ACTIVITY 6

Quantifying Scientific Uncertainty

COMPUTER SIMULATION

ACTIVITY 6

Quantifying Scientific Uncertainty

ACTIVITY SUMMARY

Students use a computer simulation to gather data from an air quality sensor under different conditions. By analyzing the data generated by the simulation, students further explore random errors and systematic errors. They learn how collecting more data and averaging can affect the accuracy of random errors but not systematic errors. Students consider how error bars can communicate scientific uncertainty in data. They use this information to consider a decision between two air sensors.

ACTIVITY TYPE
COMPUTER SIMULATION

NUMBER OF 40–50 MINUTE CLASS PERIODS
1-2

KEY CONCEPTS & PROCESS SKILLS

1 Uncertainty in data is often a result of errors. Scientific errors can be random or systematic and can lead to conclusions that are less likely to be correct.

2 Scientific methods can reduce sources of uncertainty. Techniques to reduce random error include taking repeated measurements and averaging across many samples. Techniques to reduce systematic errors include calibrating equipment more carefully and designing investigations to control for other factors that could influence the results *(confounds)*.

3 Confidence intervals, confidence levels, and error bars describe the uncertainty of data and the probability that data are accurate.

4 Probabilistic reasoning can be used to identify meaningful patterns in data *(signal)* about a phenomenon being investigated. Variations in the data *(noise)* can increase scientific uncertainty by distorting or hiding the signal.

CONCEPTUAL TOOLS

NEXT GENERATION SCIENCE STANDARDS (NGSS) CONNECTION:
Analyze data using tools, technologies, and/or models (e.g., computational, mathematical) in order to make valid and reliable scientific claims or determine an optimal design solution. *(Science and Engineering Practice: Analyzing and Interpreting Data)*

Consider limitations of data analysis (e.g., measurement error, sample selection) when analyzing and interpreting data. *(Science and Engineering Practice: Analyzing and Interpreting Data)*

VOCABULARY DEVELOPMENT

accuracy
(assumed prior knowledge)
the closeness of a measured value to a standard or true value

error bars
a visual representation of the amount of uncertainty in a measurement, presented as a ± (plus or minus) range of numbers that fall above and below the average

precision
(assumed prior knowledge)
how close measurements of the same item are to each other

reliable
(assumed prior knowledge)
able to be reproduced consistently

validation
(assumed prior knowledge)
a process of determining the accuracy of a measurement

TEACHER BACKGROUND INFORMATION

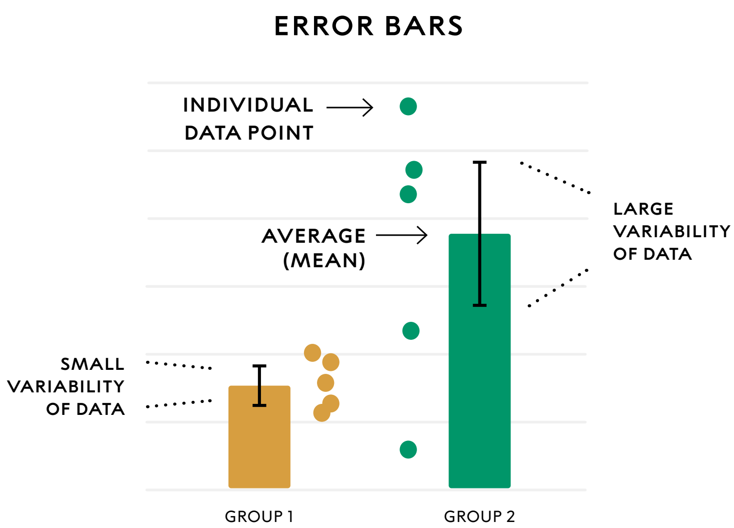
Two Graph Types Used in Computer Simulation

In the Scientific Uncertainty computer simulation used in this activity, Graph 1, "Average of All Trials," is a jitter plot. It shows data points for all trials plotted in the same place. The data points are NOT plotted over time. In order to be able to see all the data points, they are spread horizontally (slightly) so they do not overlap. This graph makes it easier to see the distribution pattern of the data points in order to determine how close the points are to one another and, therefore, how precise and reliable the photometer readings are.

Graph 2, "Current Average," also known as the running average, shows how the average of the trials changes with more trials. It shows that when there are few data points, the average has large shifts with each new measurement. As more data points are added, the average gradually shifts less with each new point and becomes more stable. The average then gets closer to the true value as more trials are added.

Error Bars: Estimating the Scientific Uncertainty

Error bars represent the variability of a data set and are drawn as lines that go above and below a point on a graph. For an average, the size of the error bars around a point is related to a measure of the variability in the data set in relation to the average and can depend on the number of measurements or samples used. The larger the error bars, the more variability and uncertainty present in the data. Although error bars can represent various statistical measures of variability (standard deviation, standard error, etc.), this computer simulation uses 95% confidence intervals to determine the error bars. Look at the following diagram, which represents error bars on a bar graph; the simulation uses line graphs. For a repeated measurement, the error bars correspond to the range of values in which you can be 95% confident of the true value falling within. Thinking about variability in data is one factor that scientists look at when considering how accurate (close to the true value) their data is, which can affect their confidence in conclusions they can make from their data.



Scientists calculate the variability of repeated measurements based on how spread apart the data points are from one another. This value can then be used to produce error bars that can represent different measures of variability (e.g., standard deviation, standard error or the mean, or 95% confidence intervals). Regardless of which measure is being used, large error bars generally correspond to more variability in the data and less certainty that the average is an accurate representation of the true value, whereas smaller error bars indicate the opposite. Note: This unit does not use the term *variability* to describe data patterns; instead, variability describes how spread apart the data points are.

MATERIALS & ADVANCE PREPARATION

FOR THE TEACHER

- VISUAL AID 6.1 "Comparing Photometer 1 and Photometer 2"

FOR EACH PAIR OF STUDENTS

- COMPUTER WITH INTERNET ACCESS

FOR EACH STUDENT

- STUDENT SHEET 6.1 "Testing Air Quality Sensors"

Arrange for classroom computer use and familiarize yourself with the Scientific Uncertainty Simulation.

TEACHING NOTES

Suggestions for **discussion questions** are highlighted in gold.

Strategies for the **equitable inclusion of diverse students** are highlighted in mint.

GETTING STARTED (5-10 MIN)

1 **Review random error and systematic error.**

- In this activity, students will investigate the effect of random error and systematic error on data. To ensure that students are prepared, ask, Which of the following is an example of a random error or a systematic error?
  - A tape measure goes through the clothes dryer and shrinks.

    This would cause a systematic error, resulting in measurements that are consistently shorter than the true value.
  - Phone polls rely on people who answer their phones when unknown numbers call them.

    This could cause a systematic error because the results may be skewed based on some common characteristic of the people who regularly answer calls from unknown numbers, such as age.
  - After asking three classmates how much sleep they got the night before, Gavin averages the results and concludes that all his classmates sleep only five hours a night.

    This situation involves random error because each person's sleep time might be different from the next, with some above the average and some below the average.
- If students are not familiar with the basics of reliability, accuracy, precision, and validation, you may want to review these concepts by referring them to the Science Review in the Student Book.

PROCEDURE SUPPORT (25 MIN)

2 Review the scenario and the goal of the computer simulation.

- The scenario presented in Part A: Testing a Photometer, Procedure Step 1 can be shared with the class in multiple ways. Read the scenario aloud to the class or have individual students read it aloud while others follow along with the text (either as a whole class or in small groups). Reading the scenario aloud can better support comprehension for many students, including neurodiverse students and emerging multilingual learners who often have more highly developed listening and oral skills than reading comprehension skills. Alternatively, students can read the scenario independently.
- Review how the Blase-Air sensor works (shown in Figure 6.1 in the Student Book). Note that this diagram is not representative of how all air sensors work; it's representative of the specific version in the simulation.
- Before beginning the simulation, introduce the online Scientific Uncertainty Simulation as a thinking tool that can help to model how different types of scientific errors can affect experimental data. Emphasize that this simulation is a concrete tool that supports concepts related to random errors and systematic errors, along with techniques for reducing the effects of these types of errors.
- You may wish to give students time to explore the features of the simulation and/or model how to use the simulation with the class to answer any questions they may have. Point out that there is no labeled *x*-axis for Graph 1 in Experiments 1 and 2 or for Graphs 1 and 2 in Experiment 3 because the data is not represented over time. (The data points are shown vertically and are only separated horizontally for better visibility.)

TEACHER'S NOTE: Data in the simulation has natural variability, so each computer running the simulation will get slightly different results. The true value for each computer remains the same from Experiment 1 to Experiment 2; it is different in Experiment 3. Additionally, the values are changed every time the website is refreshed, so students can repeat the experiments with different air samples, if they wish.

3 Support students' conceptual development as they use the simulation.

- Distribute one copy of Student Sheet 6.1, "Testing Air Quality Sensors," to each student. Remind students to record their data on the student sheet as they complete each experiment. You may find it helpful to model Procedure Steps 3 and 4 by completing Experiment 1 on a projected version of the student sheet.
- Be sure to explain what students will record in the second row of Tables 1, 2, and 3: students should continue to run trials until their Error Bar Range is ± 3 and then record their data. Point out that the number of trials run in the second row will vary. Since the data in the simulation may be slightly different due to variability in the simulation, make sure students are recording the data from their own simulation in preparation for Experiment 2.

- Circulate and assist pairs as they work through the simulation. The simulation has three experiments that students progress through, which are described in the information that follows.

TEACHER'S NOTE: The following information is provided to support your instruction. Do not share this information with students ahead of time.

Experiment 1: Investigate how well the original photometer (Photometer 1) measures PM2.5 levels.

This experiment shows that a random error causes data points to be scattered above and below the true value with no consistent pattern. Students should also see that conducting more trials can help reduce how much the average changes from one trial to the next, bringing the average closer to the true value. Students should notice that the error bars start out large and get smaller as you add more data.

Experiment 2: Investigate whether an appliance running next to Photometer 1 affects the data from Photometer 1 and, if so, how.

This experiment shows that adding a systematic error (in this case, caused by turning on the appliance) can shift the measurements in a consistent direction away from the true value. However, this shift does not go away even after adding more trials. Also, the shift doesn't affect how spread out the data points are from one another or the size of the error bars. The only way to correct it would be to get rid of the systematic error. This is accomplished in the simulation by comparing measurements from when the appliance is off in Experiment 1 to when it is on in Experiment 2.

Experiment 3: Compare data from Photometer 1 to the new photometer (Photometer 2).

This experiment helps students practice identifying how the distribution of the data points is related to error bars—students practice identifying whether the data is showing random errors and/or systematic errors. Students evaluate the claim that one of the photometers is more accurate than the other. They find that by looking through the data, both photometers have a similar accuracy, but one of them has more data variability. In terms of decision-making, the photometer with less variability in the data is a more reliable product.

- In Procedure Steps 4 and 5, consider reviewing the graphs as a class to support students' interpretation of the data. Before students discuss the questions with their partners, review Figure 6.2 in the Student Book to help students understand how Graph 1 and Graph 2 provide different information about the experiment and to clarify how to identify the true value, average, and error bars. See the Teacher Background Information for more information on graph types created by this simulation.

4 Review the results of each experiment as a class.

- Depending on your student population, you may want to stop the class after pairs complete each experiment to discuss the results and clarify key points. Alternatively, you can wait until all pairs have completed all three experiments before discussion. Sample Student Responses for discussion questions in the procedure are as follows:

Sample Student Responses, Procedure Steps 5, 7, 9, 10, 12, and 13

<table>
<tr><th>PROCEDURE STEP</th><th>SAMPLE STUDENT RESPONSES</th></tr>
<tr><td colspan="2">PART A: TESTING A PHOTOMETER</td></tr>
<tr><td>Step 5</td><td>Compare the data in Graphs 1 and 2 by discussing the following with your partner:
• your observations of each graph with your partner.
Graph 1 is easier to see how close the data pofints are to each other because you can see all of the data points at once. Graph 2 is easier to see how the average line gets closer to the true value as you conduct more trials.
• what Graphs 1 and 2 might look like if there were a) no error, b) random error, and c) systematic error.
a) no error
Graph 1: All the data points would be the same, so there would only be one spot.
Graph 2: The measurements would be the same for each trial, so the average line would be flat.
b) random error
Graph 1: The data points are spread randomly above and below the true value.
Graph 2: The average line fluctuates but slowly approaches the true value as you get more data.
c) systematic error
Graph 1: The data points are shifted in a consistent direction away from the true value.
Graph 2: The average line is shifted in a consistent direction away from the true value.</td></tr>
<tr><td>Step 7</td><td>Use the results of Experiment 1 to discuss the following with your partner:
• how the number of trials affected the true value, the average, and the error bars.
The number of trials didn't affect the true value. The average fluctuated a lot when there were fewer trials and only a little each time you conducted a new trial. The error bar range started big and got smaller as you conducted more trials.
• whether conducting more trials makes your average more accurate (closer to the true value) and why or why not.
Conducting more trials made the average more accurate because the average came closer to the true value.</td></tr>
<tr><td>Step 9</td><td>With your partner, determine whether the electric field of the appliance affected the measurements by:
a) comparing the size of the error bars for Experiment 1 (no appliance on) and Experiment 2 (appliance on).
The size of the error bars was about the same size by the end of Experiments 1 and 2.
b) comparing the averages for both Experiment 1 (no appliance on) and Experiment 2 (appliance on) to the true value.
The average for Experiment 1 was close to the true value. The average for Experiment 2 was shifted lower than the true value.</td></tr>
</table>

PROCEDURE STEP	SAMPLE STUDENT RESPONSES
PART A: TESTING A PHOTOMETER	
Step 10	**Discuss with your partner what type(s) of error (systematic, random, or both) are likely to be occurring in Experiments 1 and 2. Be sure to explain your reasoning.** *Experiment 1 had random error because the data points varied both above and below the true value without a consistent pattern.* *Experiment 2 had systematic error because the data points were shifted in a consistent direction (lower than the true value). It also had random error because the data points were still spread higher or lower randomly, even though they were all shifted lower.*
PART B: COMPARING TWO PHOTOMETERS	
Step 12	**Uses the data from Experiment 3 to determine which photometer:** *Responses may vary slightly due to variability in the simulation.* a) has an average measurement that is most accurate (closest to the true value). *The photometers had similar accuracy. Photometer 1 average was 495.9, and Photometer 2 average was 494.4. The true value was 495, so Photometer 2 is slightly more accurate.* b) had less scientific uncertainty in the data. `HINT: for each photometer, consider the size of the error bars as well as how spread out the data points are.` *Photometer 2 resulted in less scientific uncertainty in the data because the data points are much closer together than with Photometer 1. Also, the error bars for Photometer 2 were smaller within a fewer number of trials.* c) required fewer trials for the average to reach within ± 3 particulates of the true value. *Photometer 1 required 154 trials to get within ± 3 particulates.* *Photometer 2 required 24 trials to get within ± 3 particulates.*
Step 13	**As a class, evaluate the claim that the new Photometer 2 is more accurate than the old Photometer 1. Decide which photometer Blase-Air should use in their air quality sensor design.** *The claim that the new Photometer 2 is more accurate is not supported by the data in the simulation because the two photometers have very similar averages that are not that different from the true value.* (See Teaching Step 5 for the class decision regarding which photometer to recommend for Blase-Air's sensor.)

SYNTHESIS OF IDEAS (20 MIN)

5 **Review key ideas investigated with the simulation.**

- After completing Part B, give pairs a few minutes to discuss Procedure Step 13 before sharing ideas with the class. Project Visual Aid 6.1, "Comparing Photometer 1 and Photometer 2," to reference an example of the data collected from the two photometers in Experiment 3. The data that students gathered may be slightly different from the example, but the patterns will be similar. Ask, Should the company use the sensor with Photometer 1 or Photometer 2 in their sensor and why? Students should recognize that averages of the two photometers are almost the same, but Photometer 2 is more likely to result in data points that are closer together, and it needs less trials to get an average that is close to the true value. This means that Photometer 2 is better at reliably measuring the amount of particulates in the air and would be a better choice for Blase-Air's sensor.

- Use Visual Aid 6.1 to help students understand the relationship between scientific uncertainty and error bars. Ask, What is the relationship between the size of the error bars and the amount of scientific uncertainty in the data? Students should explain that the error bars correspond to how spread out the data points are from one another. Smaller error bars will indicate more certainty in the current average because the data points are closer together. Larger error bars will indicate the opposite. If you have begun a word wall, support students, particularly emerging multilingual learners, in sensemaking and language acquisition by adding the term *error bar*.

- Follow up the activity by asking, Based on the data for the two photometers, which photometer has less uncertainty in the data? Have students describe how this affects their decision and how sure they are about their choice. Student responses should indicate that Photometer 2 has less uncertainty in the data and smaller error bars, which makes students more certain about their choice. Alternatively, students might say that Photometer 1 has more uncertainty and larger error bars, which makes them more certain about their choice. As a result of this discussion, you may want to give students a moment to revise their responses to Build Understanding item 2.

- Evaluate if your students are able to identify the essential ideas of the activity by discussing their responses to Build Understanding items 1 and 4, which focus on how systematic error and random error can be identified in data, as well as how averaging multiple data points affects the different types of errors. Make sure students understand that while scientists often collect multiple measurements and average the data points, this method only helps to reduce uncertainty from random errors, not systematic errors. This is because random errors produce data points that can be higher or lower by chance, so averaging tends to cancel out their effects. The only way to reduce uncertainty from systematic error is to find and remove the source of the error or to redesign the experiment to better account for the error.

- Build Understanding item 3 is an opportunity to revisit the concepts of signal and noise from earlier in the unit. Emphasize that random errors and systematic errors can serve as a source of noise that can obscure the signal being investigated.

- Finish the activity by revisiting the Guiding Question, *How can you reduce random errors and systematic errors in data?* Use responses to this question to formatively assess the key concepts and process skills related to identifying and reducing random errors and systematic errors.

SAMPLE STUDENT RESPONSES

BUILD UNDERSTANDING

1. **Consider how the data changed during the experiments in the computer simulation. How does conducting more trials affect how close the average is to the true value when there is:**

   a random error?

   When there is random error, the data points can randomly be higher or lower than the true value. Conducting more trials makes the average change less with each trial, and the average also becomes more accurate (gets closer to the true value).

   b systematic error?

   When there is systematic error, the average is shifted in one direction away from the true value. Conducting more trials does not make the average more accurate (does not shift it closer to the true value).

2. **Error bars play an important role in communicating about scientific uncertainty in data. Describe how the size of the error bars can affect scientific uncertainty in a conclusion. Include an example from the computer simulation in your response.**

   The size of the error bars is related to how spread out the data points are from each other, and it is mostly related to the amount of random error in the sample. If the error bars are very small, it means that there is little random error in the sample, and there can be more certainty in the conclusion than if the error bars were very large. For example, in Experiment 3 of the simulation, Photometer 1 had very large error bars, and Photometer 2 had very small error bars even though they both had an average that was close to the true value. This tells me that there is more uncertainty in Photometer 1's data and maybe it is not as reliable as Photometer 2, which has very similar measurements across all the samples.

3. **Think back to the concepts of signal and noise from Activity 2: Signal and Noise. Explain how the presentation of data in the computer simulation helped differentiate the signal from the noise.**

   The graphs in the simulation help us see the signal in the data by showing us the average of the data. For example, in the simulation, photometer data points are distributed all over the place if you look at Graph 1, and it is hard to see trends in the data. The average line gives you a more useful signal because it gives you a single value to use to make comparisons. If you look at Graph 2, you can see how the average changes and how many trials the photometer requires to reach an average that is close to the true value.

④ **Figure 6.3 shows air quality data from five sensors in the same location. Each line represents a different sensor.**

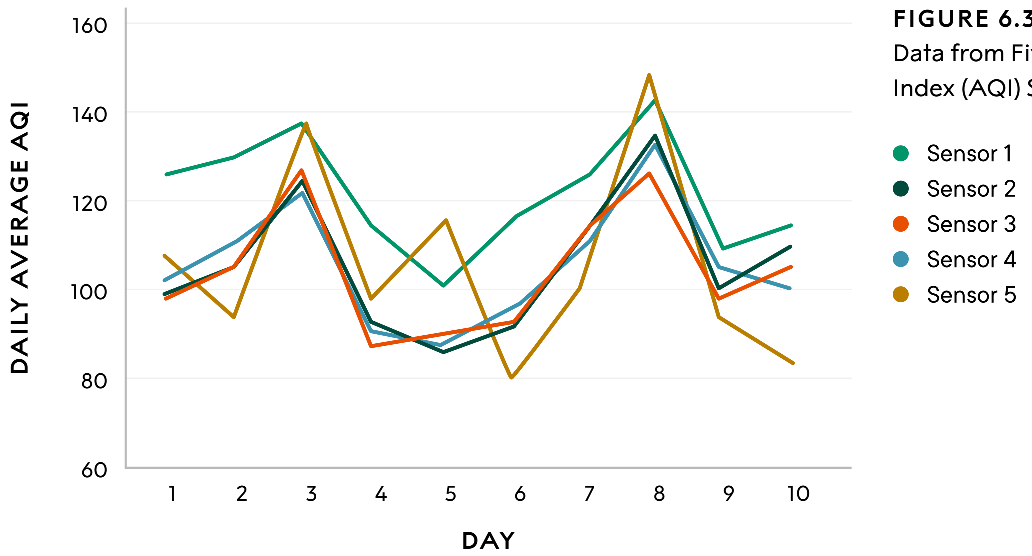


FIGURE 6.3
Data from Five Air Quality Index (AQI) Sensors

a Which sensor is most likely to have a systematic error? Explain your reasoning.

Sensor 1 because it is consistently higher than the other sensors.

b Which sensor appears to have random error? Explain your reasoning.

Sensor 5 because it fluctuates a lot and is most different from the other sensors, but not consistently below or above them.

CONNECTIONS TO EVERYDAY LIFE

⑤ **Glucose is a type of sugar that your body uses for energy. Imagine that you read about a study comparing two different glucose meters for measuring glucose levels in the blood of patients with diabetes. Each glucose meter measured the same blood sample 20 times (N = 20). Figure 6.4 shows glucose level averages, shown with error bars, that were calculated for each glucose meter.**

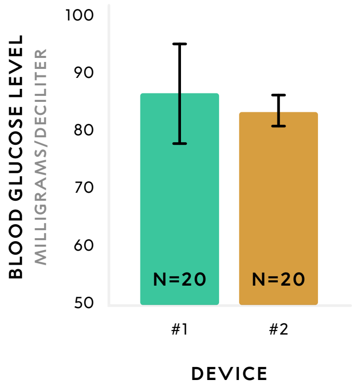
a What can you determine about the uncertainty of the data collected by each device?

You can tell that #1 has more uncertainty in the data than #2, meaning the data points are more spread out in #1. This means that there's more random error in the data for #1, and it also means that with #1, you might not always get a result that is close to the average.

b Can you be sure which device is measuring blood glucose levels closer to the true value? Explain why or why not.

You can't be sure which device is more accurate because you don't know what the true value is. It's hard to tell from this graph whether or not a systematic error is affecting the data, because a systematic error wouldn't be shown by error bars.

FIGURE 6.4
Comparing Glucose Meters



REFERENCES

Koigi, B. (April 22, 2024). *A Ugandan innovation to tackle air pollution*. Fair Planet. https://www.fairplanet.org/editors-pick/african-innovations-to-tackle-air-pollution/

Nguyen, N. , Nguyen, H. , Le, T. & Vu, C. (2021). Evaluating low-cost commercially available sensors for air quality monitoring and application of sensor calibration methods for improving accuracy. *Open Journal of Air Pollution, 10*(1). https://doi.org/10.4236/ojap.2021.101001

U.S. Government Accountability Office. (December 7, 2020). *Science & tech spotlight: Air quality sensors*. https://www.gao.gov/products/gao-21-189sp

Table 1: Experiment 1

PHOTOMETER 1

NUMBER OF TRIALS RUN	AVERAGE	ERROR BAR RANGE	TRUE VALUE
5			
		± 3	

Table 2: Experiment 2

PHOTOMETER 1 (APPLIANCE TURNED ON)

NUMBER OF TRIALS RUN	AVERAGE	ERROR BAR RANGE	TRUE VALUE
5			
		± 3	

Table 3: Experiment 3

PHOTOMETER 1 (OLD MODEL)

NUMBER OF TRIALS RUN	AVERAGE	ERROR BAR RANGE	TRUE VALUE
5			
		± 3	

PHOTOMETER 2 (NEW MODEL)

NUMBER OF TRIALS RUN	AVERAGE	ERROR BAR RANGE	TRUE VALUE
5			
		± 3	

Table 1: Experiment 1

PHOTOMETER 1

NUMBER OF TRIALS RUN	AVERAGE	ERROR BAR RANGE	TRUE VALUE
5	517	± 18.17	503
142	505.6	± 3	503

Table 2: Experiment 2

PHOTOMETER 1 (APPLIANCE TURNED ON)

NUMBER OF TRIALS RUN	AVERAGE	ERROR BAR RANGE	TRUE VALUE
5	471.8	± 12.4	503
179	483.3	± 3	503

Table 3: Experiment 3

PHOTOMETER 1 (OLD MODEL)

NUMBER OF TRIALS RUN	AVERAGE	ERROR BAR RANGE	TRUE VALUE
5	491.2	± 8.8	495
154	495.9	± 3	495

PHOTOMETER 2 (NEW MODEL)

NUMBER OF TRIALS RUN	AVERAGE	ERROR BAR RANGE	TRUE VALUE
5	495.8	± 9.1	495
24	494.4	± 3	495

VISUAL AID 6.1 | COMPARING PHOTOMETER 1 AND PHOTOMETER 2