INSTRUCTIONS FOR EDUCATOR

This ruler drop activity is a customized example of the “Setting Up Your Data for Success” template. Additional resources, including the original template, can be found on this resource’s webpage:

- The “Customizable Student Handout” document is the original template, which you can customize for any investigation. More information on how to complete and implement the customizable handout is provided in the general “Educator Materials” document.
- The “Ruler Drop Student Handout” document is a modified version of the “Customizable Student Handout” for the ruler drop investigation. This “Ruler Drop” educator document provides more information about the “Ruler Drop Student Handout” and how to implement it in class.

Note that many parts of this document are similar or identical to the general “Educator Materials” document. Skip or skim repeated sections as needed.

OVERVIEW

This activity guides students through organizing and documenting spreadsheet data using the ruler drop, a common experiment that measures reaction time. The activity focuses on decisions during data organization and collection that facilitate future graphing and analysis.

The “Ruler Drop Student Handout” document has eight parts, which can be further adjusted based on your class’s needs:

- In Part 0, students brainstorm variables that could affect reaction time, then decide which variables to investigate as a class.
- In Part 1, students begin the investigation by writing the research question and forming a hypothesis.
- In Part 2, they discuss the experimental procedure as a class and design their own data sheet.
- In Part 3, students collect measurements and then enter their data into a spreadsheet.
- In Part 4, they compare their spreadsheets with those of their classmates and reflect on their data organization decisions.
- In Part 5, they learn about some common effective practices for organizing data.
- In Part 6, the class decides on the final organization of their data, so that everyone can combine their data into one data set.
- In Part 7, they document their final spreadsheets by adding metadata.

This educator document contains multiple resources for implementing this activity with students, including the following (use the links to go directly to each section):

- background on the ruler drop
- teaching tips for deciding the structure of the activity
- procedures for implementing each part of the activity in class
- assessment guidance with sample answers for the “Ruler Drop Student Handout” questions

Additional information can be found on this resource’s webpage, including suggested audience, estimated time, and curriculum connections.
KEY CONCEPTS

- Defining a research question and hypothesis guides data collection decisions.
- Data in spreadsheets should be consistently organized (e.g., with clearly labeled columns and rows) to facilitate interpretation and analysis.
- Standard spreadsheet organization practices and metadata make data easier to share and understand.

STUDENT LEARNING TARGETS

- Practice data collection, data entry, and spreadsheet skills.
- Identify and implement effective practices for data organization and documentation.
- Describe the importance of spreadsheet practices that make data understandable and shareable.

PRIOR KNOWLEDGE

Students should be familiar with:

- generating research questions and hypotheses
- collecting and recording variables in an experiment
- entering data (numbers and text) and performing basic functions (calculating averages using a formula and sorting) in the selected spreadsheet tool (e.g., Google Sheets, Microsoft Excel)

BACKGROUND

The ruler drop is an experiment used to measure reaction time. A ruler is dropped from a known height, and reaction time is estimated using the distance the ruler traveled before it is caught (Figure 1).

Figure 1. An illustration of the ruler drop. In Step 1, one person (the dropper) holds a ruler above the hand of another person (the catcher). In Step 2, the dropper drops the ruler, and the catcher must catch it as fast as they can. The “drop distance” between the catcher’s hand and the bottom of the ruler is used to calculate their reaction time.

This test is used to measure reaction time in some clinical settings — for example, in assessing concussion recovery or nervous system function in Parkinson’s patients (Aranha et al. 2017, Del Rossi 2017, Ángel Latorre-Roman et al. 2019). It is also used in classroom explorations (Ortolano, Finn, and Ortolano 2017, Patrick et al. 2017, Bari and Aldosky 2022), often to facilitate discussions about reaction time and nervous system function in general biology and human anatomy and physiology courses.

Students can design ruler drop experiments to explore how reaction time is affected by specific variables — such as handedness, gaming frequency, caffeine intake, or sleep deprivation. Students can also use the ruler drop to evaluate how quickly people react to different types of stimuli, such as auditory vs. visual cues.
MATERIALS

- copies of the “Ruler Drop Student Handout” document
- a ruler or meter stick with standard units (inches, centimeters, millimeters)
  - You can also purchase or print “reaction time rulers” that have drop distance already converted into reaction time.
- blank sheet (paper or online document) for each student group to record data
- access to Google Sheets (or another spreadsheet program of your choice)

TEACHING TIPS

This “Ruler Drop” educator document focuses on information for the ruler drop activity. For more supporting information, including guidance on selecting an appropriate spreadsheet program and ideas for customizing the activity based on time and the comfort of your students with spreadsheet programs, refer to the general “Educator Materials” document on this resource’s webpage.

Choosing Independent Variables

In Part 0 of the activity, students brainstorm potential independent variables that affect reaction time. These variables must be possible to measure and appropriate to collect from all members of the class.

The independent variable could be a **categorical** variable that enables comparisons between multiple groups or conditions. Examples include:

- Class year
- Age; related studies include Aranha et al. 2017 and Ángel Latorre-Roman et al. 2019
- Handedness; related studies include Scharoun and Bryden 2014 and Bari and Aldosky 2022
- Caffeine intake (yes/no)
- Being an athlete (yes/no); related studies include Del Rossi et al. 2014 and Luu et al. 2021
- Cue for releasing the ruler (visual/auditory/both)

Or the independent variable could be a **numerical (continuous)** variable that measures a certain activity or attribute. Examples include:

- Caffeine intake (number of cups of coffee over the past 24 hours)
- Sleep deprivation (number of hours slept over the past week); related studies include Patrick et al. 2017 and Bari and Aldosky 2022
- Exercise (hours over the last week)
- Time since last meal
- Hours spent gaming; related studies include Luu et al. 2021

Encourage students to think about meaningful comparisons and to avoid choosing:

- variables that may require disclosing personal information (i.e., sex, gender, ethnicity, race, etc.)
- variables that could exclude some students from participating (for example, comparing the relationship between reaction time and speed-climbing a set of stairs would exclude those unable to use stairs and those uncomfortable participating)

Options for the Facilitation Approach

This activity is written using a **collaborative approach**, where students work together to generate ideas for the class investigation and are guided by the instructor via class discussion. Alternatively, you could use a **predefined approach**, where you provide students with predetermined variables to measure in the experiment. For example, you could tell students to test how handedness (catching the ruler...
with your dominant vs. nondominant hand) affects reaction time. This approach may be preferable when class discussion time is limited, or if you want to ensure that students are investigating a specific research question or using preplanned methodology.

For the predefined approach, you can skip Part 0 of the “Ruler Drop Student Handout” document and provide the background, experimental design, and data collection procedure for the class. In this case, consider using the Part 1 questions from the “Customizable Student Handout” document instead.

**PROCEDURE**

This section suggests procedures for implementing the ruler drop activity in class (use links to go directly to each section):

- **Background Section**
- **Part 0: Brainstorming a Research Question**
- **Part 1: Writing the Research Question and Hypothesis**
- **Part 2: Designing the Methods and Setting Up Your Data Sheet**
- **Part 3: Collecting Data and Spreadsheet Entry**
- **Part 4: Comparing Spreadsheets**
- **Part 5: Organizing Data in Spreadsheets**
- **Part 6: Preparing to Combine Data**
  - Data Orientation
- **Part 7: Documenting Your Data**
- **Next Steps**

Note that much of the procedure is similar or identical to that in the general “Educator Materials” document, which you can refer to for more information. Skip or skim repeated sections as needed.

**Background Section**

This section of the handout introduces the principle of reaction time and sets the stage for students to brainstorm ideas for their ruler drop investigation. You could also show students a video clip or demo of the ruler drop procedure to stimulate ideas.

This activity is written using a collaborative approach where the class brainstorms the variables for the investigation. Alternatively, you could **predefine the variables** and make the background more specific.

**PART 0: Brainstorming a Research Question**

Students brainstorm independent variables that could be tested in a class ruler drop investigation. Refer to the “Choosing Independent Variables” section of the “Teaching Tips” for guiding students toward appropriate variables. Alternatively, Part 0 can be modified or skipped if you **predefine the variables** for students.

Note the handout suggests that students work in pairs based on the ruler drop procedure, but you may modify group size based on your class’s needs. Smaller groups are preferred for this part of the activity, as groups are combined for discussion in **Part 4**.

After students have finished brainstorming, guide a discussion to select the variables for the class investigation. Ultimately, the data collected by each group will need to be combined into a class data set with a consistent organization, so all students must agree on the same variables before collecting data.
PART 1: Writing the Research Question and Hypothesis

Students record the variables, research question, and hypothesis for the class investigation. Make sure all students are planning to use the same variables and research question. They may have different hypotheses depending on what they think will happen.

Note that students may pick “reaction time” as the dependent variable to measure. In the default procedure, however, students do not measure reaction time directly. Instead, they measure drop distance and use it to calculate reaction time. You may need to clarify this distinction with students.

PART 2: Designing the Methods and Setting Up Your Data Sheet

The class should discuss and finalize the procedure for recording the independent variables. This is an opportunity to check that all students are clear on the experimental plan before collecting data.

- Make sure the class agrees to use consistent methods/rules for collecting and recording their data. This should include the number of ruler drop trials per person, as well as the specific procedure for dropping and catching the ruler.
- Additional guidance will depend on the independent variable selected. For example, if the class is comparing reactions to visual vs. auditory cues, the class could decide for students to close their eyes and listen for the same auditory signal for catching the dropped ruler.
- As described in the handout, have students practice the procedure a few times prior to recording data.
- It is recommended to not provide guidance on measurement units at this point, so students can compare the units they chose in Part 4.

The handout instructs students to collect their data on a separate sheet first before entering it into a spreadsheet program; for more rationale, refer to the “Recording Data on a Data Sheet” section of the general “Educator Materials” document. Depending on time constraints, classroom design, accessibility, and other considerations, directly entering data into a spreadsheet may be better for your students. If so, adjust the handout to omit the separate data sheet instructions.

PART 3: Collecting Data and Spreadsheet Entry

Note: Instructions for this section are equivalent to those in the general “Educator Materials” document.

Students conduct their investigation and record their data on a blank data sheet (paper or online document). After they have finished collecting data, they will enter it into a spreadsheet program of your choosing. (As a reminder, the default template for the “Student Handout” uses Google Sheets. Edit the instructions as needed if you are using a different program.)

Encourage your students to consider whether they should organize their data differently in the spreadsheet compared to their original data sheet. If you are doing this activity over multiple class periods, the end of Part 3 is a good pause point.

PART 4: Comparing Spreadsheets

Note: Instructions for this section are equivalent to those in the general “Educator Materials” document.

Students will now compare their data sheets with those of another group. Combine student groups as appropriate. For example, student pairs from the previous parts of the activity can be combined into groups of four.
Emphasize that there are no “right” or “wrong” ways for students to have initially organized their data. The purpose of comparing with another group is to learn how others approached the task and to reflect on the elements that make a spreadsheet easy to understand. Highlight the opportunities for students to reflect on their choices.

Students can make minor organizational changes to their data at the end of Part 4, but they should not spend a lot of time reorganizing it. They will decide as a class how to organize everyone’s data later in the activity.

PART 5: Organizing Data in Spreadsheets

Note: Instructions for this section are equivalent to those in the general “Educator Materials” document.

Students will now explore examples of common data organization issues that limit the use of spreadsheet functions (calculations, formulas, sorting, etc.). The “Student Handout” uses examples from a ruler drop experiment that compared reaction time between students’ dominant and nondominant hands.

You can trim or remove examples based on the spreadsheet functions that are relevant for your classroom. If desired, you can more broadly discuss effective practices for spreadsheet organization and/or include more examples not already in the activity.

- For additional examples, refer to Broman and Woo (2018) and the Data Carpentry module “Data Organization in Spreadsheets for Ecologists.”

Students can return to working in pairs or continue to work in their combined groups from Part 4. Alternatively, you can organize students into jigsaw groups, where each student group is assigned one of the four examples (Questions 19, 20, 21, or 24); after each group works through their assigned example, students recombine into mixed groups to share their insights.

Students can work through the data table examples directly in the handout. Or you could have them examine the examples in a spreadsheet program, so that they can experience the issues with the spreadsheet functions directly.

- You can ask students to attempt the tasks described in the handout (calculating means, sorting, etc.) in a spreadsheet program. This may take more time and require students to be more familiar with using spreadsheets.
  - For comparing wide-form and long-form data in Tables 4 and 5, students could compare the ease of calculating averages for each hand treatment.
- You can find spreadsheet files for the ruler drop data tables in the “Materials” box on this resource’s webpage.

Students can make minor organizational changes to their data at the end of Part 5. But, again, they should not spend a lot of time reorganizing it at this point, as they will be making final data organization decisions after a class discussion.

PART 6: Preparing to Combine Data

Note: Instructions for this section are equivalent to those in the general “Educator Materials” document.

Students should now have ideas for how to effectively organize data in a spreadsheet. They will use what they’ve learned to decide on the organization for a class data set, which combines the data from all the groups.
Engage the class in a discussion about how to organize the class data set, following the effective practices that they previously explored. If the data needs to be organized in a certain way for the graphing or analysis tools that your class is using, guide the class discussion accordingly. Make sure to cover the following decisions:

- **How to organize the columns** of the spreadsheet, including what labels/headers to use
- **What units** to use for the measurements. Students may have used different units when recording drop distance and can convert their data by creating a new column and using spreadsheet formulas; show them how to do this if they are unfamiliar.
  - Students first need to convert drop distance to centimeters if they recorded distance in inches or millimeters.
  - The conversion from drop distance in centimeters to reaction time in milliseconds comes from Weber (2017). Students may need assistance to enter the spreadsheet formula.
  - Data unit conversion is not necessary if your class used reaction time rulers that have converted drop distance to time (milliseconds).
- **Whether to orient** the data in long form or wide form (refer to the “Data Orientation” section below for details).

Students should document the column organization, labels, measurement units, etc., that the class decides should be used. Each group should then convert their data to the class organization.

**Data Orientation**

A major decision that the class must make is how to orient the data set. For data sets with multiple groups or treatments, there are two main orientations:

- **Long form**: Also called “tidy data” (Wickham 2014). Each variable forms a column, each observation forms a row, and each cell is a single measurement. In other words, each row contains all the information associated with a set of measurements for each observation (and only one observation). An example of long-form data is provided in Table 1 in the “Student Handout.”

- **Wide form**: Some variables may be “spread out” across multiple columns, where each column represents a different value of that variable (e.g., a different group or treatment). An example of wide-form data is provided in Table 5 in the “Student Handout.” In this case, the “Hand” variable is no longer a single column, but instead reported in both the “Dominant Hand” and “Nondominant Hand” columns.

Orientation can also affect the default arrangement of the variables for different graph types. Different orientations may be favored by different programs or tools.

- Long form is required by many statistical programs and tools, including BioInteractive’s Data Explorer. It is the typical orientation for most large and/or publicly shared data sets.
- Wide form may be easier to interpret when viewing a small table of data. It can be useful for some analyses (e.g., performing summary calculations) depending on the spreadsheet programs or in paired study designs (when repeated measurements are collected on the same individual).

**PART 7: Documenting Your Data**

**Note**: Instructions for this section are equivalent to those in the general “Educator Materials” document.

Students will now learn about documenting data so that it is understandable to others, including those who may want to reproduce their findings or conduct additional analyses. Clear and accurate experimental design, methodology, data recording, and data analysis enable future researchers to replicate a research study and/or apply its approach to a new population. (There may be circumstances that prevent complete reproducibility; for example, if the exact population in which the study was conducted no longer exists.)
Students should document each column of their data sets using additional descriptions, called “metadata.” The “Ruler Drop Student Handout” document shows metadata from a study on the physiological performance of Karoo bush rats. The default “Customizable Student Handout” uses data from a study of finches in the Galápagos (a data set that is also available in Data Explorer).

Next Steps

Note: Instructions for this section are equivalent to those in the general “Educator Materials” document.

At the end of the activity, students should submit their filled-out handouts and finalized spreadsheets according to your directions. You can assess their answers, spreadsheet revisions, metadata, etc., as appropriate for your class.

To create the class data set, you can either:
- Combine all the submitted group data sets into one data set yourself.
- Create a new, central Google Sheet that each group directly adds their data to. In this case:
  - Consider prepopulating the sheet with labels or an example of the final spreadsheet organization as a reminder to students.
  - You may need to check for consistency and make some corrections after all students have entered their data.
  - To prevent accidental changes, you may want to “lock” the spreadsheet once it’s done. Students can copy or download the spreadsheet for individual analyses.

Students can use the class data set to explore whether their hypotheses from Part 1 are supported via summary calculations, graphing, and/or statistical analyses. You may direct them to use additional resources and tools, such as the following:
- The web tool Data Explorer provides a quick and easy way to visualize and analyze data. It includes a variety of options for generating plots and performing statistical tests. If you plan to upload class data, please review the guidelines under “Import your own data” on the landing page of the tool.
- Provide students with scaffolded support for selecting appropriate graph types and statistical tests to visualize their data and test their hypotheses. The “Educator Materials” and “Selection Guides” for Data Explorer can be used for guidance.
- The “Spreadsheet Tutorials” can be used to help students learn about calculations or graphing in Microsoft Excel or Google Sheets.

ASSESSMENT GUIDANCE

The questions and sample answers in this section are from the “Ruler Drop Student Handout” document. Modify them as needed based on any changes made to the handout.

PART 0: Brainstorming a Research Question

1. Describe at least one motivation for studying human reaction time. How might this be relevant to you, the public, or the scientific community?
   Students may discuss how reaction time can help us learn about nervous system function, health, or human performance. Be open to a range of reasonable responses.

2. Which of your independent variables involve comparing different groups (e.g., comparing people of different ages)?
   Students should propose a categorical or group-level variable, which could include grouping by class year,
3. Which of your independent variables involve comparing the effects of a recent activity or action (e.g., comparing how much water each person drank today)?

Students should propose a measure that can be quantified for all members of the class. These can be either categorical or numerical variables.

4. Select one of the independent variables you identified above.
   a. Predict whether and how this variable would affect reaction time.
      Students should state the result that they would expect from an investigation with their selected variable.
   b. Describe why this variable would or would not affect reaction time.
      Students should provide a reason for the effect they described.

PART 1: Writing the Research Question and Hypothesis

5. List the variables that the class will measure.

   Make sure that all students are planning to measure the same variables before they start collecting data.
   a. Independent variable
      Students should list the treatment groups or measure of recent activity they plan to compare.
   b. Dependent variable
      In the default procedure, ruler “drop distance” is what students will measure. (Students may also say that the dependent variable is reaction time. In the default procedure, however, they do not measure reaction time directly. Instead, they measure drop distance and use it to calculate reaction time.)

6. Using these variables, write the research question that the class will investigate.

   Students should write the research question decided by the class discussion. The question should specifically ask about the effect of the independent variable on reaction time (as measured by ruler drop distance).

7. Write a hypothesis for the research question.

   The hypothesis should state the result that the student expects from the investigation.

8. Is your hypothesis a null hypothesis, meaning you do not expect a difference between groups or an effect based on the independent variable? Or is it an alternative hypothesis, meaning you do expect a difference or effect?

   Student answers will vary based on their response to Question 7.

PART 2: Designing the Methods and Setting Up Your Data Sheet

9. How will the class measure and record the independent variables selected in Part 1?

   Students should describe the methodology decided by the class discussion.

10. How many trials will each person do?

    Students should describe the methodology decided by the class discussion.

PART 3: Collecting Data and Spreadsheet Entry

11. What columns does your spreadsheet have?

12. What does each row in your spreadsheet represent?

   Students will need to record observations for ruler drop distances, for each person, for multiple trials.
they organize these observations across the rows and columns of their spreadsheet will likely vary at this point.

13. Why did you decide to arrange the data in your spreadsheet in this way?

*Students should be able to justify their spreadsheet organization. You should expect that spreadsheet organization will vary across student groups and emphasize that there is no “right” way to organize a spreadsheet at this point. (Later in the activity, students will modify their spreadsheets after iterative experiences with their peers, including self-reflection and exercises designed to help them discover and correct misconceptions.)*

14. What other information, if any, is needed to make your spreadsheet understandable to someone unfamiliar with the investigation?

*Students may identify variables that need additional labels or other information.*

**PART 4: Comparing Spreadsheets**

15. Consider the layout and variables in your spreadsheet and the other group’s.
   a. What are some similarities between your spreadsheets?
   b. What are some differences?

   *Similarities or differences may include the organization of the data (the placement of variables in rows and columns), the information recorded in each cell, which units are used, and the labels selected.*

16. What features of the other group’s spreadsheet help you understand the data they collected?

   *Students may identify structures or labeling used by the other group that helped explain the variables that were measured and what data were collected.*

17. Consider the units used for your measurements.
   a. What units did you use to measure drop distance?
   b. What units did the other group use?

   *For both Parts a and b, students may have measured drop distance in inches, centimeters, or millimeters.*

   c. Were the units originally labeled on your spreadsheets, and if so, where did each group place their unit labels?

   *Students should report if and where units were labeled on their spreadsheets.*

18. Data sets can combine data from multiple sources. If you were to combine your data with the other group’s, how might you change the organization or structure of your spreadsheet?

   *Student answers will vary. Note that students are only being asked to reflect on potential changes at this point. They should not make any major changes to their spreadsheets until later in the activity.*

**PART 5: Organizing Data in Spreadsheets**

*Note:* Questions and answers for this section are the same as those in the general “Educator Materials” document.

19. Table 2 shows how the first student group organized the data.

   **Table 2.** First example of how students organized the data in Table 1.

<table>
<thead>
<tr>
<th>Column A</th>
<th>Column B</th>
<th>Column C</th>
<th>Column D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Student</td>
<td>Hand</td>
<td>Time (ms)</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>Dominant</td>
<td>30.7, 17.7</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>Nondominant</td>
<td>35.8, 34.4</td>
</tr>
</tbody>
</table>
The students tried to calculate the mean (average) distance for each type of hand using the AVERAGE function in a spreadsheet program. This function needs each cell to contain an individual numerical value.

a. The AVERAGE function didn’t work with Table 2. Why do you think this is?

*Both the “Distance” and “Time” columns have multiple measurements in each cell. The AVERAGE function doesn’t work because it needs each cell to contain an individual number.*

b. Describe the data organization practice used by Table 1, but not by Table 2, that allows the AVERAGE function to work with these data.

*The data organization practice is listing only one measurement/number in each cell. Table 1 does this by adding a “Trial” column that allows the numbers in the “Distance” and “Time” measurements to be separated into single cells in individual rows.*

20. Table 3 shows how the second student group organized the data.

**Table 3.** Second example of how students organized the data in Table 1.

<table>
<thead>
<tr>
<th>Column A</th>
<th>Column B</th>
<th>Column C</th>
<th>Column D</th>
<th>Column E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Student</td>
<td>Hand</td>
<td>Trial</td>
<td>Distance</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>Dominant</td>
<td>1</td>
<td>30.7cm</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>Dominant</td>
<td>2</td>
<td>17.7cm</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>Nondominant</td>
<td>1</td>
<td>35.8cm</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>Nondominant</td>
<td>2</td>
<td>34.4cm</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>Dominant</td>
<td>1</td>
<td>11.0cm</td>
</tr>
<tr>
<td>7</td>
<td>B</td>
<td>Dominant</td>
<td>2</td>
<td>14.2cm</td>
</tr>
<tr>
<td>8</td>
<td>B</td>
<td>Nondominant</td>
<td>1</td>
<td>47.1cm</td>
</tr>
<tr>
<td>9</td>
<td>B</td>
<td>Nondominant</td>
<td>2</td>
<td>15.9cm</td>
</tr>
</tbody>
</table>

The students tried to calculate the mean (average) for all the time measurements using the AVERAGE function. Again, this function needs each cell to contain an individual numerical value.

a. The AVERAGE function didn’t work with Table 3. Why do you think this is?

*Both the “Distance” and “Time” columns have the units written in text after each number. The AVERAGE function won’t work because it needs cells with just numbers. (In general, text in cells may prevent a spreadsheet program from recognizing numerical data in the cells.)*

b. Describe the data organization practice used by Table 1, but not by Table 3, that allows the AVERAGE function to work with these data.

*The data organization practice is removing text after the numbers in cells. Table 1 does this by removing the units in the cells and writing the units in the column headings instead.*

21. Table 4 shows how the third student group organized the data.

**Table 4.** Third example of how students organized the data in Table 1.

<table>
<thead>
<tr>
<th>Column A</th>
<th>Column B</th>
<th>Column C</th>
<th>Column D</th>
<th>Column E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Student</td>
<td>Hand</td>
<td>Trial</td>
<td>Distance (cm)</td>
</tr>
</tbody>
</table>
2 A Dominant 1 30.7 250.2
3 A Dominant 2 17.7 190.0
4 A Nondominant 1 35.8 270.2
5 A Nondominant 2 34.4 264.8
6 B SecondHand: Dom 1 11.0 149.8
7 B SecondHand: Dom 2 14.2 170.1
8 B FirstHand: Nondom 1 47.1 309.9
9 B FirstHand: Nondom 2 15.9 180.0

The students wanted to group the data based on “Hand” type (Column B) for additional calculations. They wanted all the rows of dominant hand data grouped together, followed by all the rows of nondominant hand data grouped together.

To reorder the data, they used the SORT function in a spreadsheet program, which sorts rows by the values in a specific column (in this case, the “Hand” column). The function sorts from smallest to largest values for numerical variables or alphabetically (A to Z) for text variables.

a. The SORT function didn’t group the data in Table 4 the way the students wanted. Why do you think this is?

*The treatments in the “Hand” column are named inconsistently. For example, the dominant hand treatment is called both “Dominant Hand” and “SecondHand: Dom.” When sorted alphabetically, the treatments will not be grouped together because some labels have different starting letters.*

b. Describe the data organization practice used by Table 1, but not by Table 4, that allows the SORT function to group these data the way the students wanted.

*The data organization practice is naming groups/treatments consistently. Table 1 does this by using the same names (either “Dominant” or “Nondominant”) for each “Hand” treatment.*

22. Now that you’ve identified several beneficial data organization practices, what other changes (if any) would you make to your own spreadsheet and why?

*Student answers will vary based on the initial layout of their data.*

23. What is the main difference between these two orientations?

*The wide-form example (Table 4) places each trial in its own row, with separate columns for the measurements of each treatment. So the “Dominant” and “Nondominant” treatments each have their own “Distance” and “Time” columns.*

*The long-form example (Table 5) places each observation in its own row, with the treatment being listed in its own column: the “Hand” column (recorded as either “Dominant” or “Nondominant”). Thus, all the “Distance” measurements are in a single column, as are all the “Time” measurements.*

24. The orientation of the data may affect certain spreadsheet functions, such as calculating new variables or performing summary calculations. Consider the following examples based on the data in Tables 4 and 5.

a. A student wants to convert the time measurements from milliseconds (ms) to seconds (s) in a new column in the spreadsheet. Would this task be easier in Table 4 (wide form) or Table 5 (long form), and why?
It would probably be easier in Table 5. If the student uses Table 4, they would need to create two new columns: one to convert the “Dominant Time” column and another to convert the “Nondominant Time” column. If the student uses Table 5, they only need to create one new column to convert the “Time” column.

(In general, long-form orientation is easier for calculations that need to be performed for each observation.)

b. A student needs to calculate the average time for each “Hand” treatment in a new row in the spreadsheet. Would this task be easier in Table 4 (wide form) or Table 5 (long form), and why? It would probably be easier in Table 4. The student could simply add a new row at the bottom of Table 4 to calculate an average for each column, including the average time for each “Hand” treatment. Calculating the average time for each treatment in Table 5 would be more difficult, since the student would have to sort the data by “Hand” treatment and specify specific ranges of data for calculating the averages.

(In general, wide-form orientation is easier for doing summary calculations, such as averages, per treatment.)

25. Return to the spreadsheet that you made earlier in the activity. Did you originally enter your data in long form or wide form?

Student answers will vary based on the initial layout of their data. The “Data Orientation” section of this document provides more considerations that you may want to discuss with your class.

PART 6: Preparing to Combine Data

26. What columns will the class data set have? List the headers/labels for each column.

27. Which orientation (long form or wide form) will the class data set use?

Answers for both questions will depend on the data organization you decide. Make sure all students organize their data in the same way before combining into the class data set.

28. Based on this information, write a spreadsheet formula that can be used to:

a. convert inches to centimeters

\[ \text{[cell with the distance in inches]} \times 2.54 \]

b. convert millimeters to centimeters

\[ \text{[cell with the distance in millimeters]} / 10 \]

29. Based on this information, write a spreadsheet formula that will convert drop distance in cm to reaction time in ms. (The spreadsheet function for the square root is SQRT.)

\[ \text{SQRT([cell with the distance in cm]/490.5) * 1000} \]

30. Investigations with human subjects often anonymize their data. Why do you think researchers anonymize data collected from humans?

Students may describe privacy concerns related to the collection of personal information from human participants in a research study. Answers could also address the potential for researcher bias if identifying information about participants is known when analyzing the data.

PART 7: Documenting Your Data

31. Why is it important for scientific findings to be reproducible, meaning that similar results are found by different people? Give a specific example where it’s important to reproduce the result of a study.

Reproducing the results of a study helps validate those results. This could be important for confirming
high-stakes results, such as the outcome of a medical study that will affect the development of medicines and policies. It can also help check whether similar results still apply under different conditions or in different populations.

32. Based on just Table 6:
   a. What can you tell about this study so far?
      Students may infer that the study measured characteristics of the rats, including their sex, size ("Mass"), activity level ("Time moving"), and how fast they move ("Max speed").
   b. What seems confusing or requires more information to understand?
      Students may point out that there is no information about the units of the measurements or how/where the data were collected.

33. Based on Table 7:
   a. What did you learn from Table 7 that was not clear from Table 6 alone?
      Students may report that the metadata clarified how measurements were taken and which units were used, as well as provided more background on the study (e.g., dates of data collection and who collected the data).
   b. Why do you think metadata is important to include?
      The metadata is important for anyone who is trying to understand or replicate the study. It allows the people who did the study to explain what they did in more detail without making the data set labels too long or complicated.

REFERENCES


CREDITS

Written by Kaitlin Bonner, St. John Fisher University, NY; Kristine Grayson, University of Richmond, VA; Esther Shyu, HHMI

Ruler drop illustration adapted from Biologycorner.com, used under [CC BY-NC 3.0](https://creativecommons.org/licenses/by-nc/3.0)