what other information should be added to your hypothesis?

8. As a student conducting this experiment, should you be concerned about the possibility of others suspecting that bias may have influenced the results? Why or why not?

9. What does it mean if an experiment has an “n of 50?”

9. What is the significance cutoff that is generally accepted in scientific investigations?

10. Suppose your statistical test results in p=0.10. Which of the following best explains what that 0.10 means?

a. There is a 10% chance that your hypothesis is correct.

b. There is a 100% chance that your hypothesis is correct.

b. There is a 10% chance that there is a real difference between the two test groups.

c. There is a 10% chance that there is no real difference between the two test groups.

11. Suppose you are given 20 car tires that all contain air at exactly 35psi of air pressure and 20 car tires that all contain air at exactly 70psi of air pressure. When you measure them with your pressure gauge, your gauge reads exactly 50psi for the first twenty and exactly 100psi for the other twenty. This means…

a. Your gauge is precise and accurate

b. Your gauge is precise, but not accurate

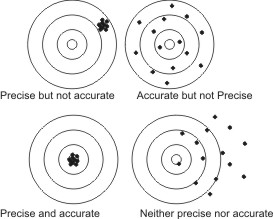
c. Your gauge is accurate, but not precise

d. Your gauge is neither accurate nor precise

12. If you were answering the question, “does outside air temperature affect tire pressure,” would your pressure gauge (from the previous question) be suitable for use in this experiment? Explain why or why not.

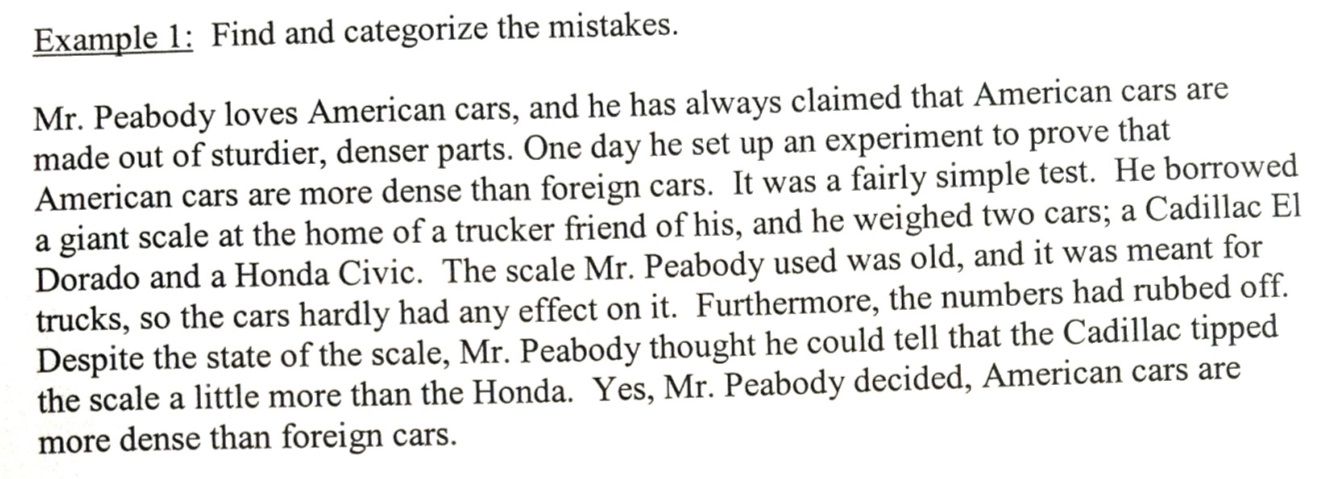
13. Suppose Gertie is trying to answer the question “does gender affect the sports that we like?” Gertie asks 10 boys and 10 girls to each list their two favorite sports. Given Gertie’s approach, why is it impossible to use either a T-Test or a Mann-Whitney test to answer the question?

14. What is the difference between correlation and causation?

1. **Scientific studies control variables**
   1. A ***variable*** is a factor that may vary. In scientific testing, the important variables are those that might affect your measurements/observations. In a simple investigation, exactly one variable should be intentionally made to vary. One other variable should be allowed to do as it pleases. The rest of the variables should be controlled (prevented from varying).
      1. The concern with extra variables (a.k.a. ***confounding* variables**) is that their variation can have confusing effects on your experimental results, thus rendering your investigation inconclusive.
      2. Potential confounding variables must therefore be ***controlled***, which means that they are *kept the same* for the sake of a fair and valid comparison. Suppose, for example, that we are trying to compare the swimming speeds of British and American Swimmers. If, during our swimming tests, the British team wears Speedos and the Americans wear Carhartt overalls, the extra variable of *type of swimming suit* would be a confounding variable that would invalidate the experiment. Similarly, if the Americans were allowed to swim with a freestyle stroke, and the British were forced to swim the butterfly, the confounding variable of *swimming stroke* would invalidate the experiment.
   2. **Manipulated variable** (a.k.a. independent variable or experimental variable): the one variable that should be intentionally varied in an experiment. This is the variable being tested. Suppose, for example, that we are trying to compare the swimming speeds of British and American Swimmers. In this case, nationality is the manipulated variable. We need to test some British swimmers and some American swimmers. For this experiment to work, swimmer nationality needs to vary.
   3. **Responding variable** (a.k.a. dependent variable): this is the variable that you measure in order to observe a possible response to varying the manipulated variable. Measuring the dependent variable tells us the *result* of varying the experimental variable. In the swimmer example, swimming speed is the responding variable, whereas nationality is the manipulated variable.
   4. **Independent and Dependent Variables:** These are alternative names that can sometimes be used in place of *manipulated variables* and *responding variables*, respectively. In general, researchers are trying to find out if the dependent variable *depends on* the independent variable.
   5. **Controlling Variables Means Eliminating Bias**:Bias is prejudice or the favoring of a certain outcome.If, for example, an American investigator is collecting the data in the swimmer experiment, he or she might cheat in favor of one nationality or the other. Controls must be in place to prevent cheating.
   6. **Controlling Variables Requires Careful Protocols**: Clear data collection protocols (steps to follow) minimize the chances that unwanted variables can sneak into an experiment. The clearer the protocol, the more confidence we can have in its implementation. When the details of data collection are left up to the measurer, it is easier for that person to introduce bias to the measurement process.
2. **Scientific studies need sufficiently large sample sizes**.
   1. ***Sample size*** tells you how much testing was done. In the swimmer example, testing one British swimmer and one American swimmer would not be a large enough sample size. Testing ten swimmers from each country would be better. Testing 1,000 swimmers would be even better.
   2. **Statistics tells us how large the sample size needs to be**. Statistical analysis boils down to this question – **what is the probability that our two groups are the same, and that the difference we are seeing is due to chance and natural variation?** If the chance of this being true is **low**, then we can be confidence that the two groups really are different.
      1. Consider the swimmer example… Suppose that we randomly select ten swimmers from each country, and using a careful and consistent protocol we find that all ten American swimmers are faster than all ten British swimmers. Given these consistent results, is it *possible* that there is really no difference between American and British swimmers? Is it possible that, out of the thousands of swimmers from each country, we just happened to select ten faster swimmers from America and ten slower swimmers from Britain – even though the two populations are, in reality, identical? Is it possible? Is it likely? What is the probability of seeing this striking difference if the populations are essentially identical? If this probability is very low, we generally reject the idea that there is no difference between the populations, and we say that it is very likely that there is an actual difference between the populations. On the other hand, if we only chose one swimmer from each country, there is no low possible outcome that could have a low enough probability to reject the idea of no difference between the populations.
      2. In many scientific studies, in order for findings to be considered “significant,” statistics must show that there is at least a 95% chance of two test groups being different. 95% confidence is called the ***significance cutoff***. What this really means, in statistical testing, is that the probability of there being no difference between the groups must be 5% or lower. This percent chance is called the **P-value**, and it is written as a decimal. A P-value of 5% is actually written as P = .05.
      3. If a study shows with 95% certainty that British swimmers are faster than American swimmers, is there a chance that the study could be wrong? Yes, a 5% chance!
      4. How big should your sample size be? If you do not know how to use statistical tools to answer this question, your sample size should be as big as you can reasonably manage. Consider this… how many times would you want someone to flip a coin so that you could decide whether or not it is really a “fair coin?” One flip clearly wouldn’t be enough.
   3. **“n”** is often used to represent the number of samples in an experiment. If you conduct a test with 50 human subjects, you have an “*n of 50*.”
3. **Solid scientific studies rely on precise and accurate numerical measurements**.
   1. **Precision**: consistency; when something is measured precisely multiple times, the measurement will always be the same. This is not the same as accuracy; you can be consistent without hitting the target. When you are looking for a potential difference between two groups, precision is arguably more important than accuracy. A good way to tell if your measurements are precise is to ask the question, “if I repeated this measurement multiple times, would I always get the same answer?” If the answer to that question is yes, then your measurement process is precise.
   2. **Accuracy**: how close measurements are to the right answer, on average.
      1. The virtue of having accurate measurements relates to logical validity, below. If our measurements are not accurate, then we may not be testing what we think we are testing.
   3. **Data need to be quantitative (numerical)**. If the results of an experiment cannot be converted to numbers in a meaningful way, then we cannot apply statistics to see if they are significant (95% certainty). We must be able to apply statistics in order to assess the strength of our conclusions.
4. **Investigations and their conclusions must be based on valid logic.** We should always ask “are we really testing what we think we are testing?” You can do everything right, but if your study is based on an illogical premise, it may be worthless.
   1. Examples:
      1. An example of a logical experiment: You want to know if sheep are heavier than pigs, so you weigh both groups. The pigs weigh more, on average, so you conclude that pigs are heavier.
      2. An example of an illogical experiment: You want to know if bowling balls are heavier than snakes. Using a meter stick, you measure their lengths and conclude that snakes are longer, so they must be heavier.
      3. There are many other ways to be illogical. We will deal with them as they arise.
   2. Do not confuse correlation with cause. A correlation between two things means that a change in one variable is accompanied by a change in the other; the two variables are tied together in some way. This does not necessarily mean that one phenomenon is *causing* the other. For example, there is a correlation between hair length and frequency of wearing dresses, but we know that neither of these phenomena causes the other.

The passages below summarize a couple of really bad investigations. Read the passages, identify mistakes, and then categorize the types of mistakes. Some categories are:

* Sufficient **Sample size**
* Controlling **Variables**
* **Logical Reasoning**
* Eliminating **Bias** (or concerns about bias)
* Precision of **Measurements**



The passages below summarize a couple of really bad investigations. Read the passages, identify mistakes, and then categorize the types of mistakes. Some categories are:

* Sufficient **Sample size**
* Controlling **Variables**
* **Logical Reasoning**
* Eliminating **Bias** (or concerns about bias)
* Precision of **Measurements**

