Designing an Experiment
Convincing evidence or not?

**Why?** There are many times in life when someone makes a statement about the effectiveness of a product or why a certain condition exists in nature. How do you know if such statements are correct? In order to decide if such claims are reasonable or not, one needs to examine all of the relevant evidence used to make such claims. One valuable tool for gathering evidence is to design an experiment. In this activity we are going to explore ways to design an experiment which will provide convincing evidence to support or reject a given claim.

**Model 1 -** Take turns reading out loud the cartoon and then each of the experiments in Model 1.

---

<table>
<thead>
<tr>
<th><strong>Friend A’s Experiment</strong></th>
<th><strong>Friend B’s Experiment</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beginning Question</strong></td>
<td>Does the color of an iPod affect the length of playtime on one charge?</td>
</tr>
<tr>
<td><strong>Procedure &amp; Tests</strong></td>
<td>One red iPod and one silver iPod were charged fully and then used until the battery was completely dead.</td>
</tr>
<tr>
<td><strong>Data &amp; Observations</strong></td>
<td><strong>Red iPod</strong></td>
</tr>
<tr>
<td></td>
<td>playback time</td>
</tr>
<tr>
<td></td>
<td>5 days</td>
</tr>
<tr>
<td><strong>Trials</strong></td>
<td><strong>iPods</strong></td>
</tr>
<tr>
<td></td>
<td>Red A</td>
</tr>
<tr>
<td></td>
<td>Red B</td>
</tr>
<tr>
<td></td>
<td>Red C</td>
</tr>
<tr>
<td></td>
<td>Silver A</td>
</tr>
<tr>
<td></td>
<td>Silver B</td>
</tr>
<tr>
<td></td>
<td>Silver C</td>
</tr>
<tr>
<td></td>
<td>Silver D</td>
</tr>
<tr>
<td></td>
<td>Silver E</td>
</tr>
</tbody>
</table>

*Note: The times for iPod Silver D were not used in this comparison because they seem to be outside the range of the other identical iPods. I assumed there must be something wrong with that specific iPod.*

The color of an iPod impacts how long the MP3 player will run on one charge and the red iPods last longer than other colored iPods.

The color of the iPod does not impact how long the MP3 player will run on one charge.

---

©HSPI – The POGIL Project  Limited Use by Permission Only – Not for Distribution  Designing an Experiment CIYvM
1. Write down some descriptive words (NO MORE than 8) which summarize your group's first impressions of each of the two experiments.

<table>
<thead>
<tr>
<th>Friend A's Experiment</th>
<th>Friend B's Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. How many iPods were tested by Friend A? How many iPods were tested by Friend B?

3. How many times did Friend A test each iPod? How many times did Friend B test each iPod?

4. a. By what process did Friend A “measure” the battery life of the iPod?
   
   b. By what process did Friend B “measure” the battery life of the iPod?
   
   c. Which of the two methods for “measuring” the battery life of the iPod provides the most reliable measure of the battery life? Explain your reasoning.

5. Take 1 minute to silently create a list of possible reasons why the red iPod in Friend A's experiment played music longer than his silver iPod. After the 1 minute of silent thinking time, work as a group to expand your list, considering as many reasons as possible.

6. Which, if any, of the problems your group listed in the previous question were minimized by Friend B's choice to test each iPod by playing the same song over and over? Provide reasons for your answers (or some reasons for using the same song over and over).

7. Which, if any, of the problems your group listed in question #5 were minimized by Friend B's choice to test multiple iPods of each color? Provide reasons for your answers (or some reasons for testing multiple iPods of each color).

8. (a) In order to have a valid and convincing claim derived from evidence gathered from an experiment, why must a person identify ALL possible conditions that could change the outcome of the experiment and then only vary ONE condition?
9. Looking at Friend B’s experiment, if she had only done her first run of data (including Silver iPod D), would she have made the same conclusion that Friend A made? Support your answer with numeric evidence.

10. Looking at Friend B's experiment, if she had only done her second run of data (including Silver iPod D), would she have made the same conclusion that Friend A made? Support your answer with numeric evidence.

11. What information became apparent after Friend B performed 3 trials of the battery length on many different iPods with the same color that was not visible by only doing 1 trial on different iPods with the same color?

12. Could Friend B have eliminated the collected data from the silver D iPod by only doing one trial of each iPod? Explain the reasons for your answer.

13. Why did Friend B include the data from the Silver D iPod on her final data table even though she chose not to use that data in calculating the average play time of silver iPods?

14. When designing an experiment, why would performing multiple trials at each condition be a wise choice?

15. Friend B concluded that the playing times of the two different colored iPods were essentially the same (only varied by 5 min on average). Do you agree or disagree with her claim? Explain your reasoning.

16. If a different experiment were performed and the researcher found that two different conditions produced playing times which also varied by 5 minutes, but the average playing times were 18 min and 22 minutes respectively, would you make the same conclusion? Why or why not?

17. When making a decision to accept or reject a claim based upon experimental evidence, does the numeric difference in outcomes provide enough information? Why or why not?
Read This!

- Scientists call the factor or condition you independently choose to explore in an experiment the **independent variable**.
- The result which depends upon the factor being investigated is called the **dependent variable**.
- The conditions that must be held constant are called the **control** variables.
- Data points which are significantly greater than or less than most of the collected data (like the minutes played by Silver iPod D) are called **outliers**. In advanced research projects, scientists must use various types of analysis to determine if the outliers should or should not be included in the set of data used in making the calculations to construct their claim.

18. Look back at Model 1. What was the independent variable under investigation? What was the dependent variable?

19. A farmer wanted to determine which fertilizer would be best to use on his field of sweat corn. He purchased 5 different brands of fertilizer and started to design an experiment in which he was going to measure how high corn plants grew after 4 weeks of being planted in buckets filled with soil from his field. What would be the independent variable in this experiment? What would be the dependent variable? What would be some control variables? Explain your reasoning.

20. A student proposed the following question for the class to investigate: Does the temperature of a tennis ball affect how high the ball will bounce?
   a) What would be the independent variable, and what would be the dependent variable?
   
   b) Identify some control variables the students would need to keep constant to ensure that any differences in bounce heights were only caused by the different temperatures of the tennis balls.
   
   c) Examine a possible table of data. Identify if there is an outlier and explain your reasoning.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>10° C</td>
<td>15.3 cm</td>
<td>14.2 cm</td>
<td>16.1 cm</td>
</tr>
<tr>
<td>15° C</td>
<td>14.5 cm</td>
<td>13.8 cm</td>
<td>17.9 cm</td>
</tr>
<tr>
<td>20° C</td>
<td>15.8 cm</td>
<td>7.2 cm</td>
<td>14.8 cm</td>
</tr>
<tr>
<td>25° C</td>
<td>14.6 cm</td>
<td>16.2 cm</td>
<td>15.7 cm</td>
</tr>
</tbody>
</table>

   d) Write a claim which can be made from the set of data from part “c”.

©HSPI – The POGIL Project  Limited Use by Permission Only – Not for Distribution  Designing an Experiment CIYvM
Extension Questions

21. A researcher was exploring the impact of drinking diet sodas sweetened with aspartame and short-term memory function.

   a) Create a beginning question, design a potential experiment to help address that question, identify the independent variable, the dependent variable, and the control variables.

   b) Since this research project involves humans, what types of precautions do you think this researcher would need to take in order to perform such an experiment?

   c) Investigate what precautions are required by your school, state government, or federal government in order to perform experiments on plants, animals, or humans.

22. Using only a penny, water, and a disposable pipette (AKA plastic medicine dropper), perform an experiment related to how many drops of water can fit on one penny.

   ◦ Create a beginning question.
   ◦ Design an experiment to address that question.
   ◦ Identify the independent, dependent, and control variables.
   ◦ Construct a data table.
   ◦ Decide how many trials should be run in order to make a convincing claim.
   ◦ Perform the actual experiment.
   ◦ Analyze your data.
   ◦ Write a claim related to your beginning question.
Teacher Guide

Learning Objectives:
Students will be able to ...
1. Identify variables which will potentially impact the outcome of an experiment and design an experiment which only changes ONE of those variables while keeping all other variables constant.
2. Construct an experimental procedure which utilizes multiple trials of specific conditions and analyze the gathered data to identify potential statistical outliers.
3. Evaluate whether the non-identical results of various trials are truly “different” by considering the magnitude of the measurements taken.

Prerequisites:
• Able to calculate averages of sets of numbers.

Assessment Questions:
1. A student wanted to explore the impact of temperature on how long a shaken can of soda would stream into the air when opened. Identify at least 4 variables which would need to be held constant when designing a procedure for this experiment.
   ANS: speed of can being opened, angle of can when opened, wind on the stream of soda, type of soda being opened, amount of time can shaken prior to opening, height of can at time of opening, opening technique consistent (no part of hand or other item obstructing stream flow), etc...
2. A student performed an experiment on the impact of the concentration of salt in soil on the growth of a wild sunflower plant. She collected the following data.

<table>
<thead>
<tr>
<th>Trial</th>
<th>0.2% NaCl</th>
<th>0.4% NaCl</th>
<th>0.8% NaCl</th>
<th>1.6% NaCl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>15.3 cm</td>
<td>14.5 cm</td>
<td>15.1 cm</td>
<td>9.2 cm</td>
</tr>
<tr>
<td>Trial 2</td>
<td>14.2 cm</td>
<td>12.8 cm</td>
<td>7.2 cm</td>
<td>10.8 cm</td>
</tr>
<tr>
<td>Trial 3</td>
<td>18.1 cm</td>
<td>17.9 cm</td>
<td>14.8 cm</td>
<td>8.9 cm</td>
</tr>
</tbody>
</table>

Which of the following data points might she consider as an outlier:
   (a) 0.4% NaCl   Trial 1
   (b) 0.2% NaCl   Trial 2
   (c) 1.6% NaCl   Trial 3
   (d) 0.8% NaCl   Trial 2
   ANS: D because the 7.2 cm value is almost half of what the other 2 trials at that condition were.
3. If the following results were used in a claim made by a researcher, which result would you consider to be most believable or able to be defended?
   (a) Average A = 345 cm; Average B = 355 cm; Experimental condition B produced greater distances.
   (b) Average A = 34.7 cm; Average B = 35.2 cm; Experimental conditions A and B were essentially the same.
   (c) Average A = 365 cm; Average B = 355 cm; Experimental condition B produced greater distances.
   (d) Average A = 34.5 cm; Average B = 365 cm; Experimental conditions A and B were essentially the same.
   ANS: B because the two values differ by .5 cm out of an average distance of 35.0 cm (only a 1.4% difference)

Teacher Tips:
• This activity is designed to get students thinking about what things they need to consider when trying to design their own experiment. Try to give students many opportunities to practice both designing their own experiment (like question number 22 in the Extension Question section) as well as analyzing other students' experimental results in order to develop their skill in experimental design.
Target Responses

1. Write down some descriptive words (NO MORE than 8) which summarize your group's first impressions of each of the two experiments.

<table>
<thead>
<tr>
<th>Friend A's Experiment</th>
<th>Friend B's Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overly simple</td>
<td>Well thought out</td>
</tr>
<tr>
<td>Not reliable</td>
<td>Clear</td>
</tr>
</tbody>
</table>

2. How many iPods were tested by Friend A? 2
   How many iPods were tested by Friend B? 8

3. How many times did Friend A test each iPod? 1 time
   How many times did Friend B test each iPod? 3x

4. a. By what process did Friend A “measure” the battery life of the iPod?
   Charging each iPod and then just letting each iPod run out of power through normal use
   b. By what process did Friend B “measure” the battery life of the iPod?
   Charging each iPod then having them play the same song over and over again until their power runs out.
   c. Which of the two methods for “measuring” the battery life of the iPod provides the most reliable measure of the battery life? Explain your reasoning.
   Playing the same song is more reliable because Friend A’s procedure does not provide convincing evidence that the color of the iPod was the reason the battery lasted longer. In Friend A’s experiment, there were many possible reasons (other than the color of the iPod) why the red iPod lasted longer than the silver iPod.

5. Take 1 minute to silently create a list of possible reasons why the red iPod in Friend A’s experiment played music longer than his silver iPod. After the 1 minute of silent thinking time, work as a group to expand your list, considering as many reasons as possible.
   The iPods were playing different types of songs, some iPods were newer than others, one could have been playing a movie or a game, the back lighting settings could have been different, etc...

6. Which, if any, of the problems your group listed in the previous question were minimized by Friend B’s choice to test each iPod by playing the same song over and over? Provide reasons for your answers (or some reasons for using the same song over and over).
   Different songs use different amount of battery energy – by having each iPod play the same song, you are more likely to see if color had any thing to due with the battery life.

7. Which, if any, of the problems your group listed in question #5 were minimized by Friend B’s choice to test multiple iPods of each color? Provide reasons for your answers (or some reasons for testing multiple iPods of each color).
   Some iPods were newer than others; by having more than one iPod of each color, one can minimize the factor of iPod age on the length of playing time, etc...

8. a. In order to have a valid and convincing claim derived from evidence gathered from an experiment, why must a person identify ALL possible conditions that could change the outcome of the experiment and then only vary ONE condition?
   Before designing an experiment, one should identify ALL possible conditions or variables which could affect the results of the question being explored. Then one can design an experiment that will only change ONE condition. Any changes in the outcome of the experiment would be attributable to that one changed condition.
9. Looking at Friend B's experiment, if she had only done her first run of data (including Silver iPod D), would she have made the same conclusion that Friend A made? Support your answer with numeric evidence.

Yes, when you find the average of each colored iPod in Run 1, the red iPods on average seem to last longer (66min)

\[ \text{iPods}_{\text{red}} = \frac{(855\text{min} + 858\text{min} + 890\text{min})}{3 \text{ runs}} = 868\text{min}_{\text{average}} \]

\[ \text{iPods}_{\text{silver}} = \frac{(788\text{min} + 906\text{min} + 808\text{min} + 640\text{min} + 866\text{min})}{5 \text{ runs}} = 802\text{min}_{\text{average}} \]

10. Looking at Friend B's experiment, if she had only done her second run of data (including Silver iPod D), would she have made the same conclusion that Friend A made? Support your answer with numeric evidence.

No, when you find the average of each colored iPod in Run 2, the silver iPods on average seem to last longer (4min).

This is close to the final conclusion Friend B made since 4min is essentially the same time of playing.

\[ \text{iPods}_{\text{red}} = \frac{(823\text{min} + 873\text{min} + 799\text{min})}{3 \text{ runs}} = 832\text{min}_{\text{average}} \]

\[ \text{iPods}_{\text{silver}} = \frac{(815\text{min} + 870\text{min} + 902\text{min} + 760\text{min} + 832\text{min})}{5 \text{ runs}} = 836\text{min}_{\text{average}} \]

11. What information became apparent after Friend B performed 3 trials of the battery length on many different iPods with the same color that was not visible by only doing 1 trial on different iPods with the same color?

The data for iPod Silver D was significantly different than the other iPods and could possibly be excluded when making the final calculations about average times. (NOTE: The teacher may want to mention that there are more sophisticated methods for dealing with outliers, but the goal of this activity was to introduce students to critically analyzing their own data to determine if something might have happened to render a set of data misleading for their research question).

12. Could Friend B have eliminated the collected data from the silver D iPod by only doing one trial of each iPod? Explain the reasons for your answer.

No, not really because only after doing 3 trials do you see that Silver D is consistently out of range of the other iPods.

13. Why did Friend B include the data from the Silver D iPod on her final data table even though she chose not to use that data in calculating the average play time of silver iPods?

She included the data because it shows what she really recorded when doing the experiment. She felt justified in excluding the Silver D data and by including the data in the data table she provides enough information for her audience to follow her thinking and to re-calculate the averages using the excluded data to see for themselves if her decision not to include the Silver D iPod was a reasonable decision.

14. When designing an experiment, why would performing multiple trials at each condition be a wise choice?

When you perform multiple trials at each condition, you are gathering enough data to better identify if any point is an outlier, also insuring that your results were truly due to the independent variable and not an artifact of your technique.
15. Friend B concluded that the playing times of the two different colored iPods were essentially the same (only varied by 5 min on average). Do you agree or disagree with her claim? Explain your reasoning.

Yes, I agree with her reasoning because 5 minutes is only about 0.6% which is essentially the same value in terms of playing time.

16. If a different experiment were performed and the researcher found that two different conditions produced playing times which also varied by 5 minutes, but the average playing times were 18 min and 22 minutes respectively, would you make the same conclusion? Why or why not?

No, I would conclude that there is a significant difference as a result of the experimental conditions, because now the 5 minutes represents a 25% difference.

17. When making a decision to accept or reject a claim based upon experimental evidence, does the numeric difference in outcomes provide enough information? Why or why not?

No, you also need to know on average what is the magnitude of the measurement being made so you can judge if the difference is important or not. 4 minutes wasn't a big deal when looking at times in the 800 minute range but was a big deal when looking at times on the order of 20 minutes.

18. Look back at Model 1. What was the independent variable under investigation? What was the dependent variable?

Independent Variable: color of the iPod
Dependent Variable: playback time

19. A farmer wanted to determine which fertilizer would be best to use on his field of sweat corn. He purchased 5 different brands of fertilizer and started to design an experiment in which he was going to measure how high corn plants grew after 4 weeks of being planted in buckets filled with soil from his field. What would be the independent variable in this experiment? What would be the dependent variable? What would be some control variables? Explain your reasoning.

Independent Variable: type of fertilizer
Dependent Variable: height of corn plants after 4 weeks
Control variables: same soil, same amount of water, same lighting, same depth seed was planted......

Since the farmer CHOSE to vary the fertilizer, that becomes the independent variable. He will check how the height which need to be the same (so that the difference in height is due only to the fertilizer) will be the control variables.

20. A student proposed the following question for the class to investigate: Does the temperature of a tennis ball affect how high the ball will bounce?

a) What would be the independent variable, and what would be the dependent variable?

Independent Variable: temperature of the tennis ball
Dependent Variable: height ball bounces

b) Identify some control variables the students would need to keep constant to ensure that any differences in bounce heights were only caused by the different temperatures of the tennis balls.

Starting point of ball, wind surrounding dropped ball, even release of ball each time, etc...

c) Examine a possible table of data. Identify if there is an outlier and explain your reasoning.

<table>
<thead>
<tr>
<th></th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>10° C</td>
<td>15.3 cm</td>
<td>14.2 cm</td>
<td>16.1 cm</td>
</tr>
<tr>
<td>15° C</td>
<td>14.5 cm</td>
<td>13.8 cm</td>
<td>17.9 cm</td>
</tr>
<tr>
<td>20° C</td>
<td>15.8 cm</td>
<td>7.2 cm</td>
<td>14.8 cm</td>
</tr>
<tr>
<td>25° C</td>
<td>14.6 cm</td>
<td>16.2 cm</td>
<td>15.7 cm</td>
</tr>
</tbody>
</table>

To determine if a set of data has an outlier, you should compare the data collected over multiple runs at the same temperature (going horizontally). The heights for 10° C (15.3 cm, 14.2 cm, and 16.1 cm), 15° C (14.5 cm, 13.8 cm, and 17.9 cm), and 25° C (14.6 cm, 16.2 cm, and 15.7 cm) appear to be similar to each other. However, the heights as 20° C (15.8 cm, 7.2 cm, and 14.8 cm) have a significant difference for the height of trial 2, 7.2 cm, which is almost half the height of the other data for that temperature. So 20° C, trial 2 could be considered a potential outlier.
d) Write a claim which can be made from the set of data from part “c”.
   If you calculate the averages for each temperature, 10°C (15.2 cm average),
   15°C (15.4 cm average), 20°C (15.3 cm average NOT using the 7.2 cm data point),
   and 25°C (15.5 cm average), there does not seem to be a significant difference
   within these temp ranges. Claim: “Between the temperature range of 10°C and 25°C,
   changing the temperature of the tennis ball does not appear to affect the height the ball will bounce.”

Extension Questions

21. A researcher was exploring the impact of drinking diet sodas sweetened with aspartame and short-term memory function.
   a) Create a beginning question, design a potential experiment to help address that question, identify the
      independent variable, the dependent variable, and the control variables.
      Answers will vary.

   b) Since this research project involves humans, what types of precautions do you think this researcher
      would need to take in order to perform such an experiment?
      Answers will vary.

   c) Investigate what precautions are required by your school, state government, or federal government
      in order to perform experiments on plants, animals, or humans.
      Answers will vary.

22. Using only a penny, water, and a disposable pipette (AKA plastic medicine dropper), perform an
    experiment related to how many drops of water can fit on one penny.

   ○ Create a beginning question.
   ○ Design an experiment to address that question.
   ○ Identify the independent, dependent, and control variables.
   ○ Construct a data table.
   ○ Decide how many trials should be run in order to make a convincing claim.
   ○ Perform the actual experiment.
   ○ Analyze your data.
   ○ Write a claim related to your beginning question.

   Answers will vary.